On the interaction of information structure and prosody: the case of Egyptian Arabic

HABILITATIONSSCHRIFT

Graz, im Dezember 2013
# Contents

**Preface** ........................................................................................................................................... vii  
**Transcription, glossing conventions and abbreviations** ...................................................................... ix

## 1 Introduction ........................................................................................................................................ 1

1.1 Information structure ......................................................................................................................... 1  
   1.1.1 Pragmatic properties .................................................................................................................. 2  
      1.1.1.1 Information status or the activation state of concepts ......................................................... 3  
      1.1.1.2 The focus-background structure or the speaker’s choice ................................................... 8  
   1.1.2 Pragmatic relations ....................................................................................................................... 16  
      1.1.2.1 FOCUS, presupposition and assertion ............................................................................... 17  
      1.1.2.2 FOCUS vs. focus ................................................................................................................ 19  
   1.1.3 Theticity ...................................................................................................................................... 23  
   1.1.4 Topic, TOPIC and relational presupposition ........................................................................... 25  

1.2 Objective and general outline of this study ......................................................................................... 27

## 2 Prosody ................................................................................................................................................. 29

2.1 Intonation: rhythm and tonal contour .................................................................................................. 30  
   2.1.1 Tune-text association .................................................................................................................. 31  
   2.1.2 Rhythm and accent ..................................................................................................................... 36  
      2.1.2.1 Rhythm in a metrical framework .......................................................................................... 38  
   2.1.3 Excursus: accent and stress ........................................................................................................ 42  
   2.1.4 Accentuation ............................................................................................................................... 45  
   2.1.5 The tonal component ................................................................................................................... 50  
      2.1.5.1 Rises, falls and the biological codes ..................................................................................... 51  
      2.1.5.2 The representation of intonation proper according to AM theory ..................................... 54  

2.2 The Intonation of Egyptian Arabic ..................................................................................................... 58  
   2.2.1 A brief overview of AM approaches to Egyptian Arabic intonation ......................................... 58  
   2.2.2 A functional model of EA intonation ........................................................................................... 60  
      2.2.2.1 The metrical level ................................................................................................................ 62  
      2.2.2.2 The tonal component .......................................................................................................... 63  
   2.2.3 The default pitch accent ............................................................................................................... 64  
   2.2.4 From tones to tunes: accent modification and accent sequencing ............................................ 71  
      2.2.4.1 Meaningful accent modification .......................................................................................... 72  
      2.2.4.1.1 Alignment and its transcription ......................................................................................... 74  
      2.2.4.1.2 Scaling ............................................................................................................................... 78  
      2.2.4.1.3 Downdrift and downstep ................................................................................................ 82  
      2.2.4.1.4 Integration and separation ............................................................................................... 85  
   2.2.5 The meaning of tones and tunes and their linguistic functions .................................................... 86
3. On the interaction of prosody and information structure ........................................ 113
   3.1 Accent as a marker of focus .................................................................................. 116
      3.1.1 Terminological problems: accent and focus, two names for the same thing or two different concepts? ................................................................. 116
      3.1.2 The focus-to-accent conundrum .................................................................. 117
      3.1.3 ‘Broad’ vs. ‘narrow’ focus .......................................................................... 118
      3.1.4 ‘Structure based’ versus ‘highlighting-based’ focus-to-accent approaches .......... 120
         3.1.4.1 Bolinger: “Accent is predictable (if you’re a mind reader)” ................ 121
         3.1.4.2 Halliday: Information structure is realized by intonation ...................... 122
         3.1.4.3 Chafe: Consciousness and information flow ........................................... 123
         3.1.4.4 Schmerling: Topic-comment structure .................................................. 125
         3.1.4.5 Ladd: ‘Normal’ stress as a signal of broad focus’ and the notion of ‘deaccenting’ .... 128
      3.1.4.6 Gussenhoven: Focus domains over semantic constituents and rule-governed accent placement ........................................................................ 131
         3.1.4.7 Selkirk: A pitch-accent-first theory and the projection of focus domains .... 134
         3.1.4.8 Lambrecht: A unified functional account of sentence accentuation .......... 136
      3.1.5 Summary and conclusion ............................................................................. 139
   3.2 Tonal marking of topic and focus ...................................................................... 141
      3.2.1 Bolinger: thematic and rhematic accents .................................................... 142
      3.2.2 Brazil: referring and proclaiming tones ....................................................... 143
      3.2.3 Gussenhoven: addition, selection and relevance testing .............................. 144
   3.3 A model of pragmaprosody ................................................................................ 146
      3.3.1 Prosodic and pragmatic preferences ............................................................ 146
      3.3.2 Prosodic preferences .................................................................................... 147
      3.3.3 Pragmatic preferences .................................................................................. 148
      3.3.4 The interaction of prosodic and pragmatic preferences .............................. 149

4 The prosody of information structure in Egyptian Arabic: pragmatic properties and pragmatic relations ........................................ 153
   4.1 Information structure in (Egyptian) Arabic: the state of the art ....................... 153
   4.2 The corpus: methods and data ......................................................................... 155
   4.3 The prosody of pragmatic properties .............................................................. 157
5 Topicality ................................................................. 261
  5.1 An excursus on word order ........................................... 263
  5.2 Topics and TOPICS .................................................. 271
  5.3 Thematic TOPICS .................................................... 277
    5.3.1 Entity topics as TOPICS ........................................ 279
      5.3.1.1 Subject-predicate constructions .............................. 284
      5.3.1.2 Packaging TOPICS: scrambling, left dislocation and pseudoclefts .................................... 288
      5.3.1.3 Reference point constructions and multiple topics ........................................... 298
      5.3.1.4 Free TOPICS .................................................. 300
      5.3.1.5 Pronominal TOPICS .......................................... 303
    5.3.2 Relational presupposition as TOPIC ............................ 308
    5.3.3 Some figures .................................................... 315
    5.3.4 Frame TOPICS .................................................... 317
      5.3.4.1 Frames with leading and closing final accents ................................. 318
      5.3.4.2 Temporal frames .............................................. 322
      5.3.4.3 Spatial frames ................................................ 325
      5.3.4.4 IF-TOPICS ..................................................... 327
  5.4 Discourse markers in TOPIC position ............................. 330
  5.5 Morpho-lexical topic markers ...................................... 333
Preface

Dwight Bolinger once characterized intonation as being “around the edge of language”. This is true for prosody in general and also of information structure. Prosody, as I understand it, relies first and foremost on universal natural processes. However, the prosodies of individual languages are not alike. That is, every language has its own way of adapting these universal processes to its overall phonological and grammatical structure. Concerning information structure, the position taken in this study is that information structural notions are in the first place ‘comparative concepts’ in the sense of Haspelmath (2010) and not necessarily universal cross-linguistic categories. Having a general cognitive basis, they are, however, useful starting points for cross-linguistic comparison. In line with Matić & Wedgwood (2012) I do not assume that there are universal linguistic categories of ‘topic’ and ‘focus’ that necessarily have a certain formal expression in every language, but rather that formal expressions may have a certain pragmatic (i.e. information structural) interpretation. Consequently, I do not assume a one-to-one relationship between information structural concepts and linguistic categories or constructions. This is perhaps even more true for prosody than for other linguistic levels. Thus, I will not be looking for how potential categories such as ‘contrastive focus’, ‘contrastive topic’, or ‘information focus’ etc. are expressed in Egyptian Arabic, but rather expect that many of these are conversational implicatures triggered by specific linguistic structures. As a corollary, the best we can hope to find looking for the relationship between information structural notions and prosodic constructions are statistical tendencies in the preferred use of certain forms to achieve certain pragmatic effects, but not systematic form-meaning relations as in the case of genuine grammatical categories.

In the present study, I have tried to fill a gap in Arabic language studies tackling two understudied linguistic phenomena in Arabic varieties: prosody and information structure. The study will therefore be of interest to students of Arabic who want to know more about the prosodic structures and their informational interpretations in Egyptian Arabic. Those who are primarily interested in the corpus study in Chapters 4-6 might want to skip some of the theoretical argumentation in Chapters 2 and 3 that provide a theoretical discussion of prosody in general and the intonation of Egyptian Arabic in particular (Chapter 2) and a synopsis of approaches to the information structure-prosody interface as well as the outline of a model of Pragmaprosody that serves as the conceptual framework for the empirical part (Chapter 3). Furthermore, the different notions of information structure will be presented in Chapter 1 and thoroughly discussed in the empirical chapters. Thus, the study may also be of interest to general linguists working on information structure and/or prosody.
Transcription, glossing conventions and abbreviations

The Egyptian Arabic examples are transcribed in a broad transcription based on the guidelines of the International Phonetic Alphabet (IPA). The so-called ‘emphatic’ consonants are transcribed as pharyngalized. Any additional velarization is not marked. In glossing examples, the following conventions obtain: affixes are separated from the stem by hyphens, clitics by an equal sign (=). In general, a simplified type of glossing is used, i.e. morphological analysis is not always indicated in the transcription. For instance, active participles are not split up into prefix and stem in the transcription, and in the glossing the stem is separated from the sign for the grammatical category by a period ([stem].PTCP), passive participles are not glossed as such at all, but are translated with English adjectives. To make the glosses less complex, tense/aspect is not glossed, as it is unambiguously marked in the verbal paradigm, prefixes indicating imperfect(ive) and suffixes indicating perfect(ive). Capital letters in the transcription indicate accentuation. Occasionally small caps are used to indicate accents with a low prominence.

The following abbreviations and symbols are used in this study:

1  first person
2  second person
3  third person
ACC  accusative; accessible
AM  autosegmental-metrical
C  consonant
CA  Classical Arabic
COLL collective
COMP complementizer
COP copula
COV coverb
dB  decibel
DEF  definite
DEM demonstrative
DIM diminutive
DM demonstrative marker
DU dual
DUR  durative
EA Egyptian Arabic
F feminine
F0  fundamental frequency
foc focus
FOC FOCUS
FUT  future
GEN genitive
GIV given
H  high tone
HAB habitual
HESIT hesitation (pause)
HON  honorific
Hz   Hertz
IMP  imperative
IN   instance noun
IND  indicative
INDF indefinite
INTENS intensive
IP   intonation phrase
L    low tone
LOC  locative
M    masculine
ms   millisecond
MSA  Modern Standard Arabic
NEG  negation, negative
NEW  new
NOM  nominative
OBJ  object
PASS passive
PL   plural
POSS possessive
PP   prosodic phrase
PRON pronoun
PTCP active participle
PRT  particle
Q    question particle
REL  relative
SBJ  subject
SEQ  sequential
SG   singular
ST   semitone
SYLL syllable
top  topic
TOP  TOPIC
V    vowel
VOC  vocative

Transcription of tones:
L    low tone
H    hight tone
^    higher
^    lower
°    same height
>    later
<    earlier
!    downstep
%    boundary tone
-    tone spreading

x
1 Introduction

1.1 Information structure

*Information structure* is a universal phenomenon of languages that serves the communicative needs of interlocutors to structure their utterances in a way that facilitates the cognitive processing of the speech flow (or a written text). Information structure puts parts of what is said or written into the foreground and other parts into the background and establishes relations between individual referents and propositions and also between the propositions themselves. Thus, information structure adds to the establishment of coherence in discourse. Optimally structured information thus helps the speaker to get his message across and makes the task of understanding easier.

Within the past decades, information structure has attracted growing interest among linguists working within different theoretical frameworks. Of course, the field of research is vast and covers many different aspects and subtopics. Ever since the beginning interest in the subject from the viewpoint of modern linguistic theories (Prague School, Halliday, Chafe), the basic notions have been defined and re-defined time and again. The term *information structure* was introduced by Halliday (1967). Although the idea that the syntactic structure of a sentence is influenced by the 'informational structure' of the message goes back to Aristotle and was already a subject of scholarly interest in the early days of modern linguistics in the work of Georg von der Gabelentz (1869), Hermann Paul (1880/1920), (Ammann 1928/1974) and Vilém Mathesius (1929) (von Heusinger 1999), it was Halliday (1967) who first assumed information structure to be an independent linguistic level, orthogonal to syntax\(^1\). Another term for the phenomenon at issue is *information packaging*. This term was coined by Chafe (1976); it metaphorically expresses the fundamental assumption that a message may be conveyed in different ways that correspond to the communicative needs of the situation and the interlocutors according to their temporary state of mind without the truth value of the information being affected. Thus Chafe's approach is a genuinely pragmatic one, which is also the approach adopted in the present study. It has been variously pointed out that some types of information structuring may lead to truth-conditional differences (cf. Krifka 2007) and there is an independent line of research treating (aspects of) information structure as an integral part of semantics (cf. von Heusinger 1999), especially the theory of *Alternative Semantics* (Rooth 1985, 1992) or the *Structured Meaning* approach to focus (Jacobs 1983; von Stechow 1991; Krifka 1992).

The current study stands in the tradition of what we might call functional pragmatic approaches, as represented by Chafe (1974, 1976, 1994) and more recent theoretical developments, such as e.g.

---

\(^1\) Cf. von Heusinger (1999) for an informative overview over the history of research on information structure.
Lambrecht (1994) and Vallduví (1992). On this view, information structure has many interactions with other levels of grammar. The present study will however be confined to the interaction of information structure and prosody. To give an account of the various developments and approaches that have emerged in the field of information structure would certainly go beyond the scope of this chapter; the aim of the following sections will rather be to concentrate on the information structural notions relevant for the present study and define them accordingly. Evidence for the choice of the categories employed and their definitions comes from the analysis of the corpus data in Chapter 4 to 6. As a matter of fact, a major distinction will be drawn between what Lambrecht (1994) called pragmatic properties and pragmatic relations, which shall be dealt with in turn in subsections 1.1.1 and 1.1.2.

1.1.1 Pragmatic properties

One major aspect of communication is the exchange of information. Communication thus involves a constant update of the speaker's and the hearer's shared knowledge or the Common Ground (Stalnaker 1974; cf. Krifka 2007). The content of the Common Ground consists of background knowledge propositions usually referred to as given information. In recent work, the computer metaphor of files and file-cards (Heim 1982) is frequently used to represent the organization of the already given information in the hearer's mind. It is assumed that the file structure is updated during conversation and information is either attached to existing file-cards or a new file-card is created that represents the new information. As Chafe (1974, 1976, 1987, 1994) has pointed out, what is assumed to be given or in the common ground between speaker and hearer is related to the hearer's consciousness or rather the speaker's assumptions about what is currently 'lit up' in the hearer's consciousness. There is ample evidence from the languages of the world that givenness, or more generally, the cognitive status or information status of specific referents or concepts may be marked linguistically, the formal correlates being essentially phonological or morpho-lexical in nature. As it seems, syntax plays a minor role in coding information status in the interlocutors' consciousness (Lambrecht 1994: 95). The phonological, or more specifically the prosodic, formal correlate of givenness are lack of prominence, mostly referred to as the deaccenting of expressions used for given denotata, which is common in some languages, for instance English (Ladd 1980, Cruttenden 2006), and pronominal coding as shown in example (1).

(1) John will not be able to attend the meeting. He is ill.

The pronoun he in the second sentence is referring back to the subject of the first sentence which can be assumed to be at the forefront of the hearer's mind at the time it is mentioned. This awareness on the side of the hearer justifies the use of a pronoun as well as the lack of prominence on that item.

The frequently referred to dichotomy of given and new, however, has also been used in another sense, referring to different, albeit interrelated concept. Baumann (2006: 35f.) identifies three major...
dimensions that have been referred to in the literature by the given/new dichotomy: (i) theme and rheme (or topic and comment) – which will be dealt with in the next subsection under the heading of pragmatic relations, ii) focus and background and finally iii) given and new in the sense of cognitive availability as referred to above. In this subsection we will take a closer look at the last two dimensions (ii and iii) and see how they fit in Lambrecht's concept of pragmatic properties.

As the present state of research on the prosody – information structure interface suggests, an adequate description of the prosodic correlates of information structure makes it necessary to distinguish between at least two types of information structural categories, one that refers to the information status of referents, i.e. of their (degree of) givenness (a notion that will have to be refined below) and the much investigated issue of marking of focus via (pitch) accent. Such a division is constitutive for Lambrecht's pragmatic theory. Lambrecht (1994: 112) proposes two distinct functional categories he calls pragmatic properties and pragmatic relations. He bases the distinction between these categories on prosodic grounds, noting that

[i]t is necessary [...] to postulate two functionally (though not necessarily phonetically) distinct types of prosodic contrast, one expressing differences in the activation states of referents, the other differences of another kind. I will argue that the existence of these functionally different types of accent correlates with the existence of two different types of information-structure categories: those indicating temporary cognitive states of discourse referents (the categories of activation and identifiability) and those indicating relations between referents and propositions. (Lambrecht 1994: 112f.)

With activation and identifiability Lambrecht refers to the two dimensions of information status, i.e. whether or not a concept is currently 'lit up' or activated in the hearer's mind or, as a prerequisite of the latter, whether it is identifiable at all. His term relations, on the other hand, refers to what has been mentioned above under (i), the theme-rheme (or topic-comment or topic-focus) dichotomy. In the present model, I make use of the basic distinction between pragmatic properties and pragmatic relations, I shall however redefine the notions according to the needs of the present study by re-establishing a differentiation between Theme-Rheme (i) and focus-background (ii) that have been lumped together in Lambrecht's model.

1.1.1.1 Information status or the activation state of concepts

Givenness or the information status of concepts has extensively been dealt with by Chafe (1974, 1976, 1987, 1994) who emphasizes the role consciousness plays in the way information is linguistically transmitted. He clearly points out that knowing something is a prerequisite for, but not the same as, being aware of it. Chafe defines given information as "that which the speaker assumes to be already

---

2 In section 1.1.2 below, the notion of FOCUS will be introduced and distinguished from focus. FOCUS corresponds to Lambrecht's notion of focus.

3 Chapter 3 will provide a review of the literature dealing with accentuation as a focus marking device.
present in the addressee's consciousness at the time of an utterance" (1974: 111). To be present in the interlocutor's consciousness means to be activated. Chafe (1987, 1994: 53) distinguishes between three activation states: active, semi-active and inactive. An active idea is salient in a hearer's consciousness or currently “lit-up”, while a semi-active one is only in its peripheral consciousness. Chafe (1994: 71ff.) relates these three cognitive dimensions to the more commonly used terminology of given-new, relating active state to given information and inactive state to new information, whereas ideas that are semi-active he calls accessible (Chafe 1994:72). Distinguishing between the speaker's and the hearer's perspectives, Chafe thinks of raising an idea to full consciousness, i.e. making it active also in the interlocutor's mind, as a process that involves effort on the part of the speaker. This effort he refers to as activation cost.

We have seen in the preceding paragraphs that givenness as a cognitive notion in the sense of the third dichotomy referred to above under (iii) is itself composed of two dimensions, namely activation state, i.e. to what extent the addressee is assumed by the speaker to be conscious of something at the time of the utterance, and identifiability, i.e. whether the addressee is assumed by the speaker to have previous knowledge of the entity referred to at all. This fact is also reflected in the lengthy quotation from Lambrecht cited above. Lambrecht integrates Chafe's idea of the separation between activation state and previous knowledge in a model of identifiability mainly based on Prince (1981).

In her seminal paper (Prince 1981), Ellen Prince devised a one-dimensional familiarity scale based on whether a certain piece of information is discourse-old/new or hearer-old/new, thereby integrating both dimensions: the knowledge or identifiability dimension and the activation dimension. The assumptions a speaker makes about a hearer's familiarity with a certain referent range from "evoked" (completely given), i.e. both discourse-old and hearer-old, over "inferrable", i.e. hearer-old, but discourse-new to "new", i.e. both discourse-new and hearer-new. Although Prince's taxonomy is more fine-grained, a tripartite classification will be sufficient for the present purpose. Lambrecht (1994: ch.3) takes the ideas presented in Chafe's and Prince's works as a basis for his integrated model that assumes an implicational hierarchy between identifiability and activation state. The following diagram is reproduced from Lambrecht (1994: 109, ex. 3.25).

As Lambrecht convincingly argues, a concept may only be activated if it is identifiable in the first place. Identifiability is thus a prerequisite for activation, a fact that is represented by the arrow in the
An unidentifiable concept is necessarily new, whereas identifiable concepts may be active, such as the referent of 'he' in example (1) above, inactive or accessible. To illustrate these three different activation states or information states let us have a look at a short text, also adapted from an example cited by Lambrecht (1994: 110).

(3) I heard something TERRIBLE last night. (Ø) remember MARK, the guy we went HIKING with (Ø)? His SISTER just died of CANCER.

The underlined expressions in the above example denote discourse referents and the capital letters represent accented or prominent items. For the time being, we will ignore the information status of other concepts, such as those denoted by verbs or adverbials and the like, but we will return to the issue of discourse referents and other concepts shortly. From the underlined expressions, I, Ø, we and his clearly refer to active or given referents. The two pronouns refer to the two interlocutors. From a syntactic (not from an informational) point of view, we find two gaps in the utterance, the first one of which also relates to one of the interlocutors, as the first elided element would have to be ‘you’, while the second one is referring back to a third participant Mark who has just been introduced to the discourse. The same is true for the possessive pronoun his. As regards the lexically coded and accented expressions, something terrible is clearly unidentifiable and therefore new information. Mark and cancer, on the other hand, are identifiable, but assumed not to be immediately at the forefront of the addressee's mind. Thus, they can and have to be activated. Mark, being a proper name, evidently points to a specific person known to the speaker and the hearer, while his sister is probably unknown to the hearer, in which case she would not be identifiable. That Mark should have a sister is nothing uncommon in itself. In her capacity of being a sister of Mark's, the person referred to becomes at least identifiable to some degree - even though not uniquely identifiable - in the map of the hearer's discourse world. Let us suppose Mark's sister's name is Lisa. If the speaker had constructed the information with Lisa as the subject of the sentence, the addressee would not have been able to identify the person the speaker had in mind. The speaker therefore proceeded gradually to the main information he wanted to transmit. He reminded the addressee of Mark first, thereby referring to the common ground between them. Only in a second step, he transmitted the information using the expression his sister as the basis for the predication, thereby anchoring the hitherto unidentifiable referent in the hearer's discourse model. Cancer, being a common disease, can also be assumed to be identifiable to the hearer and, as long as it has not been mentioned at a previous time in the discourse between the two interlocutors, it may safely be presumed to be inactive knowledge.

In a Chafean sense, Mark and cancer would have to be judged as identifiable, but inactive, while his sister would be new information, but still be accessible as Mark's sister. Even if the hearer does not have any prior knowledge of Mark having a sister, he will normally accommodate the existence of a sister of Mark's – upon hearing the possessive noun phrase his sister. This example shows that accessibility is attributable to informationally heterogeneous referents or concepts and that it is not
covered by Chafe's concept of activation alone. It makes clear that accessibility is not always directly related to activation from a previously conscious state, whereas thinking of a concept as active or inactive necessarily is a matter of consciousness. Lambrecht (1994: 104; emphasis in the orig.) suggests to think of "cognitive accessibility as a potential for activation rather than as the state of a referent in a person's mind." This entails that the term accessible covers a wider range of cognitive states than the term semi-active for which Chafe seems to use it quasi-synonymously.\(^4\) On the other hand, it makes sense to think of accessibility as a cognitive prerequisite for informational structuring of propositions, which shall be extensively dealt with in the next subsection under the heading of pragmatic relations.

But before we turn to these, I will briefly touch upon the ways an idea or a referent may be accessible to a hearer. As already noted by Halliday, there are different sources that may make a concept "recoverable" to the hearer. In Halliday's theory of information structure given is defined ex negativo as the optional remainder within an information unit after the identification of what is new or focal. For Halliday (1967: 211) the given parts are those which are "recoverable anaphorically or situationally". That is, they may either have been mentioned before within the same discourse or present in the text-external world. As far as anaphorically (i.e. textually) recoverable concepts are concerned, a certain time factor has to be taken into account. Evidently, a concept will not stay at the forefront of a hearer's mind for an infinite amount of time. There is no definite answer, how long a concept may stay 'lit-up' in the mind of a listener, nor is it clear how the distance between first and second mention should be measured (see Givón 1983 for a proposal). Chafe (1976: 33) suggests that the effect of discourse boundaries such as a change of scene would be a more promising criterion for measuring the persistence of activeness than for example the number of intervening sentences. The question how long a concept can count as completely active and therefore given and when it will be assumed to have dropped to peripheral consciousness thus remains unsolved. But one thing seems to be clear: a concept that has already been mentioned at any time in a specific discourse may safely be regarded as identifiable and accessible. Chafe (1976: 40) gives the example of a letter that was first mentioned on page 13 of a novel and referred to as 'the letter' again on page 118, showing that definiteness as a formal expression of identifiability may last over 105 pages of written text.

There is yet a third source from which a hearer may retrieve a referent or a concept in general. A hearer may be able to make inferences from discourse internal and discourse external information, invoking a cognitive "schema" as proposed by Chafe (1987: 29). Prince (1981) also refers to the 'in-

\(^4\) It is not entirely clear to me whether Chafe really assumes the concepts to be synonymous. In his monograph of 1994 he relates two different taxonomies to each other, mapping active – semi-active – inactive on the one hand to given – accessible – new on the other hand (cf. 1994: 71-81). Although Chafe is not exceedingly explicit about the classificatory distinction between these two lines, he seems to be thinking about the first one as pertaining to the cognitive state of concepts and of the second line in terms of the activation cost required to make a concept ultimately active in the hearer's mind. If this interpretation is correct, Chafe may have had in mind a similar distinction as the one I will be proposing in the present book.
between' category (between completely given and new) in her familiarity scale as "inferrables", mentioning the example of a driver being inferable from the prior mention of a bus. Chafe (1987: 29f.) gives the example of a number of concepts that have been made "semi-active" or "accessible" by evoking the schema of an ‘undergraduate class’, such as ‘student’, ‘instructor’, ‘teaching assistants’ and ‘a classroom’.

The main interest of the linguist, of course, is how the identified cognitive categories are expressed linguistically, or whether they have any linguistic correlates at all, so that it may be justified to assume a linguistic category in the first place. I have already made reference to the formal correlates of identifiability and activation in various places in the text. At this point I will give a brief overview of the correlations between these cognitive categories and their most important formal expressions as identified in the literature, following Lambrecht (1994: 107): i) presence vs. absence of accent, ii) pronominal vs. lexical coding and iii) definite vs. indefinite marking. Lambrecht (ibid.) notes that accessibility "has no direct phonological or morphological correlates", but may be coded syntactically. Recent research, however, suggests that accessibility may also be coded prosodically (Baumann & Grice 2006) or morpho-lexically (Gundel et al. 1993).

In the preceding paragraphs I have loosely referred to given ‘concepts’, ‘ideas’, ‘referents’ and even ‘information’, using these terms interchangeably. A note is in order now to differentiate between these terms and to define more strictly what is meant by the different terms. In the majority of studies on information structure, the notion of givenness is implicitly or explicitly taken to apply only to discourse referents. Referents are most typically expressed as syntactic arguments, which may be either entities or propositions. Lambrecht (1994: 74ff.) has pointed to the special cognitive status of discourse referents as being identifiable and retrievable, a property the denotata of verbs, adjectives, prepositions and the like lack according to Lambrecht (1994: 111). From a cognitive point of view, events and states more quickly change from one proposition to another, while discourse referents are more persistent, frequently being topics in longer topic chains about whom different events or states are predicated (cf. also Chafe 1994: 68). Also, arguments and predicators show a different behaviour with respect to accentuation in many languages, verbs being more prone to deaccentuation than arguments (Chafe 1974, Schmerling 1976, Gussenhoven 1983, Ladd 1996), a fact that Lambrecht interprets as "a consequence of the inherent difference in the way in which discourse-referential vs. non-referential expressions are processed" (Lambrecht 1994: 267).

On the other hand, events and states may be turned into referents when being referred to by nominal constituents, frequently pronouns, in a following clause (Chafe 1994: 68; Lambrecht 1994: 75). According to Chafe (1976: 28) givenness may also apply to verbs (or more precisely their denotata). Büring (2013: 448, ex. 5b) gives an example of a sentence with a verb that may be deaccented due to its givenness.
(4) (Don't jump! –) But I WANT to jump.

Following Chafe, I will take the position that the cognitive status of *givenness* or *accessibility* applies to all kinds of concepts. I will, however, show that arguments also have a special status in the information structure of Egyptian Arabic utterances. In Egyptian Arabic, this special status may also be signalled prosodically, albeit not by accentuation vs. deaccentuation, but by prosodic phrasing and pause and additionally by other morpho-lexical means (4.5.6).

We may thus come back to the above mentioned tripartite division of information status which I shall will be referring to as *given – accessible – new* in the present model. Information status is assumed to apply to all kinds of concepts concerning their identifiability and presumed activation state in a hearer's consciousness at the time of an utterance. Given that all three categories have been found to have formal linguistic correlates in some languages, they have been applied in annotating the corpus data of the present study. Contrary to Lambrecht, I do not perceive the *information status* of concepts as a *pragmatic property*, but rather as the cognitive basis on which the pragmatic properties of *focus* or *background* are assigned to these concepts, which will be the topic of the next subsection.

1.1.1.2 *The focus-background structure or the speaker's choice*

So far, we have been dealing with information status from a quasi-objective perspective, i.e. we have looked at the potential of a concept to be *judged* as given or accessible at a specific point in discourse, whether that potential is based on prior mention of the concept or its inferability from the discourse-internal or the discourse-external world. However, it is not the case that every concept that in principle satisfies the condition for being coded as given, i.e. expressed in an *attenuated* manner, will also be coded as such. In other words, not every fully active referent is referred to by a pronoun but may also be referred to by a full lexical NP quite frequently, nor is every (content) word referring to a given concept necessarily accentless. The reasons for this are manifold, involving different parameters, among which contrastivity has been a major concern (Chafe 1974, 1976), as in the following example cited by Schmerling (1976: 72). Schmerling uses three degrees of prominence, the weakest one of which will be ignored here, the strongest and second-strongest ones are signalled by large capitals and small capitals, respectively.

(5) JOHN insulted MARY, and then SHE insulted HIM.

Obviously, John and Mary have both been activated in the first clause and referred to by accented pronouns in the second clause. As Schmerling (p. 73) notes, anaphoricity does not directly correlate with lack of accent. In this specific example, the explanation has frequently been given in terms of *contrastivity*, meaning that John and Mary are contrasted with respect to uttering an insult. But there are various other semantic and stylistic reasons as Lambrecht (1994: 96) suggests, not to forget
prosodic ones, such as the principle of rhythmicity (Selkirk 1984, 1995; Truckenbrodt 1995, 1999; Büring 2001; El Zarka 2005) which will be dealt with in chapter 2.1.2, for accenting given linguistic material.

As Chafe emphasizes "it is important to remember that givenness is a status decided on by the speaker, and that it is fundamentally a matter of the speaker's belief that the item is in the addressee's consciousness, not that it is recoverable." (Chafe 1976: 32; emphasis added). The speaker's belief and consequently the speaker's choice is thus the relevant factor determining the actual expression of information status. The axiom of the speaker's choice has been most eloquently stated by Bolinger (1972) in his reply to attempts at accounting for accent placement in terms of syntax (Lakoff (1972) and Bresnan (1971, 1972). Bolinger (1972: 633) emphasizes that accents are "directly reflecting the speaker's intent" and that "[a]ccented words are points of information." Similarly, Schmerling (1976: 73) notes that anaphoricity is not sufficient to account for lack of prominence, she rather attributes the latter to a "lack of 'significance'", making the observation that some items are not accented or stressed because they are in Schmerling's wording "taken for granted" (ibid.). Schmerling's view thus coincides with Bolinger's in that it is the speaker who decides what s/he thinks to be significant or informative and what s/he decides to be worth receiving some degree of prominence accordingly.  Like Bolinger and Schmerling, Halliday's account of information structure rests on the premises about the relationship between accent (or tonic in Halliday's terms) and information. His information focus is marked by one accent (exceptionally two accents), which he defines as

one kind of emphasis, that whereby the speaker marks out a part (which may be the whole) of a message block as that which he wishes to be interpreted as informative. What is focal is 'new' information; not in the sense that it cannot have been previously mentioned, although it is often the case that it has not been, but in the sense that the speaker presents it as not being recoverable from the preceding discourse. (Halliday 1967: 204; emphasis added)

At the beginning of this chapter I have pointed to the fact that the given–new distinction has been used to refer to three different dichotomies in the literature, one of them being what Baumann (2006: 54) refers to as givenness proper, which has been dealt with under the heading of information status in the previous section, the second one relates to focus–background structure. Halliday's focus definition equates focus to "'new' information [...] in the sense that the speaker presents it as not being recoverable from the preceding discourse" (Halliday 1967: 204). Note that Halliday explicitly points to the fact that this information may have been mentioned before. Also his use of scare quotes indicates that he is in fact differentiating between given and new as informational properties related to a specific cognitive state in the addressee's mind and new in the sense of informative and significant. Halliday's focus definition, which is based on the study of English, is a phonological one. The sole correlate of

---

5 This is not to say that Schmerling fully agrees with Bolinger's "radical focus-to-accent" view in all its details. Although she clearly refutes the structure based accent rules of the time (p. 52), she resorts to pragmatic structure instead, introducing the idea of topic-comment structure as a factor determining sentence accentuation, differentiating between different topic accents and accents on the comment part.

9
the pragmatic concept of focus as conceived by Halliday is accentuation. Even though focus definitions vary in the literature on information structure, highlighting is the one feature all characterizations have in common, as noted by Miller (2006) in his survey of focus in the languages of Europe. Miller (2006: 122) points out that "every concept of focus has to do with giving prominence to constituents and the information they carry".

To sum up the discussion so far, it is evident that a distinction has to be drawn between focus as relevant and significant information and new information in terms of information status, which was the topic of the previous section. What is, however, of importance, is the question how an accent is related to the information status of the item that carries it on the one hand, and how it is related to another dichotomous pair of concepts that will be discussed in the following section, the notions of theme-rheme or topic-comment. In the next section, we will discuss the latter pair in relation to another concept involving focus in some yet to be defined sense, namely presupposition and assertion.

As far as the first relation is concerned, Chafe (1976: 31) suggests that given information in English is expressed in a "weaker and more attenuated manner than new information", that is phonologically with lower pitch and lower prominence and morpho-syntactically via pronominalization. We may add to these zero anaphors and different types of ellipses. Chafe also suggests that attenuation might be a universal means of expressing givenness. As I have already mentioned at the beginning of the chapter, a common way of intonational marking of givenness in English is deaccentuation. In recent years, there has been a growing body of studies challenging this assumption (Cruttenden 2006; Swerts et al. 2002; Hellmuth 2010 for Egyptian Arabic, inter alia). I will take up that question again in chapter 4, where I will argue that the Chafean concept of attenuation as a universally valid expression of background information may still be valid, even though deaccentuation clearly is not a universal phenomenon and many languages have been shown to accent given material and even though there is no perfect match of givenness and deaccentuation even in languages that apply the latter as a formal device in the expression of givenness. In chapter 4, it will also become increasingly clear why it is indeed necessary to differentiate between the two types of pragmatic properties, as suggested within the present study, the first one pertaining to the cognitive status of concepts or ideas, referred to as information status here, and the other one pertaining to their potential linguistic expression which is manifested in the focus-background structure, in the sense that focus is understood as a kind of ‘emphasis’ or ‘highlighting’ and background is what the speaker at a given moment in discourse takes for granted and therefore chooses to express in an attenuated manner.

Two major claims will be defended in this study: (1) that the focus-background structure is not entirely predictable from information status and (2) that the focus-background structure is not entirely predictable from the other two other dichotomous pairs we will have to deal with below, the topic-comment partition and the relation of presupposition and assertion involving a second - relational -
type of 'focus' which will be defined in the next subsection (1.1.2). This view entails that the focus-
background structure is in principle not entirely predictable at all, definitely not from an information
structure point of view. There will always remain a residue of unpredictability depending on semantic
and stylistic factors, which in the end is equivalent to assuming a good deal of speaker's choice in how
an utterance is eventually presented.

Coming back to the difference between information status and its expression, we may again take a
look at Halliday's notions of given and new. It is obvious that he does not use the term new in the sense
of the cognitive information status we referred to above, the same is obviously true for his notion of
given. Accordingly, Halliday views given and new as structural functions assigned to constituents in
the information unit, explicitly stating that "given is offered as recoverable anaphorically or
situationally" (Halliday 1967: 211). Of course, Halliday, like many others, does not draw a categorical
distinction between these notions in the same way he clearly differentiates between information
structure that organizes a communicative act into "message blocks" and thematization that operates on
the syntactic level of the clause. I will explain the different notions involved in Halliday's account
using one of his examples (p. 202) in a slightly adapted version and interpreting the different roles the
constituents carry in Halliday's system.⁶

(6)
   a. // MARY // always goes to TOWN on Saturdays. //
   b. // MARY // always goes to town on SATURDAYS. //
   c. // MARY always goes to // TOWN on Saturdays. //
   d. // On SATURDAYS // Mary always goes to TOWN. //

According to Halliday, the four versions above correspond to four distinct options to convey the
proposition that Mary always goes to town on Saturdays. The first choice the speaker makes is the
division of the proposition into information units. Halliday stipulates that every such information unit
coincides with a tone group, but nowhere does he make explicit what the exact criteria are according
to which the parsing of the utterance is done. He then assumes that this parsing out of message blocks
– which we may take to be an equivalent to phrasing (cf. 2.2.6) – is the basis for, but not fully
specifies the position of the tonic or the information focus. If we compare (6a) with (6b) we find the
same distribution of information units, but different locations of the information focus. We may infer
from the information structure of the two sentences that ‘Mary’ must already have been a topic of
discourse in (6d), but not necessarily in (6a-c), but we do not know whether any of the concepts
‘Mary’, ‘town’ and ‘Saturdays’ have been mentioned before when they are accented. We can only safely
assume that they have or that they are taken for granted, being inferable, if they are not accented. This
implies that givenness is a necessary, but not a sufficient condition for attenuation and that it is
possible to infer givenness from lack of accent, but not to predict lack of accent from givenness. We
could also imagine the above proposition to be articulated in the way illustrated by example (7) which

⁶ Halliday uses bold face for the accented constituents, I have used capital letters for the sake of consistency.
exhibits four accent domains. This articulation of the utterance, however, does not give us any hint as to the givenness of the constituents. I will not mark any domain boundaries here, as I take boundaries and phrasing to be an additional device, orthogonal to accent placement and accent shape (cf. 2.2.6 for a detailed discussion of the issue).

(7) MARY ALWAYS goes to TOWN on SATURDAYS.

Many studies on intonation and prosody have followed Halliday's assumption of the unmarked prosodic unit (Halliday's information unit or tone-group) as being co-extensive with a clause and marked by one single accent called the nucleus that is located at the end of the domain. A case in point in that respect is the Chomsky-Halle Nuclear Stress Rule (NSR)\(^7\) (Chomsky & Halle 1968), which – despite the massive critique levelled at it due to the wrong predictions it makes in many cases (Schmerling 1976; Bolinger 1972) – still turns up in modified and adapted versions in most generative models of intonation and information structure. But the investigation of naturally occurring speech reveals that this assumption is an empirically not verifiable abstraction. Brown et al. make the following observation:

Halliday's 'unmarked' tone group structure, with the pretonic correlating with given and the tonic with new, reflecting the Prague school's notion of the communicative dynamism of a sentence which increases as the sentence progresses, depends on a notion of information structure within the tone group. This may be an appropriate analysis for information structure in written language or the conversations of academics, which tend to be syntactically more complex than the typically paratactic language of our data. In pause-defined units there may be several foci, marked as contrastive/emphatic or new by the speaker. (Brown et al. 1980: 160; emphasis added)

The important point is the difference in methodology employed by Brown et al. and Halliday, respectively. While Halliday's analysis rests on an ill-defined notion of information unit that is mapped to a prosodic unit called the tone group, Brown et al. explicitly use a prosodic unit as a point of departure which they define as the stretch of speech occurring between pauses. As Brown et al. observe, such prosodic units may contain ‘several foci’. Obviously, Brown et al. use the notion of focus as an equivalent of accent. In this respect, their account resembles that of Bolinger (1958, 1986) who has repeatedly emphasized that every accent is itself meaningful, which is tantamount to saying every accent may be a focus. Contrary to Halliday's system, Brown et al.'s and Bolinger's accounts are thus not based on structure. There are, however, structure-based accounts of focus prosody that also assume the equivalence of focus and accent, e.g. the one proposed by Gussenhoven (1983) in his seminal paper **FOCUS, mode, and the nucleus** and Selkirk’s earlier work (Selkirk 1984, 1995). Selkirk (1984) called her model a "pitch-accent first"\(^8\) theory, trying to account for the fact that accents may occur in various locations within a domain. Selkirk explained the relationship between accent and focus as follows:

---

7 In the framework of Generative Grammar, the NSR was developed to account for the empirical fact or perhaps rather the intuition that the main prominence on sentence level frequently can be perceived on the last lexical item. A sentence or clause is taken to be expressed by a phrasal prosodic constituent. The concept will be dealt with in more detail in chapter 2 below.

8 The notion of pitch accent and other prosodic categories will be defined in chapter 2.
Our empirical claim is that in English it is the pitch accent assignment of a sentence that is directly related to its focus properties. Again roughly speaking, the presence of a pitch accent correlates with a focus (and thus with 'new information'), while the absence of a pitch accent indicates the lack of focus (or 'old information'). (Selkirk 1984: 200)

We will discuss the different models in more detail in Chapter 3.1 and refer to it again in Chapter 4.5.3 for the discussion of the Egyptian Arabic facts.

The problem that arises under the assumption of a one-to-one relation between accent and focus is the empirical observation that not every item carrying an accent can be said to be expressing 'new' information, in the sense of the speaker's declared contribution to discourse. As a solution to this problem, most accounts assume in one way or another that it is only one (the last) of such accents in a certain domain that carries the focus marking - following the same nuclear stress principle as referred to above - while the other accents are viewed as secondary. Büring (2007) for instance, has called such accents "ornamental accents". He assumes that they are assigned to given prefocal elements that are not focus-marked themselves. But this assumption is equally problematic as we will see in the course of this study, as the special treatment of the final accent in a certain domain is an arbitrary stipulation that does not hold empirically. This decision relies on the – equally arbitrary – assumption that the final accent in a domain is always perceived as the strongest, which again cannot be verified by spontaneous corpus data (cf. Brown et al. 1980: ch.5). In chapter 2.2, I will return to this issue and argue for an approach that does away with the notion of an obligatory nucleus as the correlate of FOCUS and that evaluates every individual accent in its own right instead. I will argue that it is the relative prominence of the accents that is relevant for the interpretation of the utterance. Prominences and prominence scales, however, are not only correlates of information structure, but also of other phenomena, such as the rhythmic organization of an utterance or the semantic weight of the lexical items involved. In the light of naturally occurring language data, the frequently referred to case of intonation units carrying only one accent at the end, and at most perhaps a second (and secondary) one, that is tacitly or explicitly assumed to be the default case, on the contrary, seems to be exceptional, even in English, unless an intonation phrase is sufficiently short. It has in fact been pointed by several authors that in ‘broad focus’, an accent on every lexical item is the default (Bolinger 1986; Jacobs 1988; Gussenhoven 1992; Büring 2007; Féry & Kügler 2008; inter alia). We may therefore state that the approaches to the focus-to-accent problem in the literature is solved in two different ways: One branch of approaches views accent as directly reflecting focus and another branch does not and accounts for the existence of different types of accents assigning focus status to only one of them. In Chapter 4, I will present a model that suggests a third solution. My approach makes a distinction between foci and default (i.e. rhythmic) accents, but it also does away with an obligatory focus accent that is the focus exponent for a whole domain. I shall have more to say on this issue in Section 1.1.2. In the remainder of this section, I shall clarify some notions that are important for the identification of a focus in the sense of focus accent.
The problem of the one-to-one correspondence of focus and accent is especially obvious in the existence of accented topics which are not normally assumed to be 'new' information. It has been observed that topics are mostly characterized by an accent type that differs from the one used for elements that introduce new information (Bolinger 1958, 1986; Jackendoff 1972; Brazil 1975, 1997; Uhmann 1991; Steedman 1991, 2000; Féry 1993, Büring 1997; *inter alia*). These issues will be the subject of the next subsection and will be further developed in the prosodic model presented in Chapter 2. Empirical evidence from Egyptian Arabic supplied for the assumptions made there will be provided especially in Chapter 4.

It will have become clear from the preceding discussion that a focus marked by highlighting is not confined to new information, nor is it compatible with a definition that assumes that focus is necessarily part of the comment in a topic-comment sentence such as in the question-answer pair in (8), as another type of focus definition suggests (cf. section 1.1.2):

(8) A: What can you tell me about Mary?  
B: MARY // [always goes to TOWN on Saturdays.]_comment

What the highlighting approach to focus rather is compatible with, is the notion of *contrast* (Chafe 1976). Chafe (1976: 33) gives the following example (9) to illustrate his notion of *contrastiveness*. The context for the utterance is the organization of a party and the preparations for it in which various persons are involved and the sentence asserts the fact that Ronald is the one who prepared the hamburgers and nobody else.

(9) RONALD made the hamburgers.

According to Chafe (1976: 33f.), the sentence in (9) consists of a background and a *focus of contrast*. It rests on the awareness that someone made the hamburgers, the background knowledge, and expresses the idea that there is a limited number of candidates who could have done it and the assertion that a certain candidate - in this case Ronald - is the correct one. Chafe notes that contrastiveness is expressed by higher pitch and stronger stress but that it cannot always be identified "on a phonetic basis alone" (p. 36). But he identifies a steeper fall in pitch after contrastive accents than after accents that only mark new information, noting that “[p]robably this increased prominence of contrastive focus results from an increased emotional commitment that is likely to accompany contrastiveness” (ibid.). While in (9) the contrastive item is also the 'new' information, Chafe explicitly argues that contrastivity may relate to topics and FOCI (in the sense of being the new or relevant information, cf. Section 1.1.2) alike.

Some of the more recent models of information structure that invoke a category of *contrast* differ from Chafe's approach in the type of the *alternative set* they assume. While Chafe explicitly limits this set to

---

9 Note that Chafe generally avoids the notion of focus as a structural category, using only givenness and contrastiveness as informational categories.
a certain number of candidates, others allow for an open set of alternatives. This is true for Vallduví and Vilkuna's notion of *kontrast* and Molnar's (Molnár 2006) I-CONTRAST\(^\text{10}\) which is said to imply the existence of alternatives without necessarily excluding them. Such a weak form of 'contrast' however is in no way different from a definition of focus that is widely resorted to in generative models based on *Alternative Semantics* (Rooth 1985, 1992) which defines focus as evoking a contextually salient set of alternatives (Rooth 1985, 1992; Jacobs 1988; Krifka 2007).

The position taken here is that contrast is not the only function that can be signalled by a focus. As suggested above, elements of a linguistic utterance may also be highlighted for other purposes, such as the activation of referents or semantic foregrounding. Also, whether contrast is a genuine linguistic category is a matter of some debate. Contrary to the above mentioned position held by Chafe and others, some scholars, e.g. Bolinger and Lambrecht (1994), prefer to view contrast as an interpretation arising as a result of general conversational implicatures. This view is expressed by Bolinger (1961) in a discussion of contrastive accent in English:

> In a broad sense, every semantic peak is contrastive. Clearly, in *Let's have a picnic*, coming as a suggestion out of the blue, there is no contrast with dinner party, but there is a contrast between picnicking and anything else the group might do. As the alternatives are narrowed down, we get closer to what we think of as a contrastive accent. (Bolinger 1961: 87)

For Bolinger, contrastive accent is thus a gradient concept that could readily be subsumed under the more general term 'emphasis'. Due to its vagueness and non-scientific flavour, the concept of *emphasis* is not especially popular among linguists. As a result, there have been attempts to replace it by a more constrained notion, as for example the concept of *imposed salience* (Clamons et al. 1993; Mulkern 2003) which is taken to be "indicating the importance or foregrounding the speaker chooses to give to particular discourse entities. The notion of imposed salience is used in developing more precise characterizations of contrast and emphasis." (Mulkern 2003: iii). Whether Mulkern succeeds in getting rid of the vagueness attached to the notions of *salience* and *emphasis* or not is a different issue, but two aspects are interesting in her proposal: (i) Imposed *salience* differs from 'inherent salience' (which) is equivalent to the cognitive state of referents referred to above. In fact, according to Mulkern (p. 24f.), speakers may choose to impose salience to discourse entities to (re-)rank them relative to one another and they may do so employing different types of linguistic devices, such as prosody, morphological marking and certain syntactic constructions. ii) Imposed salience is also different from *FOCUS* as a pragmatic *relation* (1.1.2) and may accordingly also be applied to topical expressions that denote an already given entity as in (8) above, and (ii). The concept has recently been used by Kim (Kim) to give an alternative explanation for the use of Korean morphological markers -(n)jun and -i/ka which have been commonly assumed to mark topic or contrast and focus, respectively.

---

\(^{10}\) Molnár (2006) proposes two different types of contrast, I-CONTRAST signaled by prosody (e.g. in English) and S-CONTRAST signaled by syntactic means (e.g. in Finnish). She argues that the tonal contour (the fall-rise) in case of a contrastive topic may operate on an open set of alternatives and does not necessarily exclude all other alternatives (i.e. is not necessarily exhaustive).
These various attempts only show the need for some notion that accounts for the enhancement of the semantico-pragmatic properties of an entity or of an event in a given discourse, independently from the pragmatic role it plays in a proposition and independently from the cognitive state associated with it in the interlocutors' minds. I believe that all the different notions mentioned in this section are in fact based on exactly this same need, which is simply to account for a special emphasis given to individual elements. The purposes for this may vary, whether they be contrast, (exhaustive) identification, activation or simply the relevance of a given entity to the speaker. As Mulkern (2003: 10) points out, the actual distribution and interpretation of the various instances of imposed salience is a matter of "conventional use and conversational implicature". Thus we are back to where we started with Halliday's characterization of focus as "a kind of emphasis" (Halliday 1967: 204) signaled by prominence.

In the present work, I will make use of this notion of focus and define it in contradistinction to background as a special emphasis given to a certain denotatum for multiple reasons, e.g. by means of prosodic prominence. In chapter 4 this notion will be further developed and sustained with empirical data from Egyptian Arabic. This type of focus encodes a pragmatic property of the element it is attached to, namely its information value or relevance as conceived by the speaker. It is essentially equivalent to the focus concept in much of the work on the prosody of information structure cited above. I will call this notion focus of interest or simply focus, following Bolinger's terminology (Bolinger 1958 and subsequent work) and define it as follows:

(10) A focus of interest is a semantic component that the speaker choses to foreground, due to the information value s/he attributes to it at a certain moment in discourse.

1.1.2 Pragmatic relations

Accounts of information structure differ as to whether they make a distinction between focus/background and topic/comment or theme/rheme (Halliday 1967; Jacobs 1988; Molnár 1993; Krifka 1991/92, 2007) or whether they conflate them into a one-dimensional model such as in Vallduví (1992), Lambrecht (1994), Erteschik-Shir (1997) or Büring (1997). As we have seen above, models that have given up the dimension of focus/background and topic/comment as two intersecting notions of information structure frequently resort to the notion of contrast (Vallduví & Vilkuna 1998) to account for topics that are put into the foreground, but also to differentiate between two types of focus, mostly called contrastive focus or identificational focus and information focus, respectively (É. Kiss 2006; Gundel 1999; Zimmermann 2007).

In the above subsection I have defined focus as a pragmatic property that is locally marked on a specific element of word-size or even smaller. In the following subsection I will propose another
notion based on Lambrecht's characterization of 'focus' as a pragmatic relation that may involve larger constituents. Such linguistic entities may range from the size of an argument, a complex predicate phrase to whole clauses and sentences. In Chapter 4 I will show that while both types of 'focus' have been used interchangeably in the literature, they in fact pertain to different linguistic functions. While focus of interest highlights individual concepts, rhematic FOCUS is one (and sometimes the only) part of a structured proposition.

1.1.2.1 FOCUS, presupposition and assertion

Halliday (1967: 204) takes the function of what he calls information focus to “reflect the speaker’s decision where the main burden of a message lies” or what is ‘new’ in a message. An important aspect of his definition is that focus may have scope over the whole information unit (p. 243). Ever since, ‘focus’ has also been associated with the information a speaker wants the addressee to add to his previous knowledge. It is this special aspect that has given rise to definitions that refer to this aspect of 'focus' and define it in some way as the complement of a presupposition (Chomsky 1970; Jackendoff 1972, Vallduví 1992; Lambrecht 1994).

A presupposition frequently contains given information (1.1.1.1) as it is that part of a pragmatically structured proposition that represents the common ground between speaker and hearer. While information status refers to specific referents, a presupposition is related to knowledge or information, i.e. to whole propositions. But presupposed propositions do not necessarily consist of given concepts. New information may also be accommodated as presupposed, if the speaker presents a proposition as if the hearer already knew it, i.e. he may only be pretending that the proposition belongs to the common ground. In the following example taken from Lambrecht (1994: 68), which is supposed to be the first sentence of a short story, it is clear that moving to Switzerland is new information. This information is, however, presented as a presupposed proposition, i.e. as if it already belonged to the common ground between speaker and hearer. This is formally indicated by the use of a before-clause, which in the unmarked case is coding pragmatically presupposed propositions. In the case at hand, the presupposition is accommodated, i.e. it is created by the very act uttering the sentence as it is.

(11) Before I moved to Switzerland I had never seen a Rolls Royce.

But more frequently presupposed propositions have given referents. Thus the cookies in (13a) below or the hamburgers in (9) above are given or at least accessible to the addressee. The presuppositions in these utterances are that someone stole the cookies (13a) or that someone made the hamburgers (9). Matić (2003: 88) proposed that it is essential to make a distinction between existential presuppositions and relational presuppositions. I will follow Matić in that assumption and adopt his definition of existential presupposition as playing "the role of a term (as opposed to predicate) on the propositional
level”, pointing to "the existence of an entity or a state of affairs" (ibid., emphasis in the orig.). A relational presupposition is defined by Matić (2003: 89; emphasis in the orig.) as "presupposed relative to the utterance", i.e. as the proposition to which the speaker relates an assertion to be made. We will see in Section (1.1.2.2) how relational presupposition is related to the notion of TOPIC as I define it here. According to Lambrecht, focus is the “semantic component of a pragmatically structured proposition whereby the assertion differs from the presupposition” (Lambrecht 1994: 213). I will refer to this kind of 'focus' as rhematic FOCUS, or FOCUS for short, and assume with Lambrecht that it plays a pragmatic role within a structured proposition. In accordance with the above assumption, I slightly alter Lambrecht's definition and define FOCUS as follows:

(12) A rhematic FOCUS is the semantic component of a pragmatically structured proposition whereby the assertion differs from the relational presupposition.

The notion of rhematic FOCUS is basically equivalent to the old category of rheme assumed by the Prague structuralists. In the literature, it has been suggested that coherent discourse is structured by implicit questions (e.g. Büring 2003). If we take as an example sentence (13a) and compare it with the segmentally identical sentence in (13b) we see that the two sentences only differ in accent position.

(13)

a. JOHN stole the cookies.

b. John stole the COOKIES.

While (13a) can only function as an answer to a question such as ‘Who stole the cookies?’ with the presupposition that some cookies have been stolen, (13b) may answer the questions ‘What did John steal?’ or ‘What did John do?’ or ‘What about John?’ with the only presupposed fact being the existence of John who is known to both interlocutors and may thus be assigned the cognitive state given or accessible, depending on the context. The important observation in such cases is that the part of the sentence that conveys the new or newsworthy information is not immediately identifiable from the formal expression of the sentence. Cases like (13a) are referred to as narrow FOCUS from a formal point of view or as argument FOCUS from a semantic point of view. The same is true for the interpretation of (13b) that assigns a focus of interest only to the cookies. In the second interpretation, however, we are dealing with broad or wide FOCUS. In relation to what is commonly referred to as the topic-comment structure of a sentence, the broad FOCUS constituent in (13b) is co-extensive with the comment about the topic 'John'. In the case of broad FOCUS the accented constituent is usually called the focus exponent in the literature and the whole predicate phrase is called the focus domain. On such a view, the problem that now arises is how to interpret the single accent on the object noun and how to arrive at the scope of focus or the focus domain. This issue has preoccupied most of the research on focus prosody during the second half of the past century.
The presumption for this discussion, however, contains two weak points. Firstly, it has been initiated by and developed on the basis of English and other West-Germanic languages, which are known for their strong tendency to deaccent given material and verbs. Secondly, it refers to mostly constructed examples, such as written sentences of various lengths, as they very rarely occur in spoken language. In spontaneous dialogue, information units, in the sense Halliday uses this term, are frequently rather short, and rarely correspond to syntactically complete sentences. They usually consist of one simple proposition or a part of it, as suggested by Chafe's "one new idea constraint" (Chafe 1994). Such sometimes only exhibit one accent, and the question of the relationship between this accent and the focus domain simply does not arise. But even where a prosodic unit is co-extensive with a more complex proposition in spontaneous speech one nuclear focus accent cannot be identified. Thus, the whole discussion about focus domains and focus exponents and the mapping rules between them is rather academic and their application to spoken language data quickly reveals that they do not hold empirically (cf. Brown et al. 1980). The whole issue of accentuation rules will not concern us here for the moment, but will be discussed in chapter 4. The important observation here is, however, that longer one-accent utterances have been given a special status in the literature on focus prosody they do not deserve (cf. Fuchs 1976: 300).

1.1.2.2 FOCUS vs. focus

Given that utterances with a broad FOCUS more often than not contain multiple accents I will take up the distinction of focus of interest and rhematic FOCUS made above and elaborate on the argumentation. First, FOCUS as part of an assertion entails that a topic is not to be included in the FOCUS domain (we will, however, refine the relationship between information structural categories later, cf. 4.5). We have, however, seen that topic expressions may very well be made prominent, especially when they are contrastive. Thus Lambrecht distinguishes between contrastive topic and contrastive focus. It has, however, also been suggested in the literature that a contrastive topic contains a focus (Molnár 1998; Krifka 2007). This assumption can be maintained if focus and FOCUS are assumed to be related, but not identical notions. The sentences in (14) illustrate different domains of the rhematic FOCUS, covering the whole sentence (14a), only the predicate phrase (14b) or just an argument of the proposition, the object in (14c) and the subject in (14d). The first two are instances of broad FOCUS, while the second two are instances of narrow FOCUS.

(14)

a. What happened?
   [PETER stole the COOKIES]_{FOC}.

b. What did Peter do?
   Peter [stole the COOKIES]_{FOC}.

c. What did Peter steal?
Peter stole [the COOKIES]_{FOC}.

d. Who stole the cookies?
   [PETER]_{FOC} stole the cookies.

The assumption of a distinction between focus as a pragmatic property and FOCUS as a pragmatic relation allows a unified account of ‘contrastiveness’, as contrastive topics can be interpreted as involving a *focus*. This is illustrated by an example (15), adapted from Krifka (2007: 7).

(15)
A: What do your siblings do?
B: [My [SISTER]_{loc} TOP [studies [MEDicine]_{loc} FOC and [my BROTHER]_{loc} TOP [is working on a [FREIGHT ship]_{loc} FOC.

Furthermore there are other instances of foregrounding that can be accounted for by the assumption of a separate notion of *focus of interest* shown in the wide FOCUS sentences in (16).

(16)
a. A: What happened?
   B: [John]_{top} [stole the COOKIES]_{FOC}
   B’: [[JOHN]_{loc} stole the [COOKIES]_{loc} FOC
   B”’: [[JOHN]_{loc} [STOLE]_{loc} the [COOKIES]_{loc} FOC
b. A: What did John do?
   B: [John]_{top} [stole the [COOKIES]_{loc} FOC
   B’: [[JOHN]_{loc} top [stole the [COOKIES]_{loc} FOC

If we wish to account for the different prosodic realisations of the given answers, we have to account for the presence of the individual accents within the larger domains. As already noted above, such accents have been dealt with in the literature in various ways. Some theories (especially the prosodically based ones) equate focus domain with accent domain (Halliday 1967; Gussenhoven 1983; Selkirk 1984; Bolinger 1986). Some theories allow for the embedding of these foci into larger focus domains, such as Selkirk’s (1984) model (see Chapter 3.1 for more details). In the present study, I take the view that the different answers shown in (16) have a different pragmatic impact. In chapter 4 the proposed distinction will be furnished with further arguments concerning the formal expression of the focus and FOCUS based on the empirical analysis of the Egyptian Arabic data.

The argument for the distinction between these two pragmatic categories thus comes from the fact that not every information unit that contains a focal accent can be interpreted as a *FOCUS domain*. If we take the example in (15) or in the example from Halliday (1967) quoted in (6d) above, the accents on ‘sister’, ‘brother’ and ‘Saturdays’ do not mark the constituents they are associated with as ‘rhetic’, i.e. they do not constitute the most relevant or most important information of the propositions. They may, however, be said to evoke *alternatives*. I believe that such alternatives can hardly be assumed to
apply to large constituents, such as whole predicate phrases or complete sentences. But the implicature that there are other alternatives to the element(s) marked by an accent is rather restricted to the immediate constituent that contains this accent. Take the accents on the rhematic parts in (13) and (15), for example. If the accent on 'freight' in 'freight ship' evokes any alternatives at all, it might possibly evoke the alternative 'passenger ship' or perhaps also 'freight train', but it seems implausible to me that it might evoke alternatives for the entire predicate, such as 'is working as an engineer' or 'owns a big company' or 'is really good for nothing'. Consequently, I believe focus to only have scope over the referent it is attached to. In chapter 4 we will explore the possibility that large FOCUS domains, such as complex predicate phrases and sentences may not contain a focus at all.

The differentiation between the two 'focus' types solves another scope problem that arises when only part of a referring expression is highlighted. This is, for example, the case with corrections. In some languages (e.g. English) it is possible to use accent to correct elements as small as a bound morpheme or even a phoneme, like in Bolinger's oft-cited example "The whisky wasn't EXported, it was DEported" (Bolinger 1961: 83). While we can definitely say that the two morphemes 'ex' and 'de' are contrasted by means of a focus accent, it makes no sense to view these constituents in terms of assertion and presupposition. No one will seriously claim that 'ex' is the FOCUS completing a presupposition that 'whisky is being 'ported".

The same is also true for even less intuitive cases of an NP containing a modifier adduced by Lambrecht (1994: 216), quoted here as (17).

(17)

Which shirt did you buy? -- I bought the GREEN one.
The GREEN one.
*GREEN.

Lambrecht argues that the FOCUS domain must be the whole NP and not just the colour expression 'green'. The answer to the question in (17) does not have to be a whole sentence, the NP is enough. But answering only with the colour expression alone is non-felicitous. According to standard focus projection rules, focus cannot project from a modifier up to the higher domain of the NP and would thus be restricted to the colour 'green' (but cf. Büring 2006 for a different view; see also 4.5.2 and 4.5.3 on that issue), which again cannot be interpreted in terms of assertion and presupposition. Only the referent, which is the 'green shirt', enters into a pragmatic relation to the proposition. Within the present account, there is no contradiction between the two analyses. While the FOCUS of the sentence is the whole argument with its property of being a shirt and being green that stands in a FOCUS relation to the whole proposition, the focus only falls on the modifier, emphasising the specific property of being green and thereby evoking alternatives, such as 'red', 'blue' or 'yellow' or rather 'red

---

11 Note that although it is usually assumed that the notion of alternatives is also applicable to larger constituents and even whole sentences, it is in practice confined to narrow focus cases.
shirt’, ‘blue shirt’ and ‘yellow shirt’. Thus again, the difference between the two focus types is one of *property* versus *relation*. While the colour has the property of being the notion of interest within the answer, the whole referent is the information required. After all, is not the colour that is the object of buying, but the whole shirt.

Similarly, the foci ‘John’, and ‘cookies’ in example (16) do not constitute information on their own. It is through their topic or FOCUS function that they acquire an informational value: the ‘cookies’ as part of the rhematic FOCUS, and ‘John’ as the topic in (16b) and (16a, B) and part of the FOCUS in (16a, B’ and B”). This idea will be elaborated in Chapter 4, where we will take a closer look on the phonological realization of pragmatic constituents.

To summarize, I assume with Lambrecht (1994) a basic distinction between two information structure categories, *pragmatic properties* and *pragmatic relations*. For the time being, it is important to differentiate between focus as the pragmatic property of being activated, contrastive, or simply relevant and interesting, constituting only a ‘point’ of information (Halliday, Bolinger) and not the information itself. Focus has the phonological correlate of highlighting the denotatum of the expression it is associated with. I thus follow Bolinger who defines focus in very general terms as the expression which is highlighted: "My position is that accents are prima facie iconic, responding to the speaker's sensation of *interest* in what he is saying plus a general desire to impress" (Bolinger 1985: 120; emphasis in the orig.).

**FOCUS**, on the other hand is the new or relevant *information* the speaker wants the addressee to add to his previous knowledge. The semantic constituent that is the FOCUS enters into a pragmatic relation with the whole proposition and thereby constitutes the information that is given about a specific topic. The canonical expression of **FOCUS** is thus by the comment in a topic-comment sentence, the "psychological predicate" of van der Gabelentz (1869) and Hermann Paul (1880) (cf. von Heusinger 1999). Thus, FOCUS has also been defined as ‘predication’ (É. Kiss 2006; Zimmermann 2012).

From a linguistic point of view, we are of course interested in the formal expression of the two categories. As far as the prosodic correlates are concerned, it is generally assumed in the literature that every focus domain is somehow marked by highlighting. In the present account I shall argue that it is only *focus* that is necessarily marked by highlighting, i.e. by prominence, while **FOCUS** may or may not involve a special prominence. As I will show in the following chapters, specifically in chapter 4, the more stable prosodic correlate of FOCUS is intonation proper, i.e. tonal shape, i.e. the quality of the pitch contour.
1.1.2.2 Topic, TOPIC and relational presupposition

Although the definitions of topic are commonly viewed as rather vague, it is generally assumed that topic is a cross-linguistically important notion of information structure. The concept of topic dealt with in this study is confined to sentence topic. Little is said about discourse topics and the issue of discourse development in general, although this is a fascinating area of research that is intimately connected to the notion of sentence topic. Thus, discourse topics are occasionally alluded to, but only where necessary. One major aspect of topicality dealt with in the present study is illustrated by the following example (18) from Krifka (2007: 41; annotation changed according to the conventions of the present model):

(18)


Definitions of topic vary in the literature. However, the most widely accepted definition is in terms of an aboutness-relation that holds between the topic of a specific sentence or clause and the whole proposition expressed by this clause (Hockett 1958; Gundel 1976; Dik 1978; Dik 1978; Reinhart 1982; Lambrecht 1994). In (18a) it is Onassis who is the topic of the sentence and ‘marrying Jacqueline Kennedy’ is being predicated about that topic; in (18b) it is the other way round. Again, it is in the first place the perspective that differs between the two sentences, not their truth-conditional value - the result being in both cases that Aristotle and Jacqueline got married. While (18a) would be expected in a context dealing with Aristotle Onassis, we expect (18b) to be uttered in a context in which Jacqueline Kennedy has been previously referred to.

In addition to the aboutness function, there is another function that has been attributed to 'topics' in the literature. This "scene-setting" function (Chafe 1976) has to do with anchoring a proposition in time and space (19a) or relating it to an individual concept, thereby delimiting its applicability (19b). Topics that fulfil such a delimiting function are for example sentence initial constituents as they occur in so-called "topic prominent" languages (Li & Thompson 1976), an example from Japanese is given in (19b). This function shall be referred to as a frame.

(19)

a. Tuesday]FRAME TOP I went to the dentist. (Chafe 1976: 51)  
b. [Sakana wa]FRAME TOP tai ga oisii  
   fish TOP red.snapper NOM delicious  
   ‘Fish, red snapper tastes good’ (Li & Thompson 1976: 468)

Similarly, Dik (1978) distinguishes topic and theme. He conceives the latter as being located in the left periphery of a sentence and not as part of the predication proper. Chafe's topics and Dik's themes are
thus left attached to the main clause and are not arguments of the proposition. This dimension of topicality is related to the general linear progression of pragmatic constituents in a sentence that has frequently been assumed (cf. the quote from von der Gabelentz below). By the same token, the Prague School’s distinction between theme and rheme served to explain facts about word order in Slavic languages. Halliday (1967) adopted the idea that the theme necessarily comes first in a sentence. In fact, Halliday distinguishes between information structure which in his account is defined phonologically, involving tone-groups as information units that contain an information focus, i.e. an accent, as noted above, and the thematic structure which pertains to syntax.

In the present study, I will also distinguish between two types of ‘topic’ - on a par with the two types of ‘focus’ presented above. But the distinction drawn here is not exactly equivalent to any of the distinctions mentioned here. In the present model, a distinction is made between entity topic and thematic TOPIC. An entity topic necessarily denotes a referent, about whom a predication as made, whereas a thematic TOPIC (or TOPIC for short) refers to one or more semantic constituents that are necessarily sentence initial and typically exhibit a certain type of intonation. TOPIC is the more comprehensive category that may contain all different kinds of constituents, such as entity topics, all kinds of frame expressions like adverbials or different types of subordinate clauses. Crucially, a TOPIC may even contain a verb, because TOPIC is used here to be partially co-extensive with what has been referred to as relational presupposition above. An entity topic may be expressed in different ways, for instance by a full lexical NP, a personal pronoun, a pronominal suffix, a possessive suffix, a demonstrative pronoun or in Egyptian Arabic, a null subject language, not at all. The definition of thematic TOPIC responds to the observation that topics are preferably sentence-initial and thus partly corresponds to Halliday’s or Dik’s theme and Chafe’s topic. In Chapter 5, the distinction drawn will be argued for in more detail and formal evidence from Egyptian Arabic will be provided.

Ever since Aristotle postulated the universal partition of judgment into two parts (double judgment), the subject (hupokeimenon) and the predicate (katēgoroumenon), this concept has been supposed to be reflected in the sentence as the physical manifestation of a judgment. Von der Gabelentz (1869) in an amazing article on comparative syntax, felt the need to distinguish grammatical subject and predicate from the pragmatic notions he called “psychological subject” and “psychological predicate”:

Was bezweckt man nun, wenn man zu einem Andern etwas spricht? Man will dadurch einen Gedanken in ihm erwecken. Ich glaube, hierzu gehört ein Doppeltes: erstens, daß man des Andern Aufmerksamkeit (sein Denken) auf etwas hinleite, zweitens, daß man ihn über dieses Etwas das und das denken lasse; und ich nenne das, woran, worüber ich den Angeredeten denken lassen will, das psychologische Subject, das, was er darüber denken soll, das psychologische Prädicat. (von der Gabelentz 1869: 378, emphasis in the orig.).

(‘What is our intention, when we talk to someone about something? We want to evoke an idea in him. I think that this involves two things: first, we have to draw the other’s attention (his thinking) to something, and second, we make him think such and such about that something. What I want the other person to think of, I call the psychological subject, and what he is supposed to think, is the psychological predicate.’)
Von der Gabelentz (p. 379) also notes that the psychological subject necessarily precedes the psychological predicate. In the rest of the article, the psychological subject is described in a way that suggests it to be largely equivalent to our category of TOPIC.

1.1.3. Theticity

The 'double judgment' referred to above was further dealt with in the philosophy-of-language debate at the end of the 19th century by Franz Brentano and his pupil Anton Marty under the heading of *categorical statement*. Brentano and Marty questioned the Aristotelian axiom that every human judgment is necessarily expressed in a proposition involving a *subject* and a *predicate* that may either be affirmed or denied, thereby involving two separate cognitive acts: presentation (*Vorstellung*) and judgment (*Urteil*). They claimed that there exists another type of judgment which is logically unstructured and does only involve one cognitive act of affirmation or denial of a presentation. In the Brentano-Marty theory the distinction pertains to thought and not to actual linguistic statements. The theory was later taken up and applied to linguistics proper by Kuroda (Kuroda 1972) who pointed to the linguistic relevance of the distinction, based on the fact that Japanese makes a distinction between topical and non-topical subjects (20). His ideas were further developed by Sasse (Sasse 1987) whose main claim was that the thetic/categorical distinction is also a valid grammatical concept. Sasse (1987) applied the notion of *categorical* statement to sentences that select a *topic* or a *predication base* about which a predication is made by means of the *predicate*. Thetic utterances, on the other hand, are characterized by the fact that they do not select such an entity, but it is the utterance as a whole that expresses the statement. Thetic sentences have come under different names such as "neutral description" (Kuno 1972), "news sentences" (Schmerling 1976), "all-new utterance" (Allerton & Cruttenden 1979; Fuchs 1980), "presentative constructions" (Bolinger 1989), "eventive" (Gussenhoven 1983) or "event-reporting" (Lambrecht 1988) sentences.

Today, the terms *thetic* and *categorical* are commonly used for two types of assertion that crucially differ in whether they select a topic (or predication base as the topic was called in the seminal paper by Sasse (1987) and predicate something about it (*categorical*), or not (*thetic*).

Typical examples for the two categories are given in (20), quoted after Sasse (1987: 512, 514).
a. THETIC ('simple')

English: *It is raining.*

dog-PRT running is
There is a dog running.

b. CATEGORICAL ('double'):

English: *[John] [is intelligent].*
Japanese: *Inu-wa hasitte iru*
do-PRT running is
The/A dog is running.'

Consequently, thetic statements are not viewed as predications by Sasse, but as simple assertions “of the existence of a state of affairs” (Sasse, 1987: 556) without the bipartite division of labour that is characteristic of categorical utterances. Speaking in grammatical categories, this implies that the constructions used will only be felicitous if they avoid the canonical topic-comment structure of categorical statements. Such constructions may either exploit morphological, lexical, syntactic or prosodic means for that purpose. What all constructions seem to have in common is that they exhibit a form opposite to or at least different from the corresponding topic-comment construction, whether it be syntactic (VS instead of SV order) order to avoid the occurrence of the subject in the canonical topic position, or “prosodic inversion”, as suggested by Lambrecht (1994), which is typical of German and English subject accented sentences like “The SUN is shining.”

Thetic statements may serve various functions. They may, for instance, give background information as in weather expressions such as ‘The sun is shining’ or ‘It’s raining’. It has also been observed that thetic statements frequently contain ‘empty’ verbs, a fact that invited an explanation of the construction on semantic grounds. However Sasse (1987) argued that the superior explanation is one in terms of pragmatics or information ‘packaging’ (p. 558) and the actual articulation is dependent on the speaker’s choice (p. 521). Aspects like informational value (prior mention, situative presence), semantic weight or ‘emptiness’ of the predicate, or special semantic features of certain verbs (e.g. verbs of appearance and disappearance), as well as grammatical properties like intransitivity or unaccusativity, which have frequently been adduced as conditioning factors for subject accentuation and ‘deaccenting’ of the predicate, can be understood as possible restrictions to the actual pragmatic choice.
As far as the different functions are concerned, Sasse (1987: 566) lists seven typical domains for thetic expressions (21):

(21)
   a. existential statements
   b. explanations
   c. surprising or unexpected events
   d. general statements (aphorisms, etc.)
   e. background descriptions
   f. weather expressions
   g. statements relating to body parts

However, the idea of theticity as a linguistic category was not generally accepted in the linguistic community. In subsequent research, the issue of theticity was further explored by Sasse and a group of fellow researchers, with a special focus on VS word order (Matras 1995). In his most recent summarizing paper, Sasse (2006) identifies some problems with previous research and offers a less ambitious characterization of theticity, not as a well-defined linguistic category, but rather as a cross-linguistically comparable phenomenon. This concept of theticity is also the point of departure for the present study and will be dealt with in chapter 6.

1.2 Objective and general outline of this study

The main aim of this study is to integrate different aspects of information structure (focus, FOCUS, topic and information, i.e. the different degrees of givenness and accessibility) in order to build a comprehensive account of the prosodic encoding of information structure. Even though the empirical basis for the theoretical model mainly comes from Egyptian Arabic, frequent reference to work on other languages will ensure a wider applicability of the ideas offered here. To achieve this goal, I will proceed as follows: In Chapter 2 I will introduce the prosodic model for the description of Egyptian Arabic prosody (intonation) that is used throughout the book and discuss some theoretical issues concerning intonational phonology. Chapter 3 will then be devoted to an overview of some works of the vast body of literature on the prosody-information structure interface. The discussion will concentrate on the important early proposals as they serve as the basis of all later (mostly formal) accounts. At the end of chapter 3, I will suggest a general conceptual model of pragmaprosody as the theoretical framework for the empirical chapters to follow. Chapter 4 discusses the general question of the prosody of information structure in Egyptian Arabic, investigating pragmatic properties (information status and focus-background) as well as pragmatic relations of topic/TOPIC and FOCUS. The empirical corpus study focuses on the issues of topicality (Chapter 5) and theticity (Chapter 6). A summary of the results will be given at the end of every individual chapter. The proposals concerning the typical realization of information structural categories made in this study are supposed to hold at
least for the language investigated, Egyptian Arabic. I believe, however, that the basic ideas underlying this account are also valid as a general conceptualization of how the interaction of prosody and information structure works.
2 Prosody

Before we can take a closer look into the interaction of pragmatics and prosody, some introduction to the prosodic structure of speech, especially what has come to be referred to as intonation is called for. Section 2.1 will be devoted to this issue. Based on the view outlined in 2.1, I will present the model for the description of Egyptian Arabic (henceforth EA) in Section 2.2 and discuss some theoretical questions. The reader who is only interested in the empirical investigation of the prosody of information structure in EA may only read the synopsis of the descriptive model in Section 2.2.2.

Prosody in its widest sense is a "parallel channel of communication", i.e. parallel to what is conveyed by the lexical content (Shih & Kochanski 2002). Common to many aspects of prosody is that they are related to muscular activity. Such a view not only includes the meaning of gestures and facial expressions like eyebrow movements (Bolinger 1986; Ohala 1996), it is also applicable to the prosodic component of sign languages (cf. contributions in Language and Speech 42 (2/3), 1999; Wilbur 2000). Although I fully concur with this wide concept of prosody, I will be using the term in a narrower sense here, referring to what in some lines of research has been called suprasegmentals (Lehiste 1970), i.e. different features that are not inherent in the segmental string but are generally conceived as being superimposed upon the segmental layer, namely pitch, loudness, length and tempo. All of these features are employed in the realization of prominence, which essentially means that some part of the segmental string that constitutes an utterance stands out against a backgrounded environment and is thus made more easily accessible to the hearer. Since early studies by Fry (1955; 126-152) it has been assumed that pitch plays a major role in the realization of prominence, an idea that has been substantiated by further acoustic investigations. Beckman (1986), for example, demonstrated that pitch is the most reliable cue to prominence whether in so-called pitch-accent languages such as Japanese, or in what has been classified as a stress-accent language like English\(^1\); the same seems to be true for Spanish (Kubarth 2009). The central role of pitch to the perception of prominence can most probably be held responsible for the now common use of the term intonation to refer to more than only the tonal aspect of prosody covering accentual properties as well. Concentrating on the tonal properties of prominence has also led to a certain neglect of other cues such as intensity and durational features in much of the research within the past decades. Specifically, the autosegmental-metrical approach to the study of intonation, which combines accentual with other melodic properties, is predominantly

\(^1\) Beckman, however, concludes from the results of her study that the traditional distinction among pitch-accent and stress-accent languages is justified, having found that despite the predominant use of pitch in both languages, English differs from Japanese in that it also makes use of the other cues, such as intensity and duration which virtually play no role at all in Japanese.
concerned with the position and shape of pitch accents, which will be the topic of sections 2.2.3 and 2.2.4

2.1 Intonation: rhythm and tonal contour

Most work done on intonation today relies on the assumption that intonational systems can be described in terms of an intonational grammar consisting of intonational phonemes on a par with segmental ones. But there are also some research (cf. Lieberman 1967; Chafe 1974, 1976; Couper-Kuhlen 1986, Couper-Kuhlen & Selting 1996; Selting 1993, 1995; Vaissière 1983, 1995; (Hirst & Di Cristo 1998; Hancil & Hirst 2013; Wichmann 2000; Xu 1999, 2005; inter alia) that in some way or the other stands in the often called universalist tradition. This tradition, which is probably most prominently represented by Dwight Bolinger, is expressed in the lengthy quotation below. Such approaches have also been called functionalist as they assume a direct interaction between meaning or informativeness and the shape of pitch contours without intervening "structure". Ladd (1996: ch.6) also called them highlighting approaches as they commonly assume a strong correlation between information value and highlighting via intonational marking of informative items.

Though intonation has many ties to the central arbitrary manifestations of human language - to syntax, phonology, and to some extent lexicon - its most intimate connections are with the general scheme of iconic nonverbal communication, particularly the now spontaneous, now simulated or ritualized gestures of the face, head, hands and body. Its meanings are based on inferences from concepts of up and down - often associated with actual up-down movements of other parts of the gestural complex - plus metaphorical extensions of those concepts. Supposed grammaticizations are contradictory unless seen as intersections of two relatively autonomous systems: word-based language, and intonational and physical gesture. (Bolinger 1983: 101, emphasis in the orig.)

However, I do not believe that Bolinger’s views are necessarily in contradiction to language-specific manifestations of prosody. It just emphasises two major claims: a) the impossibility of disentangling the linguistic and the extralinguistic, i.e. grammatical and attitudinal/emotional meaning and b) the essential difference between the iconicity of intonational gestures as opposed to the clear arbitrariness of segmental phonemes. Furthermore, it regards language-specific prosodic realizations as interactions of natural prosodic preferences and the linguistic components of a language, and not as the spell-out of syntactic structure.

I follow Bolinger in the central assumption that prosody has a life of its own which is governed by (universal) prosodic processes. These processes are very simple and have their origin in general physiological and perceptual mechanisms. Prosody is thus attributed to the physical or natural aspect of language in the sense of the distinction between physis - what is natural - and thesis - what is conventional about language - made in Plato’s Cratylus (cf. Hurch 1988: 8). Such a distinction lies at
the heart of Naturalness Theory, e.g. Natural Phonology (Stampe 1973; Donegan 1978; Donegan & Stampe 1979, 1983; Dressler 1984, 1985, 1996; Hurch 1988, 1996) which considers itself a continuation of earlier ideas about phonetic alternations prevalent at the end of the 19th century, especially in the work of Boudouin de Courtenay (Hurch 1988: 5). With Bolinger (1986, et passim) I assume that syntactic and other linguistic functions of prosody cannot be captured by arbitrary phonological rules. Prosodic phenomena such as phrasing, accentuation or melody in most cases only help encoding such functions, frequently obeying the prosodic processes at work. On the other hand, pragmatics, semantics and syntax may require prosodic structures that diverge from default prosody that is determined by prosodic processes alone. But these diverging prosodies also mostly follow general iconic principles. The way linguistic factors and prosodic factors interact, however, is language-specific, and thus specific patterns arise that may become conventionalized in the individual languages. Hence, I do not believe that prosodic domains or entities are defined in terms of syntactic or semantic categories, as has frequently been assumed in generative accounts of intonation. They, however, typically coincide with such categories. Rather, rhythmic patterns are thought of as a priori independent from the segmental material, the text, they are associated with.

Such assumptions are principally in line with the idea of tune-text association that will be introduced below. As already mentioned, the manifestations of prosody have a basically natural character, obeying natural preferences. This is not to say that the prosodic structures that can be observed in different languages are totally alike. But it is mainly through their association with text that linguistic meaning is established. Thus, language specific uses of tune arise which have been claimed to exhibit arbitrary form-function relations, as in the declarative and interrogative patterns in the North American language Chickasaw (Gordon 1999) and the declarative rises of Belfast English which have been termed “unnatural” by Gussenhoven (2004: 54). Whether this is actually the case or not, it is clear that the tune-text association is language specific and therefore may lead to misunderstandings between speakers of different languages. The main idea I will follow throughout the description of prosodic structure presented here is that it is the association of a tune with text that produces those language-specific prosodic constructions. And these constructions are an interesting subject of linguistic research.

2.1.1 Tune-text association

Tune, as conceived by Liberman (1979) only refers to the tonal component of a speech melody, while rhythm is divided into two components: a) the metrical structure as a property of the text which is conceived as "the non-tonal phonological representation" (p. 27) b) the metrical grid representing rhythm as such. The metrical structure is in the early version (Liberman 1979) represented by a binary branching metrical tree depicting a hierarchical structure of relational strong and weak positions. The
upshot of this idea is to provide a system that accounts for the intuitive and empirical fact that metrically strong syllables (primary accents) are not associated with metrically weak positions of a melody. Liberman explains melody and rhythm of natural speech starting out with chant - the vocative chant and the children's chant (Figure 1) - as a special kind of prosodic pattern occurring in natural speech.

Thus the metrical structure of the text in the original theory does not specify how the prosodic make-up of an utterance actually looks like. To use the example of song or chant for illustrative purposes, metrical structure does not determine the actual temporal patterning of the chanted or sung text. For example, it has nothing to say about the specific rhythm applied. Liberman (1979: 33) identified two different time signatures for the children's chant shown below in Figure 1.

(a)

(b)

Fig. 1: Two time signatures of the children's chant (after Liberman 1979: 33)

To arrive at the actual rhythmicized pattern, another type of structure is imposed which pertains to the human inclination to impose rhythmic structure on all kinds of sequences (cf. section 2.2.1). This structure is the metrical grid which Patel (2008: 105) calls "a mental pattern of multiple periodicities in the mind of a listener" (emphasis in the orig.).

The concept of prosody I adopt here is more closely related to music. Thus, I assume that intonation, or the melody, is a combination of the tonal sequence and rhythm. Thus the individual accents of a certain tune/melody are seen as the manifestation of rhythm and as building blocks of the tonal contour at the same time. Of course, speech rhythm does not display periodicity in the same way a piece of music does. As research over the past decades has shown, absolute isochrony in speech rhythm is not empirically supported. Nevertheless, there remains the clear intuition of rhythmicity in speech. As Patel (2008: ch. 3) suggests, the concept of rhythmicity in speech has to move away from the idea of isochronous inter-stress intervals (for stress timed languages) or other linguistic domains (the syllable for syllable-timed languages) and adopt a "richer view of speech rhythm" (p. 154). On this view, one aspect of rhythm in music as well as in speech is grouping structure, i.e. the grouping of tones or words into higher level units such as phrases. The next section will deal more extensively with the concept of rhythm.

Thus, it is the interaction of the tone sequence and rhythm or the metrical component as explained below that constitutes the melody or the tune. The melody of an utterance in natural language may in
its essential behaviour be compared to the melody of a song which also consists of a tonal contour and a specific rhythm. When lyrics are set to music, there will be adaptations either of the words or of the melody itself. In the following example of a short simple melody, the rhythmic and tonal patterning can be observed. It has to be noted, however, that the examples cited here are specific to Western musical tradition and to Western listeners. As ethnomusicologists emphasize, rhythmic and tonal structures are by no means universal, but differ from culture to culture. On the other hand, Western musical structures have found their way into many cultures as an effect of globalization and the export of Western pop music to all different parts of the world. I will therefore tentatively suggest an interaction of tonal and rhythmic structures based on what is mostly known from Western musical tradition.²

| (1/4x) | . | . | . | . | . | . | . | . | . | . |
| (1/2x) | . | . | . | . | . | . | . | . | . |
| (1x) | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| (2x) | . | . | . | . | . | . | . | . | . | . | . | . |

Fig. 2: After Patel (2008: 104, 108): rhythmic and melodic patterning in a short simple melody. Phrase boundaries are indicated below the score (p1= phrase 1, etc). The metrical structure (the grid) is illustrated by the dotted lines above the score. The relative strength of a beat is indicated by the number of dots above it.

The above example makes clear that tonal and rhythmic phrasing is not completely independent. Even if they are not congruent, it can be clearly seen that certain ‘target’ tones such as the note C, the tonic of that special contour, always coincides with a strong metrical beat (i.e. three times with the strongest beat within the measure and once with the second strongest beat of the measure). Moreover, these tones terminate the tonal phrases and are mostly longer than the neighboring notes and/or followed by a pause. While tonal and rhythmic structures have been extensively dealt with in experimental studies, their interaction is still an under-researched phenomenon. Krumhansl (2000: 172) notes that

> although it is clear that pitch and rhythm enter into complex relations with one another, the precise nature of their interaction is not yet well understood. As measured in some tasks, rhythmic structures can influence the organization of pitch information, and pitch structures can influence the formation of rhythmic patterns. The evidence to date does not suggest that one aspect is psychologically primary.

To make it clear what is meant by the autonomy of prosody and intonation in our model let me now venture a comparison of how I assume that the tonal and rhythmical components interact in actual utterances. The two examples in Figure 3 are frequently occurring intonation patterns from Egyptian

---

² The comparison of the suggested prosodic structures found in speech with those in music is made predominantly for the sake of illustration. Due to the uncertain scientific grounds we are entering, the details have to be taken with some caution, but nothing essential hinges on them. The rhythmic structures in the prosody of speech suggested below are empirically founded and thus provide good means of description.
Arabic. I assume the two patterns in Figure 3a and 3b to display different rhythms. While the contour in (3a) is perceived as one rhythmic group, the contour in (3b) is subdivided into two groups with a juncture between the two accents with equal height.

![Fig. 3: Two schematized intonation contours. Panel (a) shows a downdrifting contour throughout the whole intonation phrase, panel (b) shows a downdrifting contour interrupted by pitch reset after the second accent.](image)

This pitch reset is frequently taken to indicate some sort of phrase break in most models. One type of empirical evidence for the assumption of juncture is the fact that in two peaks of equal height, the second one is the one perceived as more prominent (Gussenhoven & Rietveld 1988). Thus, the second accent may be perceived as a beat that indicates the beginning of a new rhythmic group. Crucially, I do not assume (as is often the case in intonation models) that reset necessarily induces a phrase break. The important cue for perceiving some kind of juncture here rather is the rhythmic grouping of successive units. The comparison with musical structure suggests that pitch reset and the repetition of similar sound patterns also tends to give rise to or reinforce the perception of a certain rhythm. As the result of experimental studies so far suggest, a higher pitched tone tends to be heard as the beginning of a group (Krumhansl 2000: 161). Cooper and Meyer (1960: 34) argue that the incisive unambiguous dactylic rhythm, the "heavily marked, dance-like thumping" of the musical example below (Fig. 4) is partly due to the sharp separation of "almost independent motives".

![Fig. 4: Part of the melody of "Ach du lieber Augustin" (after Cooper and Meyer 1960:34).](image)

If these observations are correct, they are additional evidence for the suggested rhythmic grouping in the intonation pattern of Figure 3b. We will later see how grouping may be used for the indication of syntactic phrase boundaries in Egyptian Arabic. It may, however, also occur in total ignorance of the syntactic and morphological boundaries. This may be seen as evidence of the independence of prosodic structure from grammatical structure. A preference for structuring four words in two phrases containing two words each, regardless of the morphological structure, has also been reported for Japanese by Kubozono (1993), cited in Gussenhoven (2004: 187). Similar facts are reported of Catalan
where eurythmic constraints override syntactic constituency and thus cause a mismatch between syntactic and prosodic phrasing (Prieto).

Let us now proceed to tune-text-association, or to be more precise, *melody*-text-association, and for this purpose return to the illustrative example of the children's chant. The observation made by Liberman is that when lyrics are set to music, some constraints are imposed on the association process. Strong syllables of the words have to be matched with strong syllables of the melody. Thus a primary word stress may not be matched with a metrically weak syllable of the melody. Let us now see how this is accomplished in the children's chant. The association in Figure 5a is infelicitous because the word stress of *Pámela* is matched with a metrically weak tone whereas the adaptation shown in Figure 5b corrects this by subdividing the only syllable of *loves*. Note that it is not the highest tone that is associated with the metrically strong syllable but the metrically strong one, the prior only serving as a *grace note* for the major melody.

![Figure 5a](image1)

**Fig. 5:** Tune-text association of a song text with a given melody (after Ladd 1996: 52ff.). Panel (a) shows an infelicitious association, panel (b) shows a felicitous association.

One of the main differences between the assumptions made here and the common view in mainstream prosodic theory concerns the metrical structure of the text. In *Autosegmental-metrical theory* (cf. 2.1.2), it is assumed that the metrical pattern is built in a bottom-up way and thus somehow emerges from the text. Contrary to that, I assume a top-down-structure that rests on *a priori* independent accentual patterns that are associated with the lexical content. Such a view implies that the rhythmic patterns are not construed within grammatical domains such as the (morphosyntactic) word or syntactic constituents (NP, VP or clause). Rather, it is the *association* with linguistic domains that applies actual prosodic patterns to the lexical content and, if necessary, alters them in conspiracy with what is needed to indicate the linguistic functions. To make clear what is meant by these rather abstract suggestions consider the following examples from Italian based on the analysis by Hurch (1996). As Hurch argues, "[t]hrough rhythmic structures like Italian [ôòóôô], [ôòô], [ôôó], or [ôôôôô] onto which words can be mapped, like *generalmente* 'generally', *civiltà* 'civilization', *caduto* 'fallen', *parafango* 'fender', or *oreficeria* 'goldsmith's shop' respectively, are perceived as rhythmically even" (1996: 87; *italics in the orig.*), not only by speakers of Italian, but also by speakers of languages with
fixed primary accent. However the crucial point is that primary accent locations in the Italian verb inflectional categories future and conditional (part-ir-ó "I will leave" and part-ir-éi "I would leave") which are derived from the infinitive form refer to a different morphological domain - the inflectional ending - than present tense indicative forms (párt-o "I leave") which are stem-derived. Given that primary accent here fulfils morphological functions, we cannot assume that in both forms the accent locations are likewise derived exclusively by prosodic processes. Contrary to that, Hurch argues that primary accent locations frequently go against what might be expected if default prosody was at stake, and that such deviant accents more readily fulfil a morphological function. Such primary accents thus are arrived at by morphonological rules rather than by phonological processes which is what a bottom-up analysis from the smallest units up to higher levels suggests. Hurch goes on to argue that secondary accents as exemplified by the Italian word patterns given above, however, are governed by prosodic processes. Given that secondary accent locations may vary and words may select different prosodically well-formed patterns whereas primary accent location does not vary, cf. èlettricitá vs. elètricitá, gènerìcaménte vs. genèricaménte, etc., it follows that the hierarchical bottom-up derivation generally assumed in metrical models does not hold for Italian word accentuation patterns. Hurch concludes that "it is the primary accent which, if we have a dependency relation at all, determines (or at least influences) the location of secondary stress and not vice versa" (p.84). It is not possible here to go deeper into the argumentation, but the main point should be clear by now; for a more detailed argumentation cf. Hurch (1994, 1996) and Donegan and Stampe (1983).

On a par with the concept of morphoprosody just outlined, I will assume a level of pragmaprosody that essentially follows the same basic tenets of an independent prosodic structure and pragmatically governed selection and configuration of prosodic patterns (see Chapter 3.3). Being the main objective of this book, this prosody-information structure interaction will be discussed in detail for Egyptian Arabic in the empirical chapters (4-6).

2.1.2 Rhythm and accent

The surface of the ocean responds to the forces that act upon it in movement resembling the ups and downs of the human voice. If our vision could take it all in at once, we would discern several types of motion, involving a greater and greater expanse of sea and volume of water: ripples, waves, swells and tides. (Bolinger 1972c [1964]: 19)

So far, it has been tacitly assumed that (pitch) accents are the manifestation of rhythm. In this section I will further elaborate on this assumption. If accents are understood as prominence, which is the
approach I am taking here, they can be conceived as the building blocks of rhythm. The concept of accent (or stress) is notoriously complicated and its treatment will be delayed to the second part of this section. Let us start therefore with the natural preferences that govern rhythm.

Rhythm is defined by Fraisse (1982: 150) as “order in movement” and its perceptual quality is related to the ability to predict what will follow on the basis of what has been perceived. Thus, anticipation seems to be one of the most important characteristics of rhythmic perception, while absolute periodicity or isochrony obviously are not a necessary condition. Many experiments have pointed to the lack of correspondence between physical and perceptual time measures between individual events in rhythmic structure, especially in speech (Fraisse 1982: 154; Martin 1972: 499; Lehiste 1977; cf. Heliel 1976 for Egyptian Arabic). Thus, the (physical) isochrony hypothesis can be regarded as empirically falsified. Patel (2008: 159), the most recent and comprehensive study on the relationship between music and speech, suggests “abandoning a fixation on periodicity” and concentrating on non-periodic aspects of rhythm in language.

Rhythm is pertinent to human cognition in general. It has often been noticed that humans tend to impose rhythmic structure on all kinds of events (heartbeat, the ticking of a clock), grouping them into smaller subunits (Lehiste 1970: 7-8, 155; Martin 1972: 487; Fraisse 1974: 169-171, 1982: 168; Auer & Uhmann 1988: 215). The experiments conducted by Fraisse showed that listeners group identical sounds into subunits by twos or by threes (Fraisse 1982: 155). As nothing objectively specifies this grouping, it is referred to as subjective rhythmization by Fraisse. Spoken language, being a type of sound system, calls for a definition of rhythm in terms of sound patterns such as the one offered by Patel (2008: 96), who views rhythm as “the systematic patterning of sound in terms of timing, accent, and grouping”. In the phonetic literature references to rhythm in language have usually focused on the isochrony hypothesis and the distinction between stress-timing and syllable-timing (Pike 1945; Abercrombie 1967; Lehiste 1977; Dauer 1983; Auer & Uhmann 1988; Grabe & Low E.L. 2002) and are thus primarily concerned with syllable and foot structure, hence with what I will call ‘low-level’ rhythm. Another type of rhythmic behaviour is the one of accents in an utterance, exhibiting rhythm on the post-lexical level that is only indirectly related to lexical stresses that offer the structural basis for accent. Such a distinction has also been suggested by Bolinger who distinguishes between syllabic rhythm, which may be viewed as a low-level phenomenon that operates on the level of syllables and feet, and accentual rhythm that may be dubbed high-level, and which operates on the level of the utterance, that is on a level where pragmatic meaning comes into play (Bolinger 1986: 14-23, 37-73).

It is this high-level rhythm we are concerned with here. What immediately strikes our ears when listening to speakers of Italian or Egyptian Arabic, is the frequent up and down of speech melody that evokes the impression of strong rhythmicity. I have suggested elsewhere (El Zarka 2005) that there might be a typological distinction between strict-rhythm and non-strict-rhythm languages, a distinction
that is relevant for the general conception of intonation and its relationship to information structure as outlined below. Strict rhythm is instantiated by frequent and – at least impressionistically – isochronous (pitch) accents being the instantiation of that rhythm.\(^3\) If we take a look at pitch tracks of utterances in these languages, we see a declining line of ups and downs, referred to as "the typical see-saw" by Mitchell (1993: 222) that frequently characterizes the whole utterance.

In summarizing, we may say that whether a language is stress-timed or syllable-timed and whether it has a clearly defined word-level stress like English, German or Egyptian Arabic or whether it probably only exhibits phrase-level stress as has been claimed for French by some authors (e.g. Di Cristo 1991: 196) or for Moroccan Arabic (Maas 2011: 53), post-lexical higher is a phenomenon present in all languages. Thus, some kind of eurythmicity is also incorporated into Gussenhoven’s analysis of French (Gussenhoven 2004: ch.13; based on the analysis by Post 2000) as constraints on the distribution of accents. Gussenhoven (2004: 253) mentions that pitch accents are relatively frequent in French and that there is usually even more than one pitch accent per phrase. This suggests that the difference between languages noted for the ‘low’ lexical level might vanish as soon as we reach the ‘higher’ utterance level rhythm.

2.1.2.1 Rhythm in a metrical framework

Rhythm involves systematic patterning and grouping which are interrelated aspects. Grouping in prosodic articulation is at least to a certain extent constrained physiologically by the finiteness of the air stream available which defines the upper limit of an utterance to be produced in one go and thus has an influence on the grouping of the speech flow into intonation units or intonation phrases (IP) which due to their physiological basis have also been called breath groups (Liebermann 1967). The other, more important, aspect of metrical structure is truly rhythmic in nature and pertains to the recurrence of prominence in time. Metrical phonology, which is one of the fundamental pillars of the Autosegmental-Metrical Theory (henceforth AM Theory), the mainstream theory of intonation today (cf. Ladd 1996: 45-59), represents these prominences as suprasegmental features and describes them in relational terms, representing word level and utterance level stress patterns in a hierarchy of prosodic constituents (Liberman & Prince 1977; Prince 1983; Selkirk 1984; Nespor & Vogel 1986). All languages exhibit some kind of rhythmic grouping and obey some version of the Principle of Rhythmic Alternation (Selkirk 1984: 52; Liberman 1979 [1975]; Liberman & Prince 1977; Selkirk 1980, Prince 1983; Hayes 1981, 1995), According to this principle, stress clashes and stress lapses (more than two unstressed syllables) are dis-preferred. A rhythmic rendition of utterances will thus

\(^3\) Note that Bolinger (1986: 67) assumes a “mild tendency toward accentual isochrony”.

38
give rise to the emergence of binary or ternary feet (trochees, dactyls)\(^4\) (Bolinger 1986; Hurch 1996; van der Hulst 1999) producing so-called ‘secondary accents.’ Utterance level rhythmic organization – Bolinger’s accentual rhythm – is mostly attributed to phrasing and the distribution of pitch accents per phrase, which is attributed to rules or constraints in generative models (\(\phi\)-restructuring rule (Nespor & Vogel 1986), Bin(MaP) constraint (Selkirk 2000), Hammock constraint (Post 2000; Gussenhoven 2004; cf. also Truckenbrodt 1999 for an overview).

Languages may have binary (1a) or ternary (1b) rhythm (van der Hulst 1999) or both.

\[
\begin{array}{cccccccc}
\text{x} & \text{x} & \text{x} & \text{x} & \text{x} & \text{x} & \text{x} & \text{x} \\
\sigma & \sigma & \sigma & \sigma & \sigma & \sigma & \sigma & \sigma \\
\end{array}
\]

\[
\text{(1a)}
\]

\[
\begin{array}{cccccccc}
\text{x} & \text{x} & \text{x} & \text{x} & \text{x} & \text{x} & \text{x} & \text{x} \\
\sigma & \sigma & \sigma & \sigma & \sigma & \sigma & \sigma & \sigma \\
\end{array}
\]

\[
\text{(1b)}
\]

On word level, the two main functions of accent are the culminative function and the demarcative functions. Culminativity makes the individual accented items stand out against their surroundings. As every content word is accented in the default case, this allows the listener to identify the main lexical items in a sentence. Demarcativity marks the boundaries of the individual items and thus indicates where one item begins and another item ends. This function is especially important in word prosody. Many languages have predictable stress patterns that allow the listener to parse the speech flow accordingly and uniquely identify word boundaries (Hyman 2006: 234).

But the demarcative function is also valid in larger domains such as intonation phrases which frequently exhibit the strongest prominence at the end or at the beginning, and in fact to a certain extent at both ends, one end frequently being stronger than the other. We may thus construct schematic grids on that basis showing edge marking and rhythmic preferences (2). The lowest level of the grid is parsed into falling trochaic structures in both panels, we may term it the foot level. In panel (2a), the strongest prominence or the head of the structure is on the left, and in panel (2b) it is on the right side.

\[
\begin{array}{cccccccc}
\text{x} & \text{x} & \text{x} & \text{x} & \text{x} & \text{x} & \text{x} & \text{x} \\
\sigma & \sigma & \sigma & \sigma & \sigma & \sigma & \sigma & \sigma \\
\end{array}
\]

\[
\text{(2a)}
\]

\[
\begin{array}{cccccccc}
\text{x} & \text{x} & \text{x} & \text{x} & \text{x} & \text{x} & \text{x} & \text{x} \\
\sigma & \sigma & \sigma & \sigma & \sigma & \sigma & \sigma & \sigma \\
\end{array}
\]

\[
\text{(2b)}
\]

In a long utterance, the strong prominences at the edges leave us with ‘lapses’ on the higher grid levels, which invites internal rhythmic grouping comparable to beat in music. In speech, grouping parses out entities smaller than the intonation phrase (IP), which have been called phonological phrase (Selkirk 1984; Nespor & Vogel 1986) or prosodic phrase (PP), will be called here. The metrical

\[^4\] Some metrical models (e.g. Hayes 1995) only allow for binary structures, an assumption that I will not follow.
representation of the schematized intonation contour in Figure 3b above is given in (3). It contains three prosodic phrases consisting of two feet each. It may optionally have a strongest prominence at the right edge or the left edge of the whole structure as indicated by the parentheses.

(3)

\[
\begin{array}{cccc}
(x) & (x) & (x) \\
\sigma & \sigma & \sigma & \sigma \\
\sigma & \sigma & \sigma & \sigma \\
\end{array}
\]

This type of strict rhythm, which is characteristic of music, is of course not normally a feature of ordinary speech. In language, rhythmic structuring is especially pertinent to certain genres such as nursery rhymes (Maas 1999: ch. 4; Heliel 1976). Such pieces of language as shown in the example of an Egyptian counting-out game (4) help illustrate the rhythmic essence of linguistic behaviour in general, being situated at the one end of the linguistic continuum that involves only minimal information, but maximal social interaction.

(4)

```
HA:di BA:di / SI:di mhAMmad / IL-bayDA:di / fA:l-u w hAT s-t-u / KULi-u ʕaLA:-di \\
hadi badi / mister M. / the-Bagdadi / took.away-it and added-it / all.of-it on-this \\
‘Hadi-badi Mr. Muhammad from Bagdad took it away and put everything on this (side)’
```

The example shows how language has to adapt to the strict rhythm of the rhyme consisting of five phrases with two trochaic feet each, except the last one that has. To make the words fit the rhythmic structuring elisions are resorted to, such as the deletion of /u/ in the proper name muhammad which is syllabified together with the first word as si:.dim.ham.mad. In the third and the fifth phrase there are accents where they are not normally expected, i.e. on the definite marker il- and on the second vowel of ʕala ‘on’ which is irregularly lengthened for that purpose. If at all, this function word would be accented on the first syllable in natural speech.

The pattern is not totally in line to the semantico-pragmatic aspect of the sentence. If it were not for the rhythm, the last accent should go on the semantically richer, more informative di “this” which may even be interpreted as contrastive. The rhythmic pressure for a falling at the end of the utterance is partly responsible for the unexpected accent on the preposition and for the formation of a prosodic word ʕala:-di “on-this” which could then regularly be accented on the penultimate long syllable.

Although nursery rhymes admittedly belong to the poetic genre more than they belong to ordinary speech, linguistic adjustments to prosodic regularities as the ones demonstrated above are commonplace in everyday speech, too. Thus, prosody not only adapts to the requirements of syntax and semantics or pragmatics, but prosodic preferences, such as the Principle of Rhythmic Alternation, also have their say in the actual shape of tunes. Well-known examples for the shaping capacity of a rhythm is the Rhythm Rule or Iambic Reversal (Liberman & Prince 1977; Hayes 1984, 1995) which
retracts a primary accent to the site of a secondary stress in order to avoid stress clash as in English *thirteen mén* instead of *thirteenth mén*. But stress clash is not the only condition leading to rhythmic adjustments as extensively illustrated by Bolinger (1986: ch.5). Other factors are a certain pressure for isochrony and the strengthening of beginnings which putatively are both at work in the utterance *academic discipline* instead of *académic discipline*, or the protraction of the accent towards the end of an utterance to strengthen the right edge as in a *substantial increase* not *increase* (Bolinger 1986: 77).

I have also frequently observed this tendency in German when readers who are not well trained in reading tend to put the strongest or an equally strong accent on a final function word instead of the prefinal content word that should be the more prominent one. Bolinger mentions functional effects of the above phenomena, calling the retraction of an accent to the beginning “thematic-rhematic repulsion” (Bolinger 1986: 62) and the protraction of a final accent to the end “climax” (p. 75). Climactic accents as in *absolutely* or *necessarily* are said to render the adverbs more emphatic (p. 78).

While final accents as a characteristic of monotonous reading style in German, where the reader seemingly pays little attention to the semantics of the text, are a reflection of a rhythmic preference of edge strengthening, climactic accents as a functional device may also be observed. In the following example (5) from a scientific talk, the speaker uses a final accent on a function word as an indicator of continuity in addition to the tonal rise. The accent is placed on the auxiliary in a relative clause that constitutes part of the sentence topic. Having chosen a climactic marking of continuity as a rhetorical device, the speaker signals that the predication on that topic is going to follow and thus arouses the interest on the part of the listeners in what is going to come next.

(5) overheard example

.... die Entwicklung, die im 20. Jahrhundert begonnen hat, ....

.... the development REL in 20th century begun has

In the present model, a fundamental distinction is drawn between prosody proper, i.e. what relies on purely prosodic principles, such as the rhythmic preferences discussed here, and the functional exploitation of prosodic phenomena in the creation of meaning, for example in the construction of grammatical categories where prosody and morphology, syntax or information structure collaborate.

So far we have looked at accent as an instantiation of rhythm, i.e. we have only thought of accent as prominence in time. However, what causes theoretical problems in prosodic theory is the fact that accents are not only part of the rhythmic component, but also an important building block of the tonal one. To understand the intricate interplay between prominences and melodic contour, I will briefly discuss the phonetic and phonological notion of accent in the following subsection in a kind of excursus, as I will need to clarify the foundations for the argumentation to follow.
2.1.3  *Excursus: accent and stress*

Three perceptually based prosodic features have been mentioned above: pitch, loudness and length, the acoustic correlates of which are fundamental frequency (F0), intensity and duration. Add to these segmental features such as elisions, vowel reductions, and vowel quality, the acoustic reflex of which is spectral composition. These features collaborate in a not yet fully understood way to produce prominence.

There is hardly a concept in the science of linguistics that is characterized by the same terminological confusion as the concepts of prominence, accent or stress. It is therefore necessary to clarify the terminology adopted here from the outset. The problem, however, is not merely a vague use of terminology. Differences in terminology mostly also point to different underlying concepts that may have far reaching consequences for the conception of prosody as a whole. For a start, let us define the least controversial of these terms: *prominence*. Prominence refers to the impression of the hearer that some part of the speech flow is standing out, that it is stronger than other parts and promoted to the foreground. The notion of prominence, however, has not been as much debated as the terms *accent* and *stress*, which are often used interchangeably to denote the same thing, a fact that has led to real terminological confusion. I am not going to review all the different uses of these two terms here but rather concentrate on two important uses that have direct repercussions on the theory of intonation. The first one I will discuss here and return to later on in some detail is Bolinger’s notion of *stress* as an abstract lexical capacity for being accented, while *accent* is the actual physical realisation of that potential. That means that under this view stress itself has no acoustic correlates as such but is merely a potential location of such correlates of prominence to occur.

The other view, common in AM theory, conceives *accent* or rather ‘pitch accent’ as an intonational category and as a prominence-cueing\(^5\) “local feature of a pitch contour – usually but not invariably a *pitch change*” (Ladd 1996: 45). It has to be noted that pitch accents are not understood as correlates of prominence in the first place but as “building blocks of pitch contours”, i.e. as parts of intonational tunes, thus belonging (exclusively) to the tonal component. If we take the distinction between *prominence-cueing* and *prominence-lending* at face value, this amounts to suggesting that F0-movement itself does not serve to lend prominence to the syllable it occurs on, but only indicates the location of the perceived prominence. The other view, held by Bolinger, the IPO research group (e.g. t’Hart et al. 1990) and, for instance (Campbell & Beckman 1997), conceives pitch as an actually prominence-lending feature. Such a view implies that pitch (F0) is one correlate of prominence. While it is by far the most important one in some languages (e.g. *pitch accent* languages such as Japanese),

\(^5\) Ladd (1996: 50) uses the term prominence-cueing, which was suggested by Francis Nolan, instead of prominence-lending to emphasize the view that accent is part of the intonational component and not a manifestation of lexical prominence.
prominence also has other important correlates in what since Beckman's seminal study (Beckman 1986) has been called stress accent languages. Starting with Fry's pioneering work (Fry 1955, 1958), experimental studies have clearly shown that perceived prominence corresponds to a complex admixture of acoustic cues. The way the stress-accent relation is dealt with in AM theory is a theoretically driven 'teasing apart' of these cues, based on the assumption that pitch (accent) is somehow superimposed on already physically concrete prominence brought about by other features such as intensity and duration. To be sure, the answers provided by AM theory have been rather vague on this issue. One severe problem with this theory is the conflation of stress as an abstract structural property (on the lexical and post-lexical levels) and the concrete manifestation of acoustically cued prominence by all features other than pitch. As Ladd (2008: 53) puts it "[a]ctual prominence in an utterance is signalled by a complex of phonetic cues that reflect greater force of articulation and possibly rhythmic regularity. Pitch accent is an additional feature that is part of the intonation system." It is difficult to see how the "addition" of pitch should not add additional prominence. Likewise, this view is difficult to reconcile with the fact that the excursion size of a pitch accent adds to its prominence (Gussenhoven et al. 1997; 2002, 2004: 85). Baumann (2006: 11), for example, states that the "difference between stresses and accents entails a difference in the strength or degree of (postlexical) prominences". This is contradictory to the claim that the strongest stress of an utterance inevitably attracts a pitch accent, to the extent that this is supposed to mean that its prominence is already the strongest one in the utterance.

On such a view, stress is a physical correlate of prominence, but also a structural feature. Utterances have ‘stress patterns’ reflecting syntagmatic prominence relations represented in a metrical grid (or tree) (= metrical strength, i.e. an abstract property) (Hayes 1995; Ladd 2008: 58). The greatest prominence is assumed to be related to what is mostly called sentence stress. This highest degree of stress (i.e. the concrete property?) inevitably attracts a pitch accent, i.e. it inevitably has a tonal correlate. Lower prominences or stresses are not necessarily associated with pitch accents, but they may nevertheless be perceived as prominent. On the other hand, they also may be associated with pitch accents which are supposed to be secondary linguistically (i.e. less important) and phonetically (i.e. less prominent). Ladd (2008: 54) summarizes the AM view on the stress-accent relation as follows:

> It assumes that utterances have a 'stress pattern', which may involve several different degrees of perceived prominence. This stress pattern reflects a set of abstract prominence relations between the elements of the utterance. The stress pattern is manifested in a variety of phonetic cues, which are admittedly not well understood. In addition to the stress pattern, there is an intonation pattern for the utterance which is composed of a string of pitch accents and edge tones, and the pitch accents are 'lined up with the text on the basis of the prominence relations' (Pierrehumbert 1980: 102).

---

6 This assumption is reflected in Ladd's adoption of the idea of prenuclear accents from the British intonation tradition (Ladd 2008: 281-7, et passim) and the treatment of postnuclear accents as secondary association of a phrase tone (Grice et al. 2000).
Recent work on the stress-accent question has suggested that word stress may be identified in stress-accent languages in the absence of a pitch accent (Sluijter 1995 for Dutch and American English; Sluiter & van Heuven 1996 for Dutch; Ortega-Llebaria & Prieto 2006 for Catalan and Spanish). The acoustic correlates that were held responsible for the perceived prominence are duration and spectral balance. This fact was interpreted that these features are correlates of (word) stress, while F0 and overall intensity are correlates of pitch accents. Moreover, pitch accents are not always realized as pitch change. It has been shown that the clear difference in the position of a pitch fall in the English minimal pairs *pérmit* (noun) and *permit* (verb) is only relevant in declarative utterances, whereas interrogative intonation changes the F0 contour and the distinction is much less salient (Ladd 2008: 50; Gussenhoven 2004: 18). The conclusion from these facts drawn by some proponents of the AM theory is that stress is “not merely a lexical abstraction” as in Bolinger’s framework but has a “separate phonetic reality of some sort” (Ladd 1996: 50; 2008: 53).

However, some of the recent studies investigating the acoustic correlates of stress in the absence of pitch accent, have yielded opposite results (Campbell & Beckman 1997 for American English; Sadeghi 2007 for Persian). The reason for the contradictory findings may be due to language-specific differences. However, the findings are also contradictory for American English.

Whatever the exact phonetic correlates of prominence are, I do not think the conceptual distinction between stress as a word-level prominence and intonational pitch accent as a component of the tonal level to be warranted. On the view adopted in the present study, accent is the concrete realization of prominence and a word stress that is highlighted by pitch, is mostly simply less prominent than one that exhibits a rapid pitch change on the stressed syllable. This view regards prominence as a one-dimensional perceptual phenomenon and accent as the phonological category that may have a number of different phonetic correlates, the most important of which is pitch. This view also implies that accentuation is gradual and not categorical (cf. 2.1.4). If we look at what Bolinger in his later work subsumed under the cues to accent (1986: 15-21) we find length, loudness, rhythm, vowel quality but also more rarely mentioned features such as “delayed release” (onset lengthening) and “breathiness” (p. 19). It is important to recognize that Bolinger’s concept of accent can be interpreted to include other cues than pitch as well, even though he in practice admittedly neglects the other features viewing pitch as the most important and efficient one. For Bolinger, accent with all its manifestations, is the concrete physical property – the realisation of an abstract potential in real speech, whereas in AM theory accent, being exclusively tonal, it is only partly responsible for the perceived prominence.

Furthermore, the different phonetic correlates are difficult to disentangle. F0 and intensity frequently go hand in hand as stronger sub-laryngeal air force is apt to increase both intensity, which creates more loudness, and fundamental frequency via a more rapid vibration of the vocal folds, which in turn yields higher pitch. Even though duration is perhaps slightly more independent, it is also a frequent
concomitant feature of high pitch excursion. In view of these facts it seems problematic to distinguish between accent as a pitch phenomenon and stress as an ill-defined phonetic amalgam of features in addition to its abstract property of offering a potential for being accented. Consequently, Hyman (1977, 1978) also argued that

word-accent comes from intonation in exactly the same way that word-boundary phenomena come from utterance-boundary phenomena... there are no clear phonetic distinctions between those features said to characterize (primary) stress, and those said to characterize intonation. The pitch, duration and intensity variations correlating with stress also correlate with the high points of intonation...” (Hyman 1978: 453)

2.1.4 Accentuation

Another problem that arises under the Standard AM view is the existence and recognition of monotonal (pitch) accents. Thus a low monotonal accent occurring in a low monotonal environment is only reluctantly accepted, while on the other hand high monotonal accents in a high monotonal environment are readily acknowledged. That means that a prominent low and a prominent high are not treated equally in the intonational representation. This is especially common in the treatment of low-level stretches of speech, which are not accompanied by any pitch obtrusion that follow a strong pitch movement earlier in the utterance (Figure 6). Putative prominences here are only reluctantly referred to in AM theory.

Fig. 6: Low duration accents in German, from Baumann (2006: 24).

The above example in Figure 6 is analysed by the author as involving two low accents L(\(^*\))\(^7\), which, following Standard AM theory, are derived by secondary association of a phrase tone to a stressed syllable (Grice et al. 2000). This theoretical decision was resorted to in order to account for post-nuclear prominences (accents) which, according to the theory, are not supposed to exist. This means that even though there is no pitch movement whatsoever, prominence is still represented in the tonal string and no reference is made to other cues of prominence. Although Bolinger acknowledges other cues to accent he also treats stretches without any pitch obtrusions as ‘deaccented’. An approach that makes explicit reference to the different features involved in the production of prominence is the

\(^7\) The parentheses indicate secondary prominence.
model developed by Kohler (Kohler 1991, 2005). Kohler (2005: 99) calls this type of prominence \textit{duration accent} and notes that it is less salient than a pitch accent and was therefore classified as partially deaccented in the original version of the model (Kohler 1991). Interestingly, the \textit{duration accents} mentioned above occur towards the end of phrases, which might justify an analysis that derives them from original phrase accents. However, low level accents also occur as phrase-final nuclear (i.e. most prominent) accents, for example the final low level accents in German (Essen 1964; Isačenko & Schädlich 1970; Wunderlich 1988; Grabe 1998a; Grice & Baumann 2002) and in Egyptian Arabic (Ali-Askar 1992: 235; Rastegar-El Zarka 1997: 252). In terms of the phonetic realization of prominence this fact may be related to what the results of perceptual experiments conducted by Fraisse (1956) suggest, namely that initial strengthening (high intensity and high pitch) at the beginning \textit{(rhythmitisation intensive)} is paralleled to final lengthening \textit{(rhythmitisation temporelle)}, which explains the fact that final low level accents may be perceived as (most) prominent. That the low level accent is part of the tonal system at the same time is evidenced by its distinct intonational meaning (von Essen 1964; Grice & Baumann 2002).

From these facts I conclude that accent is only one linguistic dimension. Accents may be more or less prominent, their prominence being caused by an admixture of various phonetic features that may vary according to the language and even the individual speaker. Moreover, the interplay of the individual phonetic features is dependent on the strength of the prominence and on the position within an utterance or a smaller domain. In the present model, accent is a phenomenon shared by the metrical and the melodic component, thus the accent specifies intonation proper and rhythm at the same time, signalling prominence and shaping the melody at the same time.

A final issue pertains to the ability of a hearer to identify stressed syllables in the absence of any phonetic cues might be due to the rhythmic organization we normally impose on a stretch of speech, thus we perceive syllables as somehow prominent (Bolinger 1986: 17). Especially native speakers may also use their knowledge of the position of lexical stresses to imagine accents where there actually are none. It has to be noted that this anticipation of an accent points to the abstract property of the metrical component. As Ladd (2008: 57f.) also argues, this fact is comparable to music where we have on-beat and off-beat rhythm. Just like syncopation in music we may also find such deviations from expected prominence positions in the abstract grid that is created by the functioning of prosodic preferences and grammatical functions. But these deviations from the expected usually have a \textit{function}. One classical example is the expression of narrow FOCUS where one accent is much more prominent than the others, especially the post-focal one. Another example is double accentuation, which exploits the anti-
rhythmic strategy of accent clash. In (6), illustrated in Figure 7 the actual grid deviates from a
rhythmically well-formed one that would only predict an accent on the penultimate syllable of *γajru*.  

\[(6) \text{ IBRAH-B-14} \]
\[
\begin{array}{llllllllllll}
& x & (x) & x \\
\times & x & x & x & x \\
\times & x & / & x & x & x & x \\
\end{array}
\]
\[\text{a::: } \gamma a j r - u / \text{ muʔahhali:n} \]
\[\text{ah } \text{NEG-NOM } \text{qualified-PL.M.GEN} \]
\[\text{LH}^{*}\text{!L}- \text{L}-\text{LHL}^{-}\]

The actual realization, however, has an accent on the final syllable as well. Although the second accent
is almost monotonal, its duration intensity and the spectral properties render it fully prominent, if
perhaps a little less prominent than the first one. The tune similar to the vocative chant which also for
emphatic purposes involves a second downstepped accent, placed on an additional syllable, if there is
one available, e.g. Ábernáthy!, Ánná! (Liberman 1979: 30ff., Beckman & Pierrehumbert 1986: 276f.;
Gussenhoven 2004: 313ff.). AM theory generally deals uneasily with pitch accents that do not
associate with metrically strong syllables. Ladd (1996: 129-131) cites examples from Italian that are
completely parallel to the EA cases and concludes that "we seem to be dealing with a general strategy
for increasing the emotional content of the contour by suspending the normal constraints on tune-text
association." (Ladd 1996: 131, emphasis added)

Another meaningful deviation from expected grid position is exemplified by a phenomenon that I will
dub *intonational syncopation*. Rastegar-El Zarka (1997: 336) has shown that in EA formal speech, the
main accent in a prosodic word is frequently shifted from the expected primary stress position in the
content word to the preposition which according to Watson 2002 and Hellmuth 2006b is unaccetable.
An example is (7), illustrated in Figure 8.

---

8 The notational conventions will be laid out below (2.2.4.1.1 and 2.2.4.2.2).
Parallel to syncopation in music, the actual grid deviates from the expected form, i.e. the actual syntagmatic prominence relations run counter expectations and overwrite the default form.

Fig. 8: Example of an early peak with prominence shift in (from Rastegar-El Zarka 1997: 336).

This complete main prominence shift is much less frequent than another phenomenon that is strongly related to *syncopation*, as observed by Rastegar-El Zarka, namely the occurrence of an *early peak* of the fall, most frequently realized on the syllable immediately preceding the accented one. This type of contour has also been observed in German news reports by Féry (1993: 103ff.; Fig. 9a) who notes that these early peaks frequently occur in pairs – the same is true for the EA examples (Fig. 8 and 9).

Both EA examples occur at the end of a paragraph and statements that somehow say the last word, sum up or complete the whole issue that has been talked about. In the utterance in (8) (Fig. 9a), the high tones are associated with the metrically weak first syllables of both words *ʔiθ* and *ta*, respectively, whereas the lexical accent is on the heavy syllables *baːt* and *waː*, respectively. The perceptual prominence shift in (7) can be related to acoustic cues: the steep fall immediately after high tone on the accented syllable, the flat contour and – quite conspicuously - the lower intensity of the lexical (cf. Figure 9). On the other hand in the case of the early peak without prominence shift in (8), the fall of the accent is as expected spanning the word-level stressed syllable.

---

9 Annotation according to the present convention has been added.
The point I wish to make here is that the relation between these two phenomena is no coincidence. The tonal information involves in both cases an early starting fall which in turn involves downstep of the metrically strong syllable signifying "finality" and "matter-of-fact". In the early peak cases only the tonal peak is shifted and thus dissociated from other prominence cues. In the syncopation cases the whole prominence is shifted, in which case tonal and non-tonal accent cues are no longer dissociated. But I assume that this is a gradual process and judgements as to where the strongest prominence is located will probably differ among native speakers. As Ladd (1996: 172) also observed, "it is not a straightforward phonetic task to identify the most prominent word or syllable". He refers to the tonally very similar, but metrically different question contours of Italian ‘Dove vai?’ and Romanian ‘Unde mergi?’ (“Where are you going?”) (p. 173). While the Italian contour involves a (stronger) prominence on the verb (analysed as H* H*+L L% by Ladd), the Romanian contour has only one prominence on the question word (H* L L%). Ladd acknowledges, however, that Italian speakers are uncertain as to which word bears the main accent (p. 174). The striking similarity of this case and the EA examples in my opinion is strong evidence for a view that does not regard (pitch) accent and nuclear accent placement as a structurally governed categorical phenomenon, but rather as a gradient matter with clear cases on the extremes of the continuum.
In view of all the facts considered in the present subsection and for reasons that will become clear later on, I adopt Bolinger’s distinction of accent as the actual realisation of prominence and stress as a lexical abstraction to which I will mostly be referring as ‘metrical strength’ in order to avoid confusion. The other use of ‘stress’ elucidated above is dealt with under the notion of ‘syllabic rhythm’ in Bolinger’s framework (Bolinger 1986: 37ff. and 63ff., Bolinger 1981; cf. section 2.1.2). In the present book, any clear prominence will be referred to as accent whether it is cued by a pitch change or not. As already mentioned, it is, however, not always clear whether a syllable actually is accented or not. But this is implicit in the assumption that accentuation is a gradient rather than a categorical phenomenon. This leaves us with accent as part of the metrical component with its tonal shape being determined by the tonal component which is the subject of the following subsection.

2.1.5 The tonal component

Having extensively dealt with the metrical component of tune, it is time now to proceed to the second component, the melody or what has also been called intonation proper. Similar to stress and accent, the term intonation has also been used to denote different concepts, a restricted one which only pertains to speech melody, i.e. to the tonal make-up that is not primarily a prominence marker, and a comprehensive one that comprises both components: accentual prosody and melodic prosody. As already explained in section 2.1.2, I take accentual patterns to be the phonetic realisation of rhythm, which is one essential ingredient of intonation. If not marked otherwise, the use of the term intonation in this book will mostly be in that latter holistic sense, while rhythm and melody will be used to refer to the individual components. Occasionally, however, I will talk about intonation (proper) in its restricted meaning, especially with reference to intonational functions, such as the marking of speech act, sentence type or information structure, as well as attitudinal and emotional aspects. In the present study I will employ three terms: contour, accent and trendline. I will use the term contour (or synonymously with it configuration) as a cover term for the other two concepts. An accent is a contour that involves exactly one prominence. An accent is thus related to the prosodic word, and one accent stretches out from the prominent syllable of one prosodic word to the accented syllable of the next. A contour that covers more than one accent will be called trendline.

2.1.5.1 Rises, falls and the biological codes

Melody or the tonal aspect of intonation relies on its physical correlate, the fundamental frequency (F0), pitch being the perceptual effect it produces. Melody also is constrained physiologically. A stronger subglottal pressure at the beginning of an utterance is apt to produce higher frequencies than the fading out air stream at the end, thus producing a declining continuum of pitch to the bottom of a
speaker’s range. Therefore it is natural for an utterance to begin high and to end low. If we combine this behaviour with the stronger metrical positions at the beginning and end of an utterance we arrive at a contour that is essentially rising rapidly at the beginning – which lends prominence to the item this rise is associated with – and falling at the end. As energy and therefore intensity and fundamental frequency is naturally higher at the beginning and lower at the end, the beginnings of intonation phrases are rather strong. This strong prominence at the beginning is, however, opposed by the lengthening of the final part of a phrase which also produces prominence (cf. the above mentioned observation by Fraisse (1956)). This would yield a contour that interpolates between the first high peak and a low end. A common pattern, however, which is extremely frequent in many languages, e.g. the West-Germanic languages English, German and Dutch, is the hat pattern (Fig. 10a), as it was called by Cohen and t'Hart (1967). The hat pattern has another movement at the end of the contour which is falling. This also lends additional prominence to the falling part, as the fall is no simple fade-out, but rather a voluntary downward movement. Another variant of the pattern is the second one shown in Figure 10b called "suspension bridge" pattern by Bolinger (1961, 1986: 47).

Both patterns conform to the principle of end weight and demarcate an IP with prominences at both edges. As in natural speech utterances, i.e. intonation phrases, are predominantly short and do not express more than one idea (cf. Chafe's (1994) one-new-idea constraint), these patterns are expected to be very frequent in spoken language.

In longer phrases we may add to this the preference of rhythmic alternation ending up with a phrase that consists not only of a rise at the beginning and a fall at the end, but of several rises and falls – a series of pitch accents (Figure 11). A typical pitch accent consists of an upward movement to a peak and a subsequent downward movement as a manifestation of the Gestalt principle of 'figure against ground' (Bolinger 1986: 15). This default pitch accent also corresponds to the physiological mechanism of tension and relaxation, which nicely ties in with Hyman's observation that lexical accent comes from intonation, as cited at the end of Section 2.1.3. A series of (uniform) pitch accents will convey a strong rhythmic impression to the hearer. Such intonational tunes are very common among the languages of the world, e.g. Danish or Swedish. But they can also be found in many other intonation-only languages such as English or German and in Spanish varieties (Prieto et al. 1996; Sosa 1999; Kubarth 2009), although it may have a more marked status in some of these languages. However, it is the most typical declarative pattern in EA, and also in Tamil for example (Keane 2007;
pers. comm.). The pattern with more than two accents often exhibits a strong downward trend, frequently called downstep\(^\text{10}\) in the literature. I will term this pattern the \textit{wave pattern}. \(^\text{11}\)

![Fig. 11: A wave pattern with three downdrifting accents, a typical intonational tune of Egyptian Arabic](image)

Looking at the pattern illustrated in Figure 11, we may venture an interpretation of the initially cited ocean metaphor as the melodic component of intonation that consists of ripples, waves, swells and tides. While the \textit{ripples} can be related to minor phonetic effects such as consonantal perturbations which need not concern us here as they have no relation to meaning, the \textit{waves} may be compared to accents riding on the global downdrifting line which can be compared to swells. I assume that, unlike ripples, waves as well as swells, i.e. accents as well as trendlines (global patterns), carry meaning. As Bolinger (1986: 140) notes, the actual distinction between the two ultimately may prove to be a hard task.

A close look at intonation contours, whether trendlines or accents or edge movements, quickly reveals the unspectacular fact that, if partitioned into smaller recurring units, they may be said to consist of rises, falls and - looking at the hat pattern for example - level elements. This stresses the fundamental observation - to cite once more Dwight Bolinger -

\begin{quote}
that [...] intonation is fundamentally the \textit{opposition of up and down}, with meanings clustering around the poles of the opposition in accord with metaphorical extension […], also that intonation shares with physical gesture the manifestation of this opposition – up and down are carried by the facial muscles, shoulders, etc. In spite of its myriad ramifications, the system is coherent, to the extent that most if not all manifestations can be ultimately traced to the primary metaphor […] (Bolinger 1986: 22, emphasis added)
\end{quote}

The above quote implies a couple of basic observations: The first one is related to the fact that intonation is a rather iconic means of expression which will hardly be doubted by anybody working in the field. But Bolinger stressed the fact that even as a means of linguistic expression, intonation remains first and foremost a natural, i.e. non-arbitrary, instrument. The second observation relates to the fact that means of expression are very limited, namely fundamentally "the opposition of up and down." Contrary to segmental phonemes, intonational categories can only be very small in number. Given the fact that these means of expression have to cover a whole lot of meanings and functions, such as emotional, attitudinal, pragmatic and grammatical, we have to expect a great deal of overlap between them which makes intonation a rather uneasy ground for grammaticalization.

---

\(^{10}\) I use the term \textit{downdrift} here to refer to declining peaks and \textit{downstep} to the frequently occurring final falling accent without a peak.

\(^{11}\) Sosa (1999: 140) calls the pattern \textit{rebotado} ‘bouncing’.
The model I will introduce in the following section on Egyptian Arabic intonation and whose main outline has been presented in El Zarka (2011) relies on these basic assumptions. It proposes three major categories to which a gross description of intonation can ultimately be reduced: leading, closing and linking configurations that are functionally conditioned. These configurations or contours are based on the sound-symbolic use of pitch in the sense of what Ohala (1983, 1984, 1994) called the frequency code. The frequency code suggests a biological basis for the interpretation of sound that is derived from the fact that high pitch is associated with smaller sized creatures and low pitch with larger sized ones, based on the size of the larynx and the resulting height of the voice pitch. The fact that large and strong creatures by virtue of their size are conceived of as threatening and small creatures as non-threatening, has given rise to a number of secondary meanings of high and low frequencies, both paralinguistic and linguistic. Thus high pitch is associated with friendliness, uncertainty, and incredulity, which can be conventionalized to express modal and informational (or discoursal) meanings like questioning, opening a topic and continuation. Low pitch, on the other hand, conveys authoritative notions like power, decisiveness and certainty, which can be fossilized in the linguistic expression of assertion and finality (cf. Gussenhoven 2002, 2004: 80–84).

In the spirit of Ohala’s proposal, Carlos Gussenhoven (2002) suggested three biological codes that are held responsible for what is universal about intonation: Ohala's frequency code, the effort code, and the production code. The frequency code has already been explained above. The second code proposed by Gussenhoven, the effort code, is related to the amount of effort a speaker expends on the production of sound. More effort will result in larger excursions or more precise articulation (Gussenhoven 2002: 47, 50; 2004: 85-89). To get a message across, a speaker will raise his effort level. The manifestation of stronger effort is more salience or prominence. Such prominence can either result in higher overall pitch range or in wider pitch excursion on individual items. Gussenhoven only deals with tone, i.e. intonation proper, but the notion can be extended to include intensity and duration increase as well. It is intuitively clear that the effort code is relevant to the articulation of focus.

The third inherent feature of the speech production mechanism which has already been referred to at the outset of this section may be captured by what Gussenhoven (2002: 51-52, 2004: 89-92) calls the production (phase) code. This code is tied to the fact that the exhalation phase provides a limited time interval in which sound can be produced through the vibration of the vocal folds. As already noted, at the beginning of that phase subglottal pressure will be higher than in the end, the most natural consequence will be a drop in intensity and a decline in fundamental frequency which leaves us with high pitch in the beginning of phrases and low values at the end. It has to be noted that there is a certain amount of overlap between the codes. The functioning of the production code can be related to the physiological explanation of rises and falls as tension and relaxation (cf. e.g. Bolinger 1986: 194), in the sense that rises overcome the natural tendency of declination while falls follow it. At the same time, a high ending can be related to the frequency code as suggested above and thus gives rise to the
interpretation of uncertainty or incredulity. This results in a variety of informational and grammatical uses (besides the affective meanings) such as incompleteness and questioning and makes rises apt for the expression of questions and topics.

<table>
<thead>
<tr>
<th>Physiological sources</th>
<th>Biological codes</th>
<th>Universal interpretations</th>
<th>Linguistic interpretations</th>
</tr>
</thead>
<tbody>
<tr>
<td>SIZE</td>
<td>Frequency Code:</td>
<td>Affective:</td>
<td>?</td>
</tr>
<tr>
<td></td>
<td>small ~ big</td>
<td>submissive ~ authoritative</td>
<td></td>
</tr>
<tr>
<td></td>
<td>high ~ low</td>
<td>vulnerable ~ protective</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>friendly ~ not friendly</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Informational:</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>certain ~ uncertain</td>
<td></td>
</tr>
<tr>
<td>ENERGY</td>
<td>Production Code:</td>
<td>Informational:</td>
<td>Question vs. Statement</td>
</tr>
<tr>
<td>(phasing)</td>
<td>beginning ~ end</td>
<td>beginning: new</td>
<td>(e.g. H% /L%)</td>
</tr>
<tr>
<td></td>
<td>high ~ low</td>
<td>topic ~ continued topic</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>At end: continuation ~ finality</td>
<td></td>
</tr>
<tr>
<td>ENERGY</td>
<td>Effort Code:</td>
<td>Affective:</td>
<td>Continuation rise vs.</td>
</tr>
<tr>
<td>(level)</td>
<td>less effort ~</td>
<td>less surprised ~ more</td>
<td>final low (e.g.</td>
</tr>
<tr>
<td></td>
<td>more effort</td>
<td>surprised ~ more</td>
<td>H%/L%)</td>
</tr>
<tr>
<td></td>
<td>smaller excursion</td>
<td>helpful ~ more helpful</td>
<td></td>
</tr>
<tr>
<td></td>
<td>~ greater</td>
<td><strong>Informational:</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td>excursion</td>
<td>less urgent ~ more urgent</td>
<td></td>
</tr>
</tbody>
</table>

Fig. 12: after Gussenhoven (2004: 95): the three biological codes, their sources and their interpretations

As can be seen from the right column in Figure 12, there is some overlap in the phonological coding based on the production code and the effort code. Both codes may lead to falling vs. non-falling phrases, which in an AM model is represented by boundary tones. Arguably, the semantics of the linguistic interpretations are also related. Thus certainty and finality definitely have something in common. If something is certain, there is no further debate about it. Likewise if something is questioned, it calls for an answer, i.e. for continuation. However, as will be argued below in section 2.2.4.1, there seems to be some evidence for further differentiating between a mere unconscious lowering of the F0 versus its intended suspension, attributable to the production code, and the intended rapid lowering or raising of pitch, which is related to the frequency code.

2.1.5.2 The representation of intonation proper according to AM theory

In this section I will first present the phonological model of a Standard AM approach, as it is the basis for the present model that will be introduced in Section 2.2. As noted above, the present approach also relies on the fundamental assumption of tune-text association, an assumption that justifies it as an
autosegmental-metrical approach. Nevertheless, there are some major conceptual differences between AM Theory and the approach taken in the present study. These have been already alluded to in the preceding subsections and will be pointed out more clearly in the following discussion.

In AM Theory, the tonal layer contains a sequence of pitch accents (tones associated with stressed syllables) that consist of high (H) or low (L) tones. Pitch accents can be singleton tones H* or L* (the asterisk pointing to the association of that tone with the stressed syllable) or bi-tonal ones, either with a leading tone (i.e. a tone that precedes a starred tone) T T* or a trailing tone (a tone that follows a starred tone) T*T. Which of the tones is the starred one has been a matter of some debate; so has been the question whether a rise-fall in the pitch contour should be analysed as a rise LH* or as a fall H*L, as Standard AM Theory does not allow for accents that consist of more than two tones. I will return to these problems later on when discussing the intonation system of Egyptian Arabic in some detail. In a sequence of at least two accents that exhibit the downtrend illustrated in the waveform shown in Figure 1, the second accent is supposed to be lowered with respect to the first one. The phenomenon of descending accents is usually called downstep and the contrary phenomenon of accents constituting an upward trendline is called upstep.

In addition, the system contains edge tones, which demarcate the ends or beginnings of prosodic constituents. The restriction to only two tones was designed to overcome the empirical problems the level system of the American school of intonation (Pike 1945; Trager & Smith 1951), which divided pitch range (the range between a speaker’s lowest and highest frequencies) into four discrete levels, had to face. The level-approach was criticized by Bolinger (1951) as being descriptively inadequate and theoretically flawed. Bolinger noted that the four levels of the American models are too many and not enough at the same time to account for the actually occurring patterns and proposed that intonation should rather be described by configurations - a discussion that has later on become known as the levels-versus-configurations debate (between adherents of the American structuralist level approach and the ones who view that the basic elements of intonational tunes are contours such as rises and falls, as the scholars of the British School of Intonation or Bolinger) (Ladd 1996: 62). According to Ladd, the autosegmental-metrical theory can be viewed as having overcome the levels-versus-configurations issue by reducing the levels to only two and thus integrating Bolinger’s notion of pitch accent (cf. Ladd 2008: 64). The tonal component of the AM model was developed by Janet Pierrehumbert in her seminal thesis *The phonology and phonetics of English intonation* (1980). She adopted the autosegmental approach developed by Leben (1973) and Goldsmith (1976) and built her model on Liberman's *metrical theory* (Liberman 1979; Liberman and Prince 1977) and Bruce's analysis of Swedish word accents for the representation of tones (Bruce 1977). Building on these resources, Pierrehumbert was able to represent Bolinger's pitch accents in a two-tone model that translates the tone levels of the American approach into a sequence of highs and lows that are scaled relative to each other. At the same time, these tones are related to the metrical grid, accounting for the
fact that pitch accents are related to metrically strong positions. After Pierrehumbert (1980) had laid the foundation for the descriptive framework, a number of modifications and diversifications have been suggested and AM theory today is not a uniform theoretical concept. What is common to all AM models is that they are all conceived as phonological grammars consisting of single tones that are viewed as phonemes. This is in contrast to Bolinger's view who explicitly argues against the phonemic status of intonational categories (Bolinger 1986: 222). The smallest relevant unit in Bolinger's system are profiles (the former pitch accents) which he conceives as intonational morphemes. The present approach follows Bolinger in this assumption.

One important aspect of the AM approach is the assumption that pitch accents are assigned to the textual elements mediated by metrical structure, which defines strength relations between the different textual constituents (Ladd 1980, 1996; 2008). In line with the tenets of Generative Theory, the metrical structure is based on syntactic structure. This assumption is another aspect in which the present model differs from Standard AM theory as has been thoroughly discussed in Section 2.1.1 and 2.1.2.

Another principle of Standard AM theory is the strict linearity of intonational tones. This means that the existence of more global phrase level melodies (trendlines) as phonologically relevant entities is rejected. Such an assumption is an essential part of what Ladd (1983, 1996: 24) called contour interaction or overlay models (e.g. Fujisaki 1983; Gårding 1983,1998; Thorsen 1980, 1983; Grønnum 1998). These models all incorporate global trend lines into their representations. Instead, Standard AM theory views global tunes as a strictly linear string of local categorically distinct tonal events, the contour between which is unspecified and phonetically realized as transitions from one phonological event to the next without any look-ahead strategy.

With all that difference in fundamental issues between the approach taken here and the basic tenets of the AM theory, the reader will most probably and justifiably ask whether it makes sense to undertake the description of Egyptian Arabic intonation within a kind of AM model at all, which is actually intended. The reason for this decision is mainly twofold. Firstly, the belief that much of intonation is iconic and therefore universal is not tantamount to questioning the usefulness of detailed description and does not obviate the necessity of accurate instrumental investigation. To do the latter it is necessary to have a methodological and descriptive framework, which in the past thirty years has been carefully developed (not only, but most prominently) by the proponents of AM Theory. The second reason is the comparability of the descriptions. The identification of specific targets and their phonetic

12 The individual models differ regarding the existence of leading tones, e.g. the Dutch AM model (ToDI) (Gussenhoven (2005)) and the IVIE system (Grabe et al. (1998)) do not allow for leading tones. The same applies to phrase tones, which were proposed by Pierrehumbert (1980) to account for the contour at the end of phrases between the last pitch accent and the right edge of an intonation phrase that is characterized by a boundary tone.
properties has been the main issue of an extensive body of work on individual languages, many of which had not been described so far. If we take the results of the phonetically based studies as they are we are able to compare across many languages, regardless of the phonological interpretation and therefore also regardless of the theoretical premises endorsed.

Moreover, the fundamental opposition of up and down of intonation, acknowledged by Bolinger (1986: 221), is best captured by the use of only two symbols, H (high) and L (low). Even in mathematics a straight line or a line segment is represented by two specific points, as their position can unambiguously be defined within the coordinate system. As Ladd (Ladd 2006: 129) points out, drawing an analogy between diphthong gestures and pitch movements, “there is no contradiction between treating the pitch movement as a phonetic gesture and describing it phonologically in terms of its endpoints (e.g. L+H)”, thereby reconciling the work of other traditions with work done in the AM tradition. We might slightly change the individual components of that concession and state that there is no difference in treating the pitch movement as a meaningful unit and describing it phonetically in terms of its endpoints. Although I hold the belief that gestures are the relevant units in intonation, I think we have to acknowledge that a considerable body of work done in the past ten years based on initial findings by Arvaniti, Ladd & Mennen (1998, 2000) (which has become known as ‘segmental anchoring’) has shown that – under experimental ‘mechanistic’ conditions - the targets align with the text in predictable ways. Moreover, it will be shown for EA spontaneous speech below, that the alignment of the target tones is not totally random. Perception experiments have also suggested that contours in speech are mostly perceived in terms of a sequence of discrete pitches (House 1990, Mertens & d’Alessandro 1995). All this suggests that targets themselves are not unimportant. The fact that register tone languages exist that operate with specified pitch levels instead of pitch contours is further evidence for the importance of targets. Finally, even though the exploitation of pitch range is commonly dismissed as "emphatic" and therefore "paralinguistic" and therefore not belonging to linguistics, it is not yet clear whether this is actually the case.\textsuperscript{13}

Thus, I think it justifiable to consider the approach offered here as autosegmental-metrical. In the remainder of this chapter I will add some discussions of a number of specific issues that have been raised above to give some details of the descriptive framework and complete the theoretical picture.

\textsuperscript{13} Cf. the debate between Ladd and Hayes in \textit{Laboratory Phonology II} (1994).
2.2 The Intonation of Egyptian Arabic

I will start the description of Egyptian Arabic intonation with a synopsis of the prosodic model used in this study. This implies some redundancy in the text, as the fundamental ideas have partly been discussed in Section 2.1. The main discussion of the theoretical issues implicit in the model will, however, be presented in the subsequent sections. The reader who is only interested in the descriptive framework used in the empirical investigation is referred to this first section (2.2.1) that offers a concise presentation of the model without going into the theoretical issues.

2.2.1 A brief overview of AM approaches to Egyptian Arabic intonation

Egyptian Arabic has repeatedly been claimed to possess an utterly simple intonational structure (Rastegar-El Zarka 1997; Rifaat 2005; Hellmuth 2006b). The main characteristic of Egyptian Arabic intonation is the richness of pitch accents as already mentioned. Mitchell’s sketchy, but insightful remarks also point to that fact, when he mentions the “up-and-down” or “see-saw effect” in Arabic intonation (Mitchell 1993: 222). This tendency to accent all content words has also been recognized by Rifaat (1991) in his investigation of the neutral declarative sentence in Egyptian Classical Arabic and by Rastegar-El Zarka (1997) in her study of Egyptian Modern Standard Arabic (MSA). One linguistic study of the Egyptian Arabic vernacular that is devoted to the question of pitch accent distribution is by Hellmuth (2006b) who found that “in EA, the overwhelming majority of content words bear a pitch accent” (p. 69). In her corpus of EA, there were 2-4% of unaccented content words (p. 66). The successive tonal movements, i.e. the accents, mostly display a certain downdrift within the intonational domain (cf. Figure 3), i.e. each accent peak is (often considerably) lowered relative to the preceding one.

Another striking feature of EA IPs is that they can be lowered with respect to each other. Frequently, whole units of intonation are realized within a compressed pitch range. In EA intonation phrases, the peak of the last accent can be either partially or totally downstepped (2.2.4.3). These two terms will be explained below in the section on accent modification. The types of downstep have been attributed to the phonetic tendency of final lowering. Final downstep is probably more common in broad FOCUS, early FOCUS conditions and thetic utterances as opposed to narrow FOCUS of an argument in final position. Figure 13a shows a typical contour of a neutral declarative with downdrift throughout the whole contour and total downstep of the last accent, while Figure 13b depicts an utterance with narrow FOCUS on the locative phrase coded by a full-fledged rise-fall accent on this last constituent.

In recent studies, the ups and downs exhibited by the above pitch track have commonly been analysed as ‘pitch accents’ within an autosegmental-metrical (AM) framework. The most common (in some
descriptions the only) pitch accent type of EA has been analysed in three different ways: as LH for prenuclear (prefinal) accents vs. HL for nuclear accents (Rifaat 1991), as H*L (Rastegar-El Zarka 1997), as predominantly H for prenuclear accents and HL for nuclear accents by Rifaat (2005) and finally as LH* by Hellmuth (Hellmuth 2006b) who regards the falling shape of the final accent as being conditioned by the repulsion of the upcoming boundary with a low phrase accent and a low boundary tone, yielding the sequence LH* L-L%. The different analyses were at least partly due to differences in the theoretical framework applied, but it also raised the question whether the differences are also due to differences in the data themselves, especially as the three studies investigated three different varieties of Arabic in Egypt, namely Classical Arabic (Rifaat 1991), Modern Standard Arabic (MSA) (Rastegar-El Zarka 1997; Rifaat 2005) and Egyptian Colloquial (EA) (Hellmuth 2006b).

Fig. 13: typical downdrifting declarative pattern with last accent realized with downstep in panel (a) and a rise-fall in panel (b).

To find out whether this observed difference was due to register differences, Hellmuth and El Zarka (El Zarka & Hellmuth 2008) conducted a small-scale experiment with five speakers that investigated a parallel corpus of MSA and EA sentences. The results of the experiment suggest that the differences are speaker-dependent and probably due to speech style, but not to the different registers involved. It seems that an early alignment of the high peak is correlated with a more deliberate and careful pronunciation of either MSA or EA materials. Interestingly, another fact could be observed in this experimental data, namely the occurrence of another low target between a peak and the onset of the next accented syllable, thus creating a flat low stretch between two successive accents as illustrated in Figure 18 below. This can be explained as being conditioned by the long stretch of potentially unaccented syllables between the two accents, as the inter-accentual interval was designed to cover four to six syllables in order to avoid tonal repulsion from an upcoming tonal event.

---

**13**Rastegar-El Zarka’s (1997) study was carried out within a Gussenhoven-style framework (today’s ToDI annotation, cf. Gussenhoven (2005) as opposed to the ToBI annotation) while Rifaat’s (1991) and Hellmuth’s (2006) approaches are closer to Standard AM theory (ToBI annotation, (1997)).
There are various reasons to refute the analyses suggested so far. Firstly, the occurrence of two successive L-tones has to be accounted for, and an analysis involving a bitonal accent, whether it be a fall HL or a rise LH, does not provide descriptive adequacy. If the late alignment is to be taken as the basis for identifying the accent as a rise LH, it would be necessary to invoke the existence of a low boundary tone. Such a decision is not itself unreasonable and will be discussed later on. The second more general reason not to analyse the tonal contours in question as either a rise or a fall when the tonal contour is perceptually ambiguous is the arbitrariness of this decision. As we have seen above, the data are not uniform. In section 2.2.5, it will be demonstrated that there are contours that are clearly rising and others that are clearly falling, but there is also something in between, which may best be characterized as a neutral rise-fall. We will return to that issue below in the discussion of accent shapes. The remainder of this chapter will be devoted to the description of the model of intonation that will be used throughout the book and the fundamentals of which have been outlined in El Zarka (2011). This model rests on the basic assumptions about prosody outlined above.

2.2.2 A functional model of EA intonation

The present prosodic model is considered an autosegmental-metrical (AM) approach to intonation as it relies on the basic assumption of tune-text association (Liberman 1979). The fundamental idea is that prosody constitutes an independent linguistic level that interacts with other linguistic levels to produce actual utterances.

In the preceding sections the fundamental issues relevant for a formal description of intonation, such as stress, accent, rhythm, and melody have been discussed. One essential issue that has not been touched upon so far is function or meaning. AM theory considers itself a strictly phonological theory. Especially in the original version developed by Pierrehumbert (1980), the author deliberately makes no reference to meaning, on the assumption that the basic elements of intonation are phonemes, not meaningful units. It is precisely this view where the present account fundamentally differs from the standard AM approach. The model presented here is based on earlier accounts of intonation, among them most prominently the work of Dwight Bolinger (1958, 1972b, 1972c, 1986, 1989). I follow Bolinger (1986, et passim) in the assumption that the tonal categories of intonation are more comparable to ‘morphemes’ carrying certain meanings than to ‘phonemes’. This meaning is attributed to the functioning of the biological codes introduced in Section 2.1.5.1. Although intonation may be used to express grammatical relations and thus intonational constructions may have conventionalized grammatical uses, there is no principled “distinction between the grammatical and the ideophonic except as they represent extremes of a scale” (Bolinger 1986: 32). The base of intonation is still expressive and its basic function is the expression of emotion and attitude.
In the present model, intonation is regarded as a combination of two equal and a priori independent components, the metrical layer representing rhythm, i.e. prominences in time, and the tonal layer representing melody. The neat separation of the two components is, however, difficult, as both components interact, just as they do in music. Intonational tunes are composite structures built of (pitch) accents that may have acquired conventional meanings. An accent is not signalled by pitch alone but involves a bundle of prosodic features, such as pitch, duration, intensity and tempo.

The notational system used here is also adopted from AM Theory. The metrical component is represented in a metrical grid (Prince 1983, inter alia) and the tonal component will be annotated using a two-tone model (Pierrehumbert 1980). The adoption of a two-tone approach, however, does not imply that the individual tones are regarded as abstract phonemes. It is common standard that part of the metrical information is included in the tonal description as we will see below. Thus it is possible to dispense with the rather awkward grid notation for the annotation, and even in the great part of the presented examples. A metrical grid is thus only used for special illustrative purposes, e.g. to make prominence relations explicit.

2.2.1.1 The metrical level

It is assumed that speech rhythm is a natural preference that does not depend on the text and does not necessarily apply in specific prosodic domains. It is further assumed that accents – whatever their phonetic correlates may be – are the instantiation of rhythm. Accents are the physical reality of metrically strong positions that occur at more or less regular intervals and crucially have a grouping function. In EA, accents are predominantly realized by pitch. As pitch variation is what constitutes melody, it follows that pitch accents are simultaneously the building blocks of the tonal and the metrical components. Thus, the notation of pitch accents also includes the identification of word-level prominences. Apart from the rhythmic factor, the metrical component indicates the segmentation of speech. Continuous speech is parsed into isolated units, so-called intonation phrases (IP). An IP may exhibit internal rhythmic grouping into prosodic phrases (PP). However, PPs are more typical for formal registers and text genres and play no important role in the description of spontaneous speech. Thus, the notation will be reduced to the level of IPs which are separated in the transcription by a slash (/).

The function of the metrical component is to parse the speech flow into manageable chunks in terms of articulation and pronunciation, to group together what belongs together semantically. The strength of prominences is relational. While high prominences single out individual informative elements, low prominences serve to put information to the background. Thus, accent positions may in principle be grammaticized for a specific function in certain languages.
The functioning of the tonal component is complementary to that of the rhythmic component. It is used to establish coherence between individual parts of the speech on different levels and it signals attitudinal and modal meanings. Moreover, some pairings with specific morpho-syntactic structures may be conventionalized so as to indicate certain functions, such as sentence mode or TOPIC and FOCUS as in Egyptian Arabic.

In this model, four tonal categories are identified on a perceptual basis. The identified contours are considered as having a leading, i.e. prospective, closing, i.e. assertive, and purely linking function, respectively (hence the choice of the terms).

- a leading contour (predominantly rising)
- a closing contour (predominantly falling)
- a linking contour (low level tone)
- a ‘high link’ or leading-linking contour (predominantly flat, but on a high pitch level)

These four categories - with some variation within them - are deemed enough to describe the tonal patterns that occur in EA. As the tonal component is seen to be in principle independent of the metrical one, phrasing is not regarded as prior to the tonal structure but as a parallel option.

As I have pointed out above, I will use two tones (H, L) for a more fine-grained annotation of the tonal patterns. The tones are not taken to represent absolute values, but are to be seen in relation to each other. That is, an L tone in a specific position is lower than an H tone in the same position. A basic falling contour is thus transcribed as HL and a rising one as LH. We have also seen that whole tunes consist of individual accents. What then is the relationship between an accent and a longer contour? Firstly, one accent may already constitute a contour, i.e. there are leading accents and closing accents (the individual accent shapes are illustrated in Figure 23). The linking contour does not involve any prominence and consequently does not contain accents. A contour may, however, cover a number of accents. We will also be referring to such contours as trendlines. A trendline may also be leading or closing, implying that the accents are stepping up or down along a tangent that can be visually imposed on the pitch track of an utterance (see Figures 29 and 30). Finally, the high link or the leading-linking contour involves prominences, i.e. accents; these accents are however less prominent than the ones associated with a rising/falling contour. The leading-linking contour may involve a rising part associated with a word stress and continue in a horizontal line at a non-low pitch. This line may contain several accents that are associated with individual content words.

It is assumed that the leading and linking accents have communicative functions. However, accents are also correlates of rhythm and have the function to single out individual lexical items. This function
may also be fulfilled by a ‘meaningless’ accent, i.e. by one that is neither clearly rising, nor clearly falling. Such an accent will be called *default accent*. The default accent consists of a rise-fall that is associated with a lexical word in a specific way, i.e. the first rising part is associated with the metrically strong syllable (word stress) and the peak is usually reached in the next syllable. The subsequent fall then covers the whole of the content word including grammatical affixes and clitics (i.e. the prosodic word plus function words) to reach the lowest point at the beginning of the following content word. The default accent is illustrated in Figure 15.

As Figure 15 shows, the default accent is represented using the symbols LHL indicating the fall-rise. To create a leading or a closing accent, the individual tones may change their default position on hypothetical vertical and horizontal axes. Displacement on the horizontal axis is called *alignment*, i.e. alignment with the segmental string (the text), displacement on the vertical axis is called *scaling*, i.e. scaling within the pitch range exploited by an individual speaker. This displacement is noted by the use of features. Alignment features follow the tone symbol and are the following:

- \( T \) (no feature) indicates that the tone is aligned according to default
- \( T^< \) indicates earlier alignment
- \( T^> \) indicates later alignment
- \( T^* \) indicates alignment within the metrically strong syllable
- \( T^- \) indicates tone spreading

Figure 20 gives a schematic illustration of differently aligned H tones. The following features are used to indicate scaling; the scaling features precede the tone symbol:

- \( T \) indicates that a tone is scaled according to default
- \(^T \) indicates that a tone is scaled higher than default
- \( \vee T \) indicates that a tone is scaled lower than default
- \( ^\circ H \) indicates that an H tone is scaled at the same height as the preceding H
- \( ^\circ L \) indicates that an L tone is scaled at the same height as the preceding H
- \( \nabla H \) indicates downstep, meaning that the H tone is at the same height as or lower than the preceding L tone (this accent occurs predominantly at the end of IPs)
- \( L\% \) indicates that the low tone has reached the bottom of a speaker’s range (at an IP boundary)
- \( H\% \) indicates a high boundary tone associated with the final syllable of a word (at an IP boundary)

What does it mean for a tone to be scaled according to default? The length and shape of utterances are physiologically constrained by the amount of air available for articulation. During the articulation process, this air stream becomes weaker resulting in lower frequency and lower intensity at the end of an IP. That means, that a certain declination or downdrift of consecutive accents is observable across an IP. As a consequence, downdrifting accents are not marked as such, while suspension of downdrift (\( ^\circ H \)) or upward trend (\( ^\vee H \)) are annotated, being the marked options. This is not a trivial issue, however, and downdrift and downstep, i.e. the further lowering of a peak resulting in a level tone or a falling gesture will be discussed in 2.2.4.1.3.

63
The above features collaborate to turn a default accent into a configuration that signals meaning. Thus, the alignment feature ‘>’, for instance, turns a default accent into an unmistakably leading one (LH>). Table 1 shows the possible realisations of leading, leading-linking and closing accents.

In Chapters 4-6, the tonal categories established in this chapter will be used in the descriptions and the two-tone system with accompanying features will be applied in the prosodic transcription of the examples.

<table>
<thead>
<tr>
<th>tonal category</th>
<th>transcription</th>
</tr>
</thead>
<tbody>
<tr>
<td>leading accents</td>
<td>LH&gt;/_____L/ /H/ %</td>
</tr>
<tr>
<td></td>
<td>L^H&gt;/_____L/ ^L/ ^H/ (H)%</td>
</tr>
<tr>
<td></td>
<td>L^H^L</td>
</tr>
<tr>
<td>leading-linking accents</td>
<td>LH^-</td>
</tr>
<tr>
<td></td>
<td>LH/ _____^L/ ^°H</td>
</tr>
<tr>
<td>closing accents</td>
<td>HL</td>
</tr>
<tr>
<td></td>
<td>LHL%</td>
</tr>
<tr>
<td></td>
<td>LHL&lt;</td>
</tr>
<tr>
<td></td>
<td>LH^L</td>
</tr>
<tr>
<td></td>
<td>!HL</td>
</tr>
<tr>
<td></td>
<td>H^L</td>
</tr>
</tbody>
</table>

Tab. 1: Leading, leading-linking and closing accents.

2.2.3 The default pitch accent

Coming back to the issue of accent vs. stress: As already noted, the present approach draws a fundamental distinction between potential accent position and accent without making the distinction between stress as prominence brought about by other phonetic features than pitch and accent as ‘pitch accent’. This implies the assumption that pitch itself is prominence-lending. Although EA is not a tone language or a pitch accent language, word accent can be regarded as phonetically realized first and foremost by pitch. In EA, the default pitch accent (see below) that signals no special meaning has the function to single out individual lexical items. If a focus accent (see Chapter 4.5.4) is used, the structure of this accent is not radically different, it is only a modified version of the basic default type. In tone and pitch accent languages the strategies employed are different. While in Stockholm Swedish, a peak after the lexical accent can be identified if the item is focused, which Bruce (1977, 2005) and Gussenhoven & Bruce 1999 identify as an additional focal high tone (or sentence accent, Bruce 1977), the Southern and central dialects of Swedish do not mark focus by an additional tone, but use a simultaneous gesture with a wider pitch range (Bruce 2005: 421) (Figure 14).
Two other examples that are illustrative here are Japanese and Chinese. In Japanese accented words, in contrast to unaccented ones, the final rise for questions is only produced after the accentual fall which has been analysed as H*L (Beckman & Pierrehumbert 1986; Gussenhoven 2004). Thus, the addition of the intonational tone does not affect the lexical tone in any way (Abe 1972: 338). By contrast, Chang (1972) describes different sandhi phenomena of intonational rises and falls with the four lexical tones of the Chengtu dialect of Szechuan.

In EA, as will be shown in Chapter 4, focus is mostly marked by relatively stronger excursion and occasionally or by an additional accent. Thus, the function of singling out individual words and focus are stacked. If we do not wish to leave focus marking by wider excursion to the phonetic component while analysing the additional accent as phonological as is explicitly proposed in Gussenhoven (2004), we have to admit that the exploitation of pitch range is phonological, too. Under this view, pitch accents in ‘intonation-only languages’ are not treated as fundamentally different from those in so-called pitch accent languages and not even from tone in tone languages. The phonetic properties of the accentual prosody and the melodic prosody (intonation proper) may be fused or may remain discernible. It is also important that strengthening a pitch accent and arriving at different degrees of accent strength is not a matter of pitch excursion alone, but may also be brought about by other different alignment of the tones and other prosodic features such as intensity and duration. In the following, I will argue that EA has a basic accent that has no special meaning, but is only a manifestation of word stress and an instantiation of rhythm, while meaningful accents are modifications of the latter. I call this basic accent type ‘default’ accent.

In accordance with the preference of maximal tonal contrast, following the general Gestalt-principle of ‘figure against ground’, the default accent in EA is a *rise-fall* (Figure 15). The alignment of the two
flanking L tones and the H tone serve the two functions of accent as identified above, the *demarcative* function and the *culminative* function. The culminative property serves to mark the item being accented as sufficiently important to be foregrounded or singled out and the demarcative property helps to identify the borders of an item and thus facilitates parsing the speech flow.

Fig. 15: the default pitch accent of EA.

If enough segmental material is available and under 'neutral'\textsuperscript{15} conditions, i.e. for example in an experimental reading task as is frequently carried out in experimental production studies, all lexical words carry such a default accent. The typical alignment of the first two tones has been extensively studied by Hellmuth (2006b) who found that the L is stably aligned with the beginning of the stressed syllable, while the position of the H is less reliable. The results of the present study show that the position of the second L is also highly variable, but given that there is enough segmental material available between the H and the beginning of the next accented word, it will often seek the beginning of the following word, thus serving the *demarcative* function. The position of that L seems to carry an ostensibly higher functional load than the position of the first L, participating in signalling focus position (Rastegar-El Zarka 1997, see Chapter 4) or at least influencing the degree of cohesion or separateness between two successive accents (see 2.2.6).

Thus, the Egyptian Arabic 'pitch accent' can also be called as an accentual phrase as in El Zarka (2011), similar to what has been suggested for pitch accent languages like Japanese (Beckman & Pierrehumbert 1986) or Korean (Jun 1996), French (Jun & Fougeron 1995, 2000) (note that the latter clearly is not a 'pitch-accent' language). It has to be pointed out, that other models, although they do not use the term accentual phrase, rely on the same default accent concept assumed here. Thus, our EA default pitch accent is largely equivalent to the *prosodic phrase* as described by Vaissière (1983), Hirst's *tonal unit* (Hirst 1998) or to the *p-phrase* suggested for German by Truckenbrodt (1999). But nothing hinges on the terminological choice. Whether this configuration is called accent, accentual

\textsuperscript{15} I assume that 'neutral' intonation is an intonation that is only used in mechanistic reading tasks or in citations as argued by Schmerling (1976: 52) who notes that "the more a citation is divorced from context, the more it is divorced from semantics - the utterer of the citation is not producing a real, natural-language sentence with both a phonetic and a semantic representation; he is producing a string of phones. Putting an example sentence in a context forces an informant to consider what the sentence means."
phrase, tonal group, minor phrase, profile, or anything the like, the basic facts are the same. There is some default configuration that largely follows the *Gestalt* principle of ‘figure-against-ground’ whose exact shape and alignment is a language specific issue.

As already noted, in many languages pitch is the most important cue to prominence. This is certainly true for ‘pitch accent’ languages such as Japanese or Swedish, but also for many intonation languages, Egyptian Arabic being a prototypical example of this type. Consider that the rise quite stably starts at the beginning of the syllable of the word that is supposed to carry the word stress. As Modern Standard Arabic (MSA) is a variety that is never spoken as a native language, the stress patterns or stress rules of the native varieties are imposed on it by native speakers. This is also done by speakers of Egyptian Arabic, in which stress position is almost completely predictable from the syllabic structure of the word (Mitchell 1993; Hayes 1995; El Zarka 1999; Watson 2002). But as the word patterns that arise in MSA are sometimes highly artificial, the position of the main word stress may vary. A deviant accent position can be traced in the pitch track by a sharp rise at the beginning of the syllable perceived as prominent. Figure 16 shows pitch tracks of the word *balah-at-an* ”date-S.F-ACC.INDF” with two different stress positions as *BA.la.ha.tan* following the stress pattern of the pausal form (i.e. is the form without case ending) or *ba.la.hA.tan* with penultimate stress, following a preference for trochaic rhythm.

![Fig. 16: Two different accentuations of the syntagma *ti:na.TAN wa-balah.tan* "a fig and a date" as *ti:na.TAN wa-balah.tan* in panel (a) and *TI:na.tan wa-BA.la.ha.tan* in panel (b).](image)

In the following subsection we will see that this primary rise is usually present in different accent types. That is, after this rise, which covers the stressed syllable, the pitch may take different directions: it may *rise*, *stay level* or *fall*. I will argue that this second part of the contour signals the intonational function and thus is the intonationally important part of the accent, while the rise that covers the stressed syllable mainly lends prominence to that syllable and thereby the whole accented item.
Since Arvaniti, Ladd and Mennen’s (1998) influential study on Greek prenuclear accents, the assumption has become common in Standard AM theory that a pitch accent is defined with respect to the alignment of the tones with the segmental string and the identification of the starred tone as the most stably aligned tone. But the issue is far from being settled, not even within Standard AM theory. The different alignment studies conducted by Arvaniti, Ladd and Mennen resulted in different analyses of the Greek prenuclear accent as L*+H (Arvaniti & Ladd 1995) and the suggestion of four other possible analyses (Arvaniti et al. 2000), namely L+H*, L*H*, LH and (LH)*, which would be compatible with the experimentally established phonetic findings. The authors conclude that phonetic alignment cannot be the basis of identifying the starred tone and that the notion of "phonological association" of tones has still to be refined and its phonetic exponents defined with more rigor. The problematic issue of the phonological representation of pitch accents is reflected in the fact that a rise across a metrically strong syllable has been analysed in all possible different ways in the literature. Let us take the different analyses of the Spanish rising accents as an example. While some analyses did not differentiate between different types of rises and labelled them all as H* (Prieto et al. 1995; Pilar et al. 1996; Nibert 2000), others posited two distinctive phonological units L+H* and L*+H, depending on the alignment of the H tone (Sosa 1999; Face 2001; Beckman et al. 2002). But this leaves us with an ambiguity in assigning the star to a tone, being based on alignment facts in the case of the H tone, but not in the case of the L tone. Consequently, there have also been analyses labelling the rise across a metrically strong syllable as (L+H)* (Face 2001).

These examples may suffice to show the severe problems Standard labelling conventions encounter, even within one language and even more so across different languages, not to mention the problems of identifying what is phonological and what is allophonic in a certain language or variety. The proliferation of fine-grained alignment-based phonological distinctions on the one hand and the inconsistency of the conventions on the other hand, ultimately lead to the situation that what is, for example, commonly represented as L*+H in German might in fact be no different than what is usually transcribed as H* and L+H* in English, as noted by Atterer & Ladd (2004: 192). These authors therefore argue that "[i]t is difficult to escape the conclusion that we are dealing with a phonetic continuum of alignment" (ibid.) comparable to that of the voice onset time of consonants. As cross-linguistic comparison of the phonetic facts shows a continuum of values, they conclude that the observed alignment distinctions between languages and dialectal varieties should rather be viewed as a distinction in the phonetic realization of one category than as different phonological categories. Considering these facts, it seems doubtful that any phonological units on a par with segmental phonemes can be parsed out of the pitch continuum at all. In the approach taken here, a phonological representation in the strict sense of the word is therefore dispensed with in favor of a systematic phonetic transcription that is supposed to be the realization of the broad functional categories identified below in section 2.2.4.1.
Another problematic issue is that of rising vs. falling pitch accents, as what is commonly represented as a rise (L+H*) in a ToBI framework (e.g. Beckman & Elam 1997) partially corresponds to a fall in a ToDI framework (Gussenhoven 2005). A comparison of the alignment facts in different languages and dialects suggests that the rise covering a the metrically strong syllable, which is the basis of the rising pitch accent discussed above, is astonishingly parallel in all the languages investigated (Figure 17). The beginning of the rise mostly aligns roughly at the beginning of the accented syllable and the end roughly at the end of the accented syllable or shortly after, with language- and dialect specific variation. This suggests that this rise in all cases can be considered a 'lexical' prominence-lending feature. If this is correct, the Standard AM analysis transcribes as intonational pitch accent what is an important cue to the position of the accented syllable. Contrary to that, the already mentioned ToDI framework (Gussenhoven 2005) or the IViE system (Grabe et al. 1998) pay attention to the accentual foot and rather transcribe the peak and what follows it. They thus keep in line with the British tradition and with Bolinger’s profiles. In Standard AM approaches such an immediate fall is represented by a phrase tone delimiting an intermediate phrase (Beckman & Pierrehumbert 1986) whose sole function seems to be to account for the existence of an L tone or - in configurational terms - the existence of a fall. At least in EA, such low tones as in the example above do not automatically give rise to any kind of phrase break, not even a weak one.

In EA, this closing L tone typically seeks the beginning of the next morpho-lexical word as illustrated by the following example from a reading task taken from El Zarka and Hellmuth (2008); the initial syllables of the morphosyntactic words and the corresponding L tones are marked in boldface and italics. The stressed syllables are capitalized, and the alignment of the tones and the text are illustrated in the interlinear annotation. The pitch track can be seen in Figure (18).
Fig. 18: A typical intonation contour of a declarative sentence read in a neutral style with a large number of inter-accentual syllables. The relevant portions containing the low level stretches are indicated by an arrow.

The observations made above are not to be understood as a claim that the EA default accent could be divided into two parts, the first one serving a quasi-lexical function and the second one an intonational one. Of course, in intonation-only languages the type of pitch movement itself is not specified in the lexicon. As will be shown in the next section, accents in EA do not necessarily exhibit this rise, but there are also high level-tones and falling tones without preceding rises, and more rarely also low level-tones. In a rise or a rise-fall the initial part belongs to the accent as a whole, which serves two functions: a truly post-lexical one, namely the intonational function, and a ‘quasi-lexical’ one, in the sense that it helps identifying the lexical item by identifying the accented syllable. At the same time, it helps creating a consistent rhythm. However, the EA pitch accent does not have any paradigmatic function on the word level like in Swedish or Japanese, but only a syntagmatic one. I further assume that the realization of the rise lends the item it is associated with more prominence than a level H tone – the steeper this rise is, the more prominent is the accent. The recognition of pitch obtrusion or change as a major cue to prominence implies that larger pitch excursions create stronger prominences. Thus, the rise (LH) is in fact an intensified high level tone (H).

To summarize, the present model refrains from positing abstract phonemic pitch accents, but rather stays close to the acoustically measurable facts. The actual turning points are annotated and their alignment is identified in relation to the established alignment facts in the default case. The analysis

**16** The parentheses indicate that there is no peak visible in the contour (cf. section 2.2.4.1.2 for details).
assumes that this default accent does not itself signal any meaning, whereas deviations from default fulfill linguistic functions. The default pitch accent in EA is a rise-fall as illustrated in Figure 15, consisting of a rising part that lends prominence to the stressed syllable and thereby to the whole of the lexical item and a falling part that covers the accentual foot whose modification is used to signify pragmatic meaning.

2.2.4 From tones to tunes: accent modification and accent sequencing

So far, I have only dealt with the smaller intonational units, which have been informally referred to as tones, configurations, or ‘pitch accents.’ A tune, being coextensive with an entire IP, by definition, consists of at least one such configuration, mostly more. Tunes are usually regarded as coherent, conventionalized contours that signal a certain meaning, such as the contradiction contour or the surprise-redundancy contour in English (Liberman & Sag 1974, Sag & Liberman 1975), for instance. Some languages, most notably languages that make use of pitch for lexical distinctions, seem to rely on trendlines to express intonational meanings rather than on accent shapes. This is reflected in the overlay models of intonation that specify certain trendlines or phrase components on which the individual (lexical) accents ride. Unsurprisingly, such models have been developed for tone or pitch-accent languages, such as Japanese (Fujisaki 1983) or Swedish and Chinese (Gårding 1983, 1998). Similar is also Grønnum’s model for Danish intonation (Thorsen 1980; Grønnum 1998). For the purposes of speech synthesis, overlay models and linear models are both in use. But for a linguistic analysis, the problems with the overlay models pointed out by Ladd (1996: 27ff.), namely to provide a “quantitative definition” of the phrase component and thus establish a category with some predictive power have to be taken seriously. On the other hand, if it turns out to be the case that slopes categorically vary with, for example, different sentence types, as suggested by Grønnum’s investigation of read speech (Thorsen 1980, Grønnum 2010)\(^\text{17}\), a strictly linear model will not be apt to capture the regularities, at least not without a considerable amount of preplanning. In any case, the Danish findings call for some sort of extrinsic representation of pitch range as suggested by Ladd (1996) that deals with pitch range as being orthogonal to pitch accent type (cf. the analysis of downstep below).

For Egyptian Arabic, we also have evidence that trendlines or phrase components do play a role (Norlin 1989; Rifaat 2005), but to my knowledge there is no evidence that the specification of

\(^{17}\) Note that the investigation of spontaneous speech in Danish (Grønnum & Tøndering (2007)), even though it also shows that slopes generally vary across sentence types, it does not show the same categories as those established for read speech. The results might also be compatible with two distinct categories – the declarative one falling to the bottom of a speaker’s range and the interrogative one falling less steeply or not falling at all. This latter case is reminiscent of Bolinger’s distinction between B profile and A profile.
categorically differing slopes is needed. The analysis of the data shows that the rising and falling trendlines that systematically occur may also be captured within the same three basic categories that are characteristic of the individual accents as already introduced above and elaborated in section (2.2.4.1), the leading the linking and the closing configurations. Descriptively, they can be handled by a linear notation using diacritics for upstep and downstep (2.2.4.2 and 2.2.5). In the following subsection I will present the different shapes one pitch accent may take, for the time being, without systematic reference to meaning. It has to be emphasised, though, that it is intonational meaning that is predominantly responsible for the modifications of the basic accent shape.

To begin with, there are phonologically conditioned modifications concerning the first and the second L-tone. Depending on the textual material, the first L-tone can be missing at the beginning of an intonation phrase as an instance of truncation when there is not enough segmental material available. However, if the L is realized in such cases, this will add to the salience of the accented item. It has been claimed that some languages tend to truncate the tonal information rather than squeeze it onto the sparse sonorous segments available, while others prefer to compress the tones (Ladd 1996). For example, while English has been shown to prefer the realization of a full fall at the end of an IP, German seems to truncate the fall (Grabe 1998b). At least at the end of a contour, which contains the important tonal information, Egyptian Arabic is of the compressing type, as falls are completely realized, even over very little sonorous material. The EA data investigated in Chapter 4.4.1.1 suggest that truncating and compressing are probably speaker-dependent.

2.2.4.1  Meaningful accent modification

In accordance with the initially identified three types of contours, the leading, the linking and the closing contour, I will show how one single accent will be shaped conforming to these three meaningful categories. Figure 19 shows a schematized representation of the accent shape over a disyllabic word with the syllable structure CVCCVC.

![Figure 19: Schematic representation of a pitch accent in Egyptian Arabic](image-url)
The first rising part within the shaded area represents the prominence-lending movement related to the accented syllable, and the second part varies due to intonational function. The dashed lines represent the various realizations of a leading tone, the dotted lines those of a linking tone, and the solid line represents the closing tone.

In El Zarka (2011) I only distinguished between two types of accent "leading" and "closing" on a par with Bolinger's B accent and A accent. The linking contour was reserved for deaccented or only slightly prominent parts of the tune, thus not representing an accent. I thus subsumed all level contours that contained prominent items and therefore accents under the heading of leading. A closer examination of the data, however, has suggested that we probably should distinguish between level accents and rising accents as they frequently relate to different meanings, types of continuation and questioning, respectively. This also ties in with Gussenhoven's distinction between the linguistic meaning attributed to the production code as continuation versus finality and the frequency code as question versus statement (2.1.4). Different categories of continuity have also been put forward by Delattre (1966), who distinguishes between "continuation mineure" and "continuation majeure", and von Essen (1964: 210ff.), who differentiates between the "interrogative Tonführung" (interrogative tone) and "progresdiente Tonführung" (continuous tone).

Rifaat (2005) made the observation that in EA accents may be tilted towards the left or the right. He suggests to describe pitch accent types in terms of "peak features". He assumes the existence of

a ‘default’ or an ‘unmarked’ accent [H], which has a projected peak aligned with the middle of the stress group, the shift on horizontal or vertical axes would account for all other accent types. Thus, a shift leftward on the horizontal axis results in an early peak or [HL] accent or rightward on the same axis yields a late peak or [LH] accent. (Rifaat 2005: 55)

This observation can be explained by the fact that the alignment of the H tone with the segmental material may differ. Rifaat's default accent is basically the same as the one suggested here (Fig. 15). In the 1980ies, Ladd (1983) and Gussenhoven (1984), having in principle adopted Pierrehumbert's (1980) two-tone approach to English intonation, developed a representation of different accent shapes using features that applied to the peaks (and less so to the valleys) related to specific meanings. As Ladd (1996: 98) points out, these features where intended as a kind of "cross-classifying device for accent types". Ladd, for example, used the feature [± downstep] for the representation of downdrifting or downstepping contours to account for the similarities among all these contours, whereas in Pierrehumbert's (1980) system downstep was treated as the phonetic consequence of the occurrence of a triggering L tone within a preceding accent.

---

18 Rastegar-El Zarka (1997) made use of some of the features proposed by Ladd in her description of the intonation of Modern Standard Arabic as spoken by Egyptians.
In El Zarka (2011) I proposed to make use of such features to account for the observation that in natural speech the tonal targets are not that stably aligned with the segmental string as suggested by the experiments. I suggested that the modifications brought about by these features turn the 'default' LHL accent into either a leading or an unambiguously closing one. The first type of features that are employed in the present model are alignment features, i.e. the horizontal shift of the peak or the valley in Rifaat’s terms. The second type of modification concerns the scaling of the individual tones. Although Rifaat does not elaborate on this second type of modification, his basic ideas seem to point to the same assumption, as suggested by the description in the above quotation.

It is important to note that the theoretical status of these features differs from those used by Ladd (1983) and Gussenhoven (1984). These authors regard the features as a phonological cross-classifying device to make distinctions among a family of accents, in this case falling accents. Ladd used the feature [± delayed peak], for example, to account for what has been known as scooped fall as an intensified variant of the plain fall (Ladd 1983: 731). Similarly, Kohler (1991) established a three-way distinction of meaningful peak alignment (peak synchronization) in German, identifying an early, a medial and a late peak with their suggested meanings of "established fact", "new fact" and "emphasis on a new fact" and "contrast" to "given fact" (Kohler 1991: 160).

Contrary to Kohler, I do not assume that peak alignment on its own signals a specific meaning. I rather view it as one of a bundle of features that are used in combination to derive meaningful configurations from the default accent type. In the present model, the features used are not ‘phonological’, i.e. I do not assume that one feature alone turns an accent into a different phonological category. I use the features as a description of the phonetic realization of rising and falling accents and trendlines. Although a scooped fall, for example, also occurs in my EA data, it is not possible to relate it to any specific meaning. Given the large gap between what is known about English and EA intonation today, it seems premature to establish a system accounting for these nuances of meanings in Egyptian Arabic.

A consistent annotation of spontaneous speech along these lines reveals that ‘prenuclear’ (i.e. prefinal) accents are not all alike, as suggested by Hellmuth (2006b), but that the choice of the accent is dependent on semantic and pragmatic factors. Ample evidence for this claim will be given in the empirical chapters about information structure. As I have already pointed out earlier (El Zarka 2011), concerning the intonation of EA, we are still in a stage of knowledge, where it is necessary to register all details of differences we can detect in the phonetic realization of intonation contours before a phonological representation - if this is desirable at all - can be established, keeping in mind that “[i]n the gradient world of intonation, everything that is detectable is potentially significant.” (Bolinger 1986: 225).
Taking the default alignment as identified by Hellmuth (2006) as a basis, it can be observed that the alignment of the peak is highly variable. If the H tone is realized in the vicinity of a syllable's end, i.e. either within the sonorant part of the stressed syllable or within the first sonorant segment of the following one, it will be transcribed as a simple H. A medial peak realized approximately in the middle of the vowel of the accented syllable is marked with an asterisk as $H^*$, if necessary. The peak may, however, be aligned way beyond the accented syllable, frequently close to or at the boundary of a constituent (mostly a lexical word), thus producing a constituent with an overall rising contour. This fact will be transcribed by a diacritic ‘>’ as $H^>$. Adversely, the H may also be aligned fairly early in the accented syllable, or - more often - before the accented syllable. Instances of such an early peak are notated with the diacritic ‘<’ as $H^<$.

Sometimes we find that the peak is quite regularly aligned, roughly with the same segmental position as in the default accent, but then the F0 shows no declination for a certain stretch of speech, sometimes up to the following peak, which in this case is at the same height as the preceding one. In such cases, the diacritic ‘−’, usually used to indicate tone spreading, is used: $H^−$.

Medial and early peaks usually coincide with (rising-)falling contours, enhancing the falling part of the contour. Figure 21 shows different alignments of the H in an initial accent on a topic constituent in panel (21a) and in an accent on a narrowly FOCUSED constituent in panel (21b). As can be seen, the FOCAL constituent shows a bundle of phonetic traits besides the different alignments of the H tone. Its overall duration is longer and the accented syllable is characterized by higher intensity. The signal

---

19 Note that in Standard AM theory the asterisk is used to mark the starred tone of a pitch accent. As the discussion in section 2.2.3 shows, the theoretical status of the starred tone is rather controversial. In a model that assumes the primacy of contours, the assumption of a phonological category ‘starred tone’ makes no sense. In the present model, the star is therefore used as an alignment feature.
also clearly shows that the early alignment is not due to the upcoming L tone. The L is reached within
at the end of the vowel /a/ in the following word ka:n "was" in both cases. Note that the H and the L
tones exhibit approximately the same pitch height in the TOPIC and FOCUS constituents. Assuming
that there are no boundaries and boundary tones at play, the difference that matters is really the whole
accent contour that covers the semantic constituent, namely rising or (rising)-falling. This also
suggests that the alignment is a voluntary choice to produce a contrast between rising and falling
accent and that it is not dependent on other phonological phenomena. Of course, such late alignments
frequently coincide with phrase breaks where an analysis that uses a phrase tone or a boundary tone
could be considered. But this is dismissed for two reasons, firstly the phrase break is no necessary
condition and secondly, there are optional boundary tones (see 2.2.6.2) that show a distinct phonetic
behaviour.

Medial peaks, which are aligned within the vowel of the accented syllable instead of the default
position towards the end, can most frequently be observed in closing accents. This may be due to the
larger effort expanded on the articulation of the accented syllable so that the rise is produced faster
than the normal speed. Early peaks, by contrast, i.e. peaks that are aligned at the onset of the accented
syllable or even before, are a means of enhancing the falling quality of the closing accent and are
related to the phenomenon of final downstep. They may also be regarded as an option controlled by
the speaker to enhance to nuance of finality and completeness.

Alignment features may also be observed in low tones. As already noted, the L tone at the beginning
of the accented syllable has rather stable alignment properties, serving as the onset of the accentual
'lexical' rise. The second L at the end of the fall (called L2 in the following) is more variable and
carries more functional load, as will be shown in Chapter 4.4.1.1. Suffice here to say that the
alignment of the L2 may be late and thus coincide with the onset of the next accented syllable. Note
that I have claimed above that in a default accent the L2 is aligned with the beginning of the next lexical word. This word may begin with a metrically strong syllable. In this case L2 and L1 coincide. The two L tones may, however, also merge, if the second word does not start with an accent. In such cases I analyse the accents as ‘linked’ – in this case, the first accent will be represented as a rise LH.

Contrary to that, L2 may be aligned much earlier, sometimes even at the end of the accented syllable itself, thereby producing a very abrupt fall. In this case L2 is transcribed as L*. Such a configuration is typical for narrow FOCUS. Thus the pragmatic impact of early L2-alignment is rather high.

The last alignment phenomenon that will be presented here is the alignment of L1 within the accented syllable. In parallel to the use of the asterisk with a peak, an L1 aligned within the accented vowel will be transcribed as L*. In EA, L* tones are more frequent as a variant of a fall than as the starting point of a rise. This means that rises of the type L*H, which are common in many languages, such as Danish (Thorsen 1980; Grønnum 1998), German (Wunderlich 1988; Uhmann 1991; Féry 1993; Grice & Baumann 2002), in some varieties of Spanish (Beckman et al. 2002) and Neapolitan and Palermo Italian (Grice et al. 2005), are rather rare in EA. In this language the L* mostly occurs at the end of a contour (Rastegar-El Zarka 1997: 251-253) as a manifestation of total downstep (Fig. 13a), which will be dealt with in the next section, and in a certain ‘implicational’ contour, known in the literature on English intonation as the surprise-redundancy contour (Sag & Liberman 1975). This tune is realized in Egyptian Arabic by the last two accents of the contour, the first of which falls low, while the second one rises quite high with a subsequent slump or fall. An example is given in Figure 22.20 This contour is extremely frequent in the corpus, especially as it is a regular feature of elicited sentences in the experiments for some speakers. The speakers seem to be bored by the task and express there uneasiness with an intonation that says ‘I’ve told you before’ or ‘You should know by now!’ or ‘Isn’t that obvious?’. In these cases, the alignment of the low tone mostly coincides with a stressed syllable, but it by no means has to, which constitutes strong evidence for the primacy of the contour. The crucial feature of the tune is the inverted (Bolinger 1958, 1986) pre-final accent with a final one whose peak is significantly raised, whose final L being crucially non-low. The typical redundancy tune in EA may thus be represented as L* ^H^L.21.

20 Bolinger (1986: 298ff), in a brilliant deconstruction of Sag & Liberman’s "surprise redundancy" contour analyzes that pattern within his overall tension-relaxation scheme as a combination of the profiles C+A, with the C for “restraint” and an A on which the speaker “explodes” (p. 301-302). Similarly, the EA contour can be understood as a mixture of uncertainty and questioning on the first part, followed by some kind of assertiveness, indicated by the fall from high. This fall is mostly suspended or truncated and thus again loses some of its force, a force that would be fully borne out, if it were falling from high to very low.

21 The symbol ^H is used to transcribe a high tone that is higher than the preceding one, while ^L denotes a truncated fall with a final low that does not go down to the expected low value. The use of these diacritics will be explained in the next section.
Fig. 22: Redundancy contour: /fuf-na:gi f-malʕab il-manjal (see-1PL Nagy in-playground DEF-Manyal) ‘we have seen Nagy on the playground of Manyal’.

The second use of the star notation is for cases where there are no peaks or valleys, but where there nevertheless is a clearly audible prominence on a lexical item despite the monotonous contour. Of course, this transcription is theoretically unsound to a certain extent. If we wished to strictly abide by the theory we should not mix up levels, and such prominences should only be represented in the metrical component. However in practice, for the sake of simplicity, the metrical component is not normally represented and the information about where the prominent items are may thus be included into the intonational representation. Usually, it is not necessary to indicate prominence because accents are always prominent. However, in some cases there are additional tonal events that indicate an accent in a metrically weak position. In such cases, the star is also used to indicate that accent, as in double accentuation indicating focus (see 2.1.4 and 4.5.4). Thus, the star notation serves two ends. The first one is to represent the alignment of peaks and valleys identifying the turning points as being aligned within the stressed syllable and to indicate prominences wherever the accent is located on a metrically weak syllable. This last notation combines the melodic and the rhythmic properties of the contour and can be compared to a musical notation that indicates the necessity of assigning a note of the same pitch an own beat.

### 2.2.4.1.2 Scaling

Let us now turn to the second family of features, the scaling features. Scaling is pertinent to the vertical position of a tone within the pitch range exploited by certain speaker’s, i.e. its actual pitch height. Belonging to pitch range variation, the representation of scaling is an area of major dissent, as it pertains to the perennial problem of disentangling what is linguistic and what is paralinguistic in intonation. In AM models, scaling is not usually included in the phonological representations. Given
that pitch range variation is frequently related to emphasis, which is regarded as paralinguistic, Ladd (2008: 209) states that “AM theory – like every other approach to intonation [...] - has a problem with pitch range.”

Following Bolinger’s (1951) critique of the American structuralist model of intonation with its four phonemic levels and his subsequent configurational account (Bolinger 1958) with the option of gradient pitch range variation, early AM approaches have all regarded pitch range phenomena as orthogonal to and thus not included in the phonological representation (Ladd 1994). More recently, however, attempts have been made to account for pitch height phenomena. Most noteworthy is Ladd’s suggestion of a metrical representation of pitch range ((Ladd 1990, 1993, 2008), based on the assumption that relational pitch range phenomena such as downstep and upstep (i.e. falling or rising trendlines) have to be accounted for. In an earlier paper (Ladd 1983) had used two features for the description of pitch range variation: [+ downstep] and [+ raised peak]. While he later on replaced the downstep-feature by the metrical representation, he dismissed the representation of an overhigh tone that involves more prominence as belonging to the gradient and thus phonetic part of intonation.

The position advocated here follows Ladd in the basic assumption that there are linguistically meaningful pitch range effects. I assume that linguistically significant pitch range phenomena are of two different types. (1) The first type is related to downstep and upstep as a purely relational feature, affecting tonal levels in a purely syntagmatic way. I assume that such relative height effects parallel the distinction between accent types. (2) The second type, however, in addition or rather in combination with the syntagmatic relative height effects involves prominence relations among the accents and thus also involves vertical variations of tonal space. In the present model, the scaling features are used to derive both types of pitch range variation. That is, the present model explicitly includes certain pitch range manipulations that produce prominence effects in the tonal description. We will turn to that issue below in Sections 2.2.5.1 and 2.2.5.3.

In the following, I will present the scaling features and show examples how they are applied to the single tones of the accent. The features are used to describe the shapes of both rising and falling accents. Every single tone of the accents is transcribed not only according to its relational value as being high (H) or low (L), but information about the relative height difference to the preceding tone is provided by the scaling features. This procedure allows the representation of downstep, same height and upstep and at the same time of pitch range expansion or compression. What is not included in the annotation are overall register differences as well as clearly paralinguistic pitch range manipulation.

---

24 ‘Pitch level 4’ of the American structuralist model.
such as generally larger pitch excursions in vivid speech. The notation applied is thus strictly relational, even in cases where prominence-lending pitch range manipulation is indicated.

To see how this works, let us first turn to the representation of the meaningful accent types suggested above. Following Bolinger (1986: 165), I define leading and linking accents, both exemplars of the predominantly rising type, as not abruptly falling after the accented syllable. This may either be realized phonetically by late peak alignment as suggested above or by a plateau to the next peak, or else by a slow fall. Frequently this fall does not arrive at a very low target, but goes down only halfway before it rises again, thus constituting a 'sag' between two peaks. This type of fall is annotated using the diacritic "^" as ^L following an H tone. Such high-scaled L tones may serve as the starting point for an upstepped accent. If they are final in an IP, they may indicate non-finality or at least less assertion and finality than a low falling L. The fall of a closing accent, may on the other hand be enhanced by the lower scaling of L2 which will be represented as ^L.

In the next section we will turn to global shapes of tunes or trendlines and it will be shown that it is important whether an accent peak (less so a valley) is at the same height, higher or lower than the preceding one. For that purpose the same diacritic will be applied to H tones: "^H" for a drastically lower peak and ^H for a higher one. Note that a certain downdrift will be assumed to constitute the default, i.e. the lowness of the H will only be represented when it is markedly lower than usual. The same height of neighbouring H tones will be indicated by the diacritic "°" as °H.

When two L tones occur in juxtaposition (i.e. L2 followed by L1 of the following accent), it can be observed that they always stay at approximately the same pitch level (cf. Figure 18). The pitch of a closing L2 of an accent thus spreads to the L1 of the following accent, perhaps taking a mild declination factor into account. So the relationship between these two tones is not annotated. The following table sums up all different variants of the leading, linking and closing accents, employing the different features.

The downstep diacritic "!" introduced in Ladd (1983) will be used for a downstepped accent, which in EA is almost exclusively IP-finally. There are rare occurrences of terraced-shaped downdrifting contours, but mostly downstep only pertains to the final accent of an IP. Unlike Ladd’s downstep diacritic, "!" does not signify that a following accent is lower than a preceding one, which, as I have already noted above, is not annotated in the case of the usual downdrift. To indicate a significant lowering of a following accent, the scaling feature "˅" is used. The downstep diacritic indicates that an

25 Most ToBI-based accounts draw a phonological distinction between two types of pitch accents involving a local F0 minimum between two F0 maxima, one of which is considered as L+H* with a phonological low target and the other simply as H*. In the second case the F0 minimum is considered as a phonetically governed "sagging transition", following the original AM model (Pierrehumbert 1980). Ladd & Schepman (2003) offer experimental evidence against this assumption and suggest that both types of phonetically differing accents should be viewed as belonging to the same accent category.
H tone does not involve a peak, but either consists of a level tone spanning the accented syllable with followed by a fall (\(\text{HL}\)) or, in combination with an early peak, the fall spans the accented syllable (\(\text{H'L}\)) and continues to the end of the IP. A third type of downstep (total downstep) involves a low level tone with high prominence which is transcribed as \(\text{L}^*\) (see 2.2.4.1.3 for more details).

The notation system still includes another diacritic related to pitch range variation. The \(\%\) sign will be used to mark optional boundary tones. An \(\text{L}\%)\) is transcribed when the bottom of a speaker's range has been reached and an \(\text{H}\%)\) indicates a high rising boundary (see 2.2.6 for more detail).

<table>
<thead>
<tr>
<th>leading accents</th>
<th>linking accents</th>
<th>closing accents</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\text{LH}^\uparrow)</td>
<td>(\text{H}^\uparrow)</td>
<td>(\text{HL}^\uparrow)</td>
</tr>
<tr>
<td>(\text{LH}^\downarrow)</td>
<td>(\text{L}^\downarrow)</td>
<td>(\text{H}^\downarrow)</td>
</tr>
<tr>
<td>(\text{LH}^\text{(L)})</td>
<td>(\text{LH}^\downarrow)</td>
<td>(\uparrow\text{HL})</td>
</tr>
<tr>
<td>(\uparrow\text{HL})</td>
<td>(\text{(H)L}^\downarrow)</td>
<td></td>
</tr>
</tbody>
</table>

Fig. 23: Schematic representation of leading, linking and closing accents.

In Section 2.2.3, the default pitch accent was introduced as a tri-tonal accent. We have also seen above, that in AM theory there is a strong reluctance to accept the existence of tri-tonal accents. However, such an accent type was proposed by Grice (1995) for Palermo Italian. Grice (1995: 121) assumes a binary branching structure for the tritonal accent with a weak and a strong part. She argues that the rise and the fall are both essential elements of the nuclear rise-fall accent in Palermo Italian questions.\(^{26}\) She thus represents the nuclear accent of Palermo Italian as a tri-tonal accent with LHL (Grice 1995: 121).

\(^{26}\) It has to be noted that Grice later changed her analysis of Palermo Italian to be in line with the assumption of Standard AM Theory that does only allow for maximally bi-tonal accents (Grice et al. (2005)).
I make use of Grice’s suggestion to represent the *leading* vs. the *closing* accent in EA metrically, as shown in Figure 24. While in a closing accent the falling gesture is the strong element which carries the intonational meaning (Fig. 24a), in a leading accent it is the rise (Fig. 24b). Any subsequent fall will not be realized as an abrupt deliberate downturn, but only as a gradually declining line interpolating between the H and the next low target or spreading up to the following H target. Given, that the flanking L tones are in both cases the weak constituents, they may also be missing altogether, leaving us with a rising LH and a falling HL accent at the extremes of the continuum.

![Fig. 24: Representation of the closing pitch accent (PA) (panel a) and the leading and leading/linking pitch accents (panel b) in EA.](image)

### 2.2.4.1.3 Downdrift and downstep

As we have seen above, EA intonation phrases frequently exhibit an overall downward trend with a downdrifting trendline of the peaks and to a lesser extent of the valleys. This phenomenon is usually called *downstep* in the literature on intonation (Pierrehumbert 1980; Ladd 1996; Gussenhoven 2005).\(^{27}\) In the older Africanist tradition, however, the term *downdrift* was used for the assimilation of a tone to the preceding one, i.e. an H tone is lowered in relation to a preceding one after an intervening L tone. The term *downstep* was reserved for the spontaneous occurrence of a lower H tone (Pierrehumbert 1980: 147; Gussenhoven 2004: 101). I will keep with the older term *downdrift* for the characterization of the typical EA declarative contour.

We have also seen that there is a natural downward trend of a phrase due to physiological factors attributed to the *production code* by Gussenhoven (see 2.1.4.1). The tendency for the pitch to gradually decline has been attributed to a physiologically motivated factor called *declination*. Declination has also integrated into quantitative models of intonation such as that of Pierrehumbert (1980). However, downdrift and declination are in fact difficult to disentangle (cf. also Ladd 1996: 74).

---

\(^{27}\) Downdrift has therefore also been referred to as *automatic downstep* (e.g. in Gussenhoven 2004).
In the present model of EA intonation, the assumption is made that downdrift typically affects phrases as a whole, as the declarative contour in Figure 13 above shows. On a par with the default accent, downdrift is the default trendline of a sequence of accents, following the mechanisms of the production code. Thus, downdrift is not included in the tonal notation. However, a steeply falling trendline is also assumed to have an explicit function, namely the indication of assertion and finality. As already noted, it is not trivial to differentiate between the simple downturn attributable to the production code and a voluntary extra downward movement in a sequence of accents. Marking assertion by a closing contour entails that it falls to low, somewhere in the vicinity of the bottom of a speaker's range. However, the natural downdrift yields approximately the same result, because as soon as a speaker has used up his breath when speaking, there will be no energy left to produce high frequency values. Therefore, the notational convention here is not to mark downdrifting accent, but only to mark suspension of downdrift and upstep. However, the issue calls for more fine-grained phonetic studies.

Another issue is the downstep of the final accent (Figure 25). This realization of the final accent may in fact be regarded as the unmarked option in a declarative contour. This type of downstep is a variant of the closing accent, i.e. a voluntary decision. The difference of the a contour involving final downstep and a contour without final downstep is illustrated in Figure 13 above that shows two downdrifting contours only differing in the last accent. The utterance on the right has a final accent with a peak that is at least as significantly lowered as the peaks before, whereas the utterance on the left has a final accent with (total) downstep (see below). In Rastegar-El Zarka (1997), I therefore made a clear-cut distinction between downdrift as a natural phenomenon of whole utterances and meaningful downstep at the end of the phrase signifying finality and "matter-of-fact" (Rastegar-El Zarka 1997: 323, 332). Ladd (1996: 75f.) also illustrates downstep in contrast to non-downstep using a two-accent phrase with a downstepped final accent. He notes that the downstepped accent is a separate phonological choice adding “a nuance like finality or completeness” (Ladd 1996: 76) without diminishing the prominence of the accent. The “matter of fact”- meaning of downstep was also pointed out by Bolinger (1986: 48).

Given that the lowering of a peak in relation to the preceding one will be taken as default and therefore not indicated, the downstep sign (!) introduced by Ladd (1983) will be used for transcribing cases where there is no peak on the accented syllable. Figure 25 is a schematic illustration of the three types of final downstep that occur in EA intonation phrases. Figure 25a illustrates a downstepped accent with a level tone associated with the stressed syllable, downstep with an early peak is illustrated in (25b), in which case the fall starts before the stressed syllable. Figure 25c shows an accent with total
downstep\textsuperscript{28}, the only discernible peak being in the prefinal accent while the accented syllable itself is on a low monotone, which is annotated as L\textsuperscript{*}, as indicated above.

\begin{table}[h]
\centering
\begin{tabular}{ccc}
(25a) & !HL & (25b) & !H'L & (25c) & L* \\
\end{tabular}
\end{table}

Fig. 25: Three types of final downstep: partial downstep (panel a), downstep with early peak (panel b) and total downstep (panel c). The shaded area represents the accented syllable of the downstepped accent.

The three variants of downstep all belong to the family of closing contours and thus all signal finality, completeness and additionally convey a matter-of-fact nuance. They all exhibit a larger degree of the fall, which is still enhanced by final lowering and hence exhibit a larger degree of finality than the non-downstepped accents. Final downstep is definitely very frequent in the languages of the world. It has also been described for German by B. Peters et al. (2005) who note that at the end of conversational turns, the final fall is especially deep and realized in combination with an early peak: "[...] dass turnfinale Gipfel überwiegend stark fallen, wodurch in Kombination mit der frühen Gipfelposition der Tiefton zusätzlich fokussiert und damit die Funktion des Abschließens verstärkt wird." (Peters et al. 2005: 18). Usually, this type of accent occurs at phrase end, but repeated !HL accents that form a downstepping terrace contour may be occasionally found, expressing what Ladd (1980, 1983) called stylization, adding a the nuance of stereotype to the utterance.

Downstepping terracing contours are usually regarded as a variant of what is called downdrift in the present study. Both contour types have been treated as a phonological choice or as an intonational morpheme (van den Berg et al. 1992) in the description of the intonation of West-Germanic languages. In EA, we have seen that the downdrifting contour is the default contour for neutral declarative utterances. The difference in the markedness of the contour in the two types of language is due to the way lexical prominences are articulated phonetically in these languages. Thus, the downdrifting pattern of EA is the equivalent of the hat pattern in English. However, in the model suggested here, both types are a 'morphological', i.e. a meaningful option.

\textsuperscript{28} According to Gussenhoven (2004: 105), this term was used by Meeussen (1970).
When accents are concatenated they show different degrees of cohesiveness. Bolinger invokes the dimension of separateness and connectedness between accents (or profiles), noting that “[p]rofile B differs from A along the dimension of ‘connectedness-separateness’ and all the semantic inferences that can be drawn from this. Where A singles things out, B ties them in” (Bolinger 1986: 166).

In this study, I will be using the terms integration and separation to account for the syntactic and semantico-pragmatic functions of the contours (4.5.5, ch. 6). Formally, cohesiveness may be described as linking between accents. Gussenhoven (1984) invoked a special ‘tone-linking rule’ to account for the cohesiveness between individual accents in British English intonation, I will apply Gussenhoven’s linking rule to account for the different types of concatenations found between accents in EA. In Egyptian Arabic we can observe similar types of linking as the ones Gussenhoven identified in English (Figure 26).

Assuming three different accent categories, the closing type, the leading(-linking) type and the default accent, each of these types may in principle be concatenated with one of the other two types, resulting in nine potential configurations. Figure 26 illustrates the three types that are relevant for the integration-separation dimension. It is, of course, the shape of the first accent that is relevant for the indication of cohesiveness, while the shape of the second accent does contribute to this dimension. (26a) shows a concatenation of two closing accents with no linking, signifying maximum separation and thus giving each of the semantic items its own weight, (26b) is the typical case of two default accents with merged L tones signifying loose connectedness, whereas in the case of complete linking (26c, d) a leading-linking and a leading accent are concatenated with a closing accent, respectively, producing the well-known hat pattern and the roof pattern (see 2.2.5.2). This final type of linking signifies maximum connectedness and is used to indicate what will be called Integration I in Chapter 4.5.5. In the next section (2.2.5) we will turn to some uses of these integrated contours.

In addition, we can observe that the linking may happen in a way as to produce a leading a closing trendline. A special case of linking the final two accents is final downstep as depicted above. Figure 27 is a schematic illustration of the different types of linked accents that can be observed in EA. I do
not assume though that the accents are linked by a phonological rule, just as I do not assume that the 
default accent is an underlying representation. Closing and leading accents are in the first place 
perceptionally established categories and, as already pointed out, alignment and scaling features are at 
play in realizing these categories. Thus, it is predominantly the realization or non-realization of an L2, 
its association with the segmental material and its scaling value that produce the auditory impression 
of a clearly rising versus a clearly falling accent. At the two ends of the continuum stand the plain rise 
without any L2 and a concomitant late peak alignment, to be represented as LH\textsuperscript{\neg} contrasting with the 
different types of clear fall, to be represented as (L) HL\textsuperscript{\textdagger}, (L, (L', !HL or (H)*L\textsuperscript{\textdagger} or (H)L\textsuperscript{\textdagger}. There is, 
of course, a certain amount of overlap and ambiguity in the in-between-area, which will be discussed 
in more detail in section 2.2.7. The default accent seems to be pragmatically 'neutral'.

While the accents in Figure (27a-c) are concatenations of leading accents, the ones in (21d-f) 
concatenate closing accents.

<table>
<thead>
<tr>
<th>Leading/Linking Tendlines</th>
<th>Closing Tendlines</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a) LH\textdagger L\textsuperscript{\neg}H</td>
<td>(e) (L)HL HL</td>
</tr>
<tr>
<td>(b) LH L\textsuperscript{\neg}H</td>
<td>(f) (L)HL !HL</td>
</tr>
<tr>
<td>(c) L\textdagger H\textsuperscript{\neg}H</td>
<td>(g) (L)H L\textsuperscript{\textdagger}</td>
</tr>
<tr>
<td>(d) LH\textsuperscript{\textdagger} H\textsuperscript{\neg}</td>
<td>(h) (L)H (!H\textsuperscript{\textdagger})L</td>
</tr>
</tbody>
</table>

Fig. 27: Concatenation of pitch accents producing leading (-linking) and closing trendlines.

2.2.5 The meaning of tones and tunes and their linguistic functions

In Section 2.1.5, I have already pointed out that the two identified accents and trendlines, leading 
(rising) and closing (falling), carry the iconic meanings related to the sound-symbolic functioning of 
the frequency code. The third tonal contour, the linking contour, is basically a flat tonal stretch that 
interpolates between two meaningful units and often characterizes what has been called ‘deaccented’ 
in accounts of English intonation (section 1.1.4). It may thus be attributed to the reversed effect of the 
effort code.

One important assumption of the present account is the functional equivalence of a single accent and a 
whole trendline. According to this assumption, function is signalled by the whole contour associated
with a linguistic constituent. If the constituent only consists of one prosodic word, say a one-word topic, the choice of the accent will be functionally determined. If the constituent is longer, the same function may be signalled by the trendline associated with that constituent, which may be the result of a series of upstepping (leading trendline) or downstepping (closing trendline) default accents, whereas the high linking contour is a result of a series of H accents that may be regarded as default accents articulated in a compressed pitch range. In the following paragraphs, the individual accent types will be discussed with the implication that the same is also true for longer contours. Examples for meaningful trendlines will be given in Section (2.2.5.1.).

As noted in Section 2.2.2, the ideas presented here are based on earlier work on intonation, mainly the work of Bolinger (1958, 1986), Brazil (1975, 1997) and Gussenhoven (1984). The identification of the three pitch accents is comparable to the suggestions made by Brazil (1975, 1997) who distinguished a referring tone (which in his system is a rise or a fall-rise), the proclaiming tone (a fall or a rise-fall) and a neutral tone. In Bolinger's and Gussenhoven's accounts there is no neutral tone. Bolinger's B profile covers the whole string of speech up to the following accent. With its basic notion of “being left up in the air” and “connectedness” (Bolinger 1986: 166, 277-292, 306-313) it is comparable to both, the leading and the high linking accents proposed here. In Gussenhoven's AM proposal the inter-accentual intervals are realized by phonetic interpolation while the variable beginning of phrases is accounted for by boundary tones.

The leading accent of the present model comes in two main versions. The first type usually starts at the beginning of the metrically strong syllable and rises to a point that is far beyond the accented syllable (LH’), frequently in the vicinity of the end of the lexical item it is associated with. In the second type, the rise starts from within the accented syllable after a fall from a higher level ((H)L*H). This type is clearly less frequent than the other type. Leading accents are frequently located at the end of an IP to signal continuation or questions. As will be demonstrated in Chapter 5, certain types of TOPICS are also realized with leading accents. At the end of an IP a leading accent may also involve a boundary tone (LH^H%) (see 2.2.6). Crucially, a leading accent does not have to be followed by a boundary, which obviates an analysis that derives the rising property of a tone from an obligatory phrase tone, thus supporting the assumption of an a priori independence of tonal contours from phrasing.

Functionally, the EA leading accent is related to Brazil's referring fall-rise, which he identifies as the typical referring contour, while he regards the simple rise as an intensification of the latter. Gussenhoven rejects the idea of intensification and proposes another view that ascribes two different meanings to the two accent types, namely "selection" for the fall-rise and "relevance testing" for the simple rise (Gussenhoven 1984: 201f.). The idea of ‘selection’ suggests that the "selection of a Variable from the background” (Gussenhoven 1983: 384) can be interpreted as the selection of a topic.
Unlike these two authors, Bolinger identifies the B profile, i.e. the simple rise, as the thematic accent (Bolinger 1986: 279). In EA, it is also the simple rise that is primarily used as a transducer to topic function, while the fall-rise is much more restricted, as is the low accented syllable in general. The empirical investigation of topics in the EA corpus (Chapter 5), however, suggests that a fall-rise (or ‘inverted’ accent) is predominantly used in a special type of contrastive topics, namely the final one in a series of parallel topic-comment constructions, thus pointing in the direction of Gussenhoven’s ‘selection from background’. However, this observation clearly needs further investigation. The differences between the functions of rises and fall-rises in the cited literature may also language-specific, as Brazil and Gussenhoven both refer to the intonation of British English while Bolinger refers to American English. If this is correct, EA is more similar to American English.  

The linking contour is a low-level tonal stretch, such as the one arising between two low tones. We may find it at the beginning of phrases, where the pitch is not really low, but mid-level, which is the typical starting point for the beginning of intonation phrases. Low linking contours may also arise at the end of phrases after a focal accent for signalling narrow FOCUS. In addition to the system laid out in El Zarka (2011), I propose a further (sub-)category, the ‘high linking’ contour here that is different from mere interpolation, involving clear prominences. The function of the second type of linking is somewhere in-between the downtoning of low links and the continuity function of the leading contour type. Thus, this contour has been illustrated under the heading of ‘leading-linking’ in Figure 27d. The high link may either contain the ordinary accentual rise with a level continuation to the next accented syllable or a slight 'sag' between the peaks or it may also be lacking that rise altogether. The justification for that contour to be different from a rising one comes from its pragmatic use on given items that are not necessarily selected as topics and from the fact that it is not used in questions, but frequently at the ends of phrases that are not associated with complete sentences, suggesting continuation. The leading and the high linking accents have a lot in common, both are used to signal continuation and both are used for ratified topics. In Bolinger’s system they are subsumed under the heading of the B profile. Likewise, they are both comparable to the different rising tones of the British tradition to which Cruttenden (1986: 163) assigns a universal “open” meaning as a cover term for all non-assertive and continuative meanings. However, Section 5 will provide more evidence for the distinction between the two categories related to the realization of different types of TOPICS.

Finally, the closing accent is comparable to Bolinger’s A profile to which he ascribes “separateness”, hence ‘singling out’ and “assertiveness” (Bolinger 1986: 164, 166; 277-306), to the proclaiming tone postulated by Brazil (1997: 67ff.) and the falling accent indication “addition” in Gussenhoven (1984: 201), and finally to the falling nucleus of the British tradition (Cruttenden 1986: 99ff.), the meaning of

29 I once had a student of Arabic who had been married to an English woman for a long time. He used to articulate topics with a complex falling-rising intonation which always stroke me as utterly un-Arabic.
which Cruttenden (1986: 163) refers to with the cover term “closed” including all assertive and non-continuous meanings. EA closing contours are associated with declarative utterances and related to FOCUS constituents as will be shown in Chapter 4.

2.2.5.1 Trendlines and patterns: relative height and the syntagmatic exploitation of tonal space

In this section, we will have a look at trendlines and whole patterns or tunes. Sequences of pitch accents may acquire a conventionalized meaning as their basic components together with their underlying meanings may collaborate to produce a conventionalized interpretation of the whole tune. Such is the case in the aforementioned hat pattern, which will be discussed in some detail below.

Egyptian Arabic utterances, especially declarative sentences, are typically falling throughout the whole utterance as already noted above. In contradistinction to these are utterances that exhibit a continuous upward trend or a suspension of downdrift (2.1.4). In the following I will try to show that downstep or downdrift and upstep and updrift are essentially manifestation of the same categories, i.e. leading and closing contours (i.e. accents OR tunes), respectively, both being ultimately related to the functioning of the frequency code in collaboration with the production code, thus expressing the same basic meanings such as finality and assertiveness or continuation and questioning (Figure 28 and 29).

An observation along these lines was also made by Bolinger (1972a, 1986) who notes that in a succession of rise-falls “[t]he effect of having the second higher, or lower, than the first is similar, despite the intervening drop, to that of a simple rise or fall.” (Bolinger 1986: 238).

An observation along these lines was also made by Bolinger (1972a, 1986) who notes that in a succession of rise-falls “[t]he effect of having the second higher, or lower, than the first is similar, despite the intervening drop, to that of a simple rise or fall.” (Bolinger 1986: 238).

Fig. 28: Sequences of two leading (panel a) and closing accents (panel b). The dotted lines represent the end of the lexical constituents (content words) the accents are associated with.

Fig. 29: A leading (panel a) and a closing tangent in a sequence of accents.

The problem of accounting for global trendlines in AM theory has already been mentioned. We have seen above that the so-called ‘overlay’ models which have mainly been developed for pitch accent
languages where the accent shapes of the lexical accents are specified in the lexicon understandably resort to another component of the intonation contour that encodes meaning, namely the trendline on which, in line with Bolinger's wave metaphor, the individual accents ride like the waves on the swells. Bolinger (1972a) shows numerous examples of relative height phenomena producing rising and falling trendlines which he attributes to "[a] partially grammaticized layer [...] of 'controlled' affective meanings" (p. 137). The intonational effect Bolinger discusses is a mismatch between grammatical form (e.g. imperative, interrogative) and tonal contour (overall rise or fall). When such a mismatch exists, intonation may override grammatical meaning and a new meaning will arise by inferential interpretation.

For Egyptian Arabic, Norlin (1989: 48) reports a suspension of the expected downward trend in ‘neutral questions’ (i.e. yes/no questions) plus upstep of the final accent. Rifaat (2005: 55), in his sketch of Egyptian-style MSA intonation also briefly notes the need of a phrasal component for the representation of this variety, and Rastegar-El Zarka (1997: 358) claims that wh-questions of the same variety start at a higher pitch level than declarative utterances without exhibiting a large pitch drop after the question word which would be indicative of a narrow focus on the question word. Such cases may well arise in echo questions where the wh-word is relatively more prominent than the backgrounded remainder of the question. In MSA wh-questions - contrary to questions in EA - the wh-word is always in initial position, and the intonation contour of ordinary wh-questions exhibits an overall fall to a relatively low value, thus being similar to wh-question contours in English or German. As I have pointed out above, the proponents of an overlay model claim that different trendlines signal different types of sentence mode, notably declarative and different types of questions. This indeed would be a problem for a linear approach as Standard AM theory, at least as long as the theory does not allow for any preplanning. However, as far as the results of the present study show, the only requirement for a descriptive framework seems to be a device to account for relative height from one accent to the next, i.e. to note whether that accent is higher, lower than or at the same height as the one before. This descriptive convention does not obviate the theoretical assumption that it is the whole contour that expresses the meaning and that this contour may very well be pre-planned. An example that illustrates that fact is the overall rising intonation phrase for a topic constituent and the overall falling contour for the comment part shown in Figure 30.
Fig. 30: Rising topic phrase *il-bajjaʕi:n hɛna:k* (DEF-vendors there) followed by a falling contour on the comment *bijɣallu kullə ħa:ga* (they.make.expensive all/everything) ‘The vendors there make everything expensive.’

The crucial point here is that the suggested phenomenon does *not* involve prominence differences between the individual accents apart. Different peak height in the upstep cases does *not* correlate with different metrical strength. We will return to that issue in section 2.2.5.3 where we will compare these cases with pitch range variation that *does* involve different prominences.

### 2.2.5.2 *Integration: the hat pattern and the roof pattern*

This section presents some linguistic functions of *Integration I* in EA utterances. At this point I will not yet get into the matter of pragmatic uses, this being the main topic of the following chapters of this study. I rather start with an illustration of some types of syntactic constituents that are frequently marked by the integrated pattern, thereby introducing the notion of integration in order to make use of it in the chapters that investigate the interaction of information structure and prosody (ch. 4-6).

In the following, I will use the transcription system developed above in addition to the more illustrative representation using schematic lines as in section 2.2.4 above. I will start with the hat pattern that consists of a leading-linking and a closing accent. It mostly combines only two content words in EA, and may only very rarely stretch out over more than two lexical items, if the medium one carries a linking tone for pragmatic reasons. Thus the hat pattern arises in comparatively short constituents whose parts closely belong together, such as the combination of a first name and a second name (12), for construct phrases, which are the semantic equivalent of a compound in Germanic languages (13), and generally for head-modifier constructions such as a verbal adjective and its complement (14), or verb series such as a combination of a coverb and a content verb as in example (15), which shows a suspended fall at the end signalling continuation.

(11)

\[
\begin{align*}
\text{HA:n} & \quad \text{LUT} \\
\text{LH} & \quad \text{!HL}
\end{align*}
\]

‘Hani Lotfi’
worst DEF-scenarios
‘the worst-case scenario’

full bee.COLL
‘full of bees’

and she remained she runs
‘and she kept running’

Syntactically, the hat pattern is very frequently used in singling out a prepositional phrase by tying together a preposition with its noun. Prepositions, belonging to the class of function words, mostly are not accented, but they may well be accented for rhythmic and/or emphatic purposes (16 and 17).

inside tree
‘inside a tree’

on DEF-frog
‘[searching] for the frog’

As the following examples show, the hat pattern may also occasionally cover short sentences, consisting of a topic and a comment (17) or a verb and an object (18) and as example (19) shows, it may also occasionally consist of three (or very rarely more) accents, the middle one being a linking accent.

‘What does it have to do with me?’ or ‘I don’t care.’
(18) \[ ma-ti\textsuperscript{\text{-}MIL-f\text{-}S\text{-}T} \]
\[ LH \textsuperscript{\text{-}H} \textsuperscript{\text{-}H} \textsuperscript{\text{-}L} \]
don’t make noise
‘Don’t make a noise!’

(19) \[ b-jib\textsuperscript{\text{-}aj\text{-}Z\text{-}G\text{-}ar\text{-}a} \]
\[ LH \textsuperscript{\text{-}H} \textsuperscript{\text{-}H} \textsuperscript{\text{-}L} \]
it becomes his reaction like the-carrot
‘His reaction will be like that of a carrot.’

In general, the hat pattern is used to combine two in most cases approximately equally strong accents to express their connectedness. The construction indicates that the thus prosodically articulated constituent forms a unit in semantic and syntactic terms or that the constituents involved stand in a pragmatic relation of topic-comment to each other. This construction stands in contrast to a construction where one of the two accents is subordinated to the other. Such a construction will exhibit the prosodic subordination of either the first or the second accent. This involves less prominence of the subordinated item, which is expressed in the tonal component through compression of tonal space (see 2.2.5.3). This type will later be called Integration II (4.5.5). Another possibility is to integrate the item into the rising part, or more frequently, the falling part of an accent with concomitant reduction of intensity.

The roof pattern or pointed-hat pattern is a variant of the hat pattern consisting of a leading and a closing contour. It is more frequently to be found stretching over two phrases that somehow belong together, frequently also involving rhythmic structuring into two parallel parts. The roof pattern seems to be an intensification of the hat pattern, i.e. the leading and closing parts are intensified, respectively. I do not believe the difference to be categorical, though. When two accents meet at the top of the ‘roof’, the first one mostly has a late-aligned peak and the second one a downstepped accent with an early peak (21 to 23). The roof pattern basically occurs in the same contexts as the hat pattern, like a construct phrase, such as NPs involving numerals (21). In examples (22 and 23) two phrases (NP + PP) are linked; example (24) illustrates a parallel structure of two clauses with a strong semantic connection between them. This also applies to (25), where the relative clause at the beginning is the subject of the whole sentence. Finally, example (27) illustrates an alternative question with a leading first part and a falling second one. Similarly to the example in Figure 30, it shows a leading contour containing more than one accent and a closing one of the same type with a break in-between. Generally, a break between the two parts is more likely than in the hat pattern, but it is by no means obligatory. The roof pattern in example (26) does not exhibit a break between the topical temporal adjunct imba:rih ‘yesterday’ and the rhematic main clause ifiare:t t’ama:t’im ‘I bought tomatoes’.
Apart from accented falls, the same melodic shape arises when function words are cliticized to the accent bearing content word to their left. Compare (28) to (25), where the two functional morphemes, the complementizer and the relative particle, are characterized by a linking contour interpolating between the two accents. The difference is due to the first accents in the two utterances. While the first accent in (25) is associated with a closing contour, the first accent in (28) is rising, causing the function word to be integrated in the falling part (see Figure 13).
(27) HA:ni kan bijiSab i l-gine:na
H. was he.is.playing in-the-garden
LH— LH L L*—%
‘Hany was playing in the garden.’

The above examples show how Integration I (the roof and the hat pattern) is used to indicate high semantic cohesiveness between semantic items, regardless of the syntactic constructions applied. This integration is indicated by the tonal shape of the contour and not necessarily by phrasing. Consider that the examples in (24) and (25) both involve independent clauses without a phrase boundary between them. In (26) and (27) the first constituent, associated with a leading contour, is a frame TOPIC in both cases, but it is only separately phrased in (26). The difference between a roof pattern and a simple accent is the number of prominences associated with the rise-fall. While a simple accent has only one prominence, a roof pattern involves two. In Chapter 4-6 the pragmatic exploitation of the integrative patterns will be thoroughly examined. However, even the above examples make it clear that an approach that regards syntactic and prosodic patterns as independent and furthermore differentiates between tonal shape and accentuation as prominence is superior to an approach that defines prosodic phrases in terms of syntax an regards pitch accents as markers of prosodic structure.

2.2.5.3 Accentuation and prominence: relative weight and the vertical (paradigmatic) exploitation of pitch range

We are now going to tackle the second part of the issue of pitch range variation discussed in connection with the representation of trendlines in section 2.2.5.1 above. As already noted several times, AM theory in general does not readily include the variation of pitch range in a phonological representation. And indeed, the distinction between what is paralinguistic and what is linguistic, for instance between emphasis and the prominence relations made use of for the indication of focus, is probably an impossible task. Ladd (1990, 1996: 279, 2008: 307) suggested a metrical representation of tonal space also to be applicable to the Anna/Manny relation in the well-known experiment by Pierrehumbert and Liberman (Pierrehumbert 1980; Liberman & Pierrehumbert 1984) elicited the production of two different intonation contours for the utterance Anna came with Manny. The intonational difference was one of information structure. For the experiment, the FOCUS was controlled by two context questions: 1) What about Anna? Who did she come with? producing an answer with Manny as narrow FOCUS (Fig. 31a) and 2) What about Manny? Who came with him? inducing narrow FOCUS of Anna (Fig. 31b). The experiment involved multiple repetitions with varying degrees of emphasis (on a scale from 1-10). The results show remarkable stability in the relation between the peak values of the two accents and remarkable stability of certain tonal targets, such as the final low value in the ‘answer’ part. In the standard approach, the actual pitch values are a
matter of phonetic implementation, i.e. they are calculated from the relative prominence value determined in the metrical component and local emphasis. As Ladd argues, the model thus cannot make any predictions about pitch relations. Furthermore, under the assumption that pitch variation is freely variable as a phonetic correlate of emphasis, the observable stability in pitch height relations under different degrees of overall emphasis remains unexplained and just a remarkable coincidence.

![Diagram](image)

Fig. 31: Two renditions of the utterance *Anna came with Manny* with two different FOCUS structures: narrow FOCUS of the initial argument *Anna* (panel a) and narrow FOCUS of the final argument *Manny* (panel b); from Pierrehumbert (1980: 266, 270).

Clearly, Ladd has a point in suggesting that the relationship between the two accents is linguistically meaningful. However, the relationship between the accents as wholes differs from the relation between downstepped accents in an overall downdrifting contour, for example. For one thing, that the answer to the question asking for the propositional object could also exhibit a second peak lower than the first one, i.e. a downstepped accent, in a rather detached, matter-of-fact rendition. The pointed-hat accent on Manny in the above example rather seems to be a matter of involved and more emphatic articulation. Nevertheless, the relative weight relation in metrical terms and the shape of the accents are the same in both utterances. Thus what is linguistically relevant as an expression of information structure seems to be the weight or prominence relation and accent shape rather than the peak relation alone.

30 ‘Metrical’ here has the original meaning of metrical stress (Liberman & Prince 1977 and subsequent work), not the high-low relationship suggested by Ladd.
Let us first see what another position within AM theory has to say about that matter before we turn to the case of EA and the solution suggested in the present approach. As already noted, Gussenhoven and his colleagues take an approach to downstep that is in a sense similar to Ladd's position. Contrary to the Pierrehumbert model, this approach does not consider intonational downstep as a matter of accent type and its phonetic implementation, but rather as a morphological choice pertaining to whole phrases (van den Berg, Gussenhoven & Rietveld 1992; Gussenhoven 2004: 107ff., 307ff.).

Gussenhoven clearly distinguishes between this type of sequential lowering of pitch values from other 'vertical' pitch range effects such as highlighting a focused item. Pitch range expansion of focus accents like for instance in Japanese are handled as part of the phonetic implementation that takes place in a structural context (Gussenhoven 2004: 204). Thus he proposes two steps for focus marking in Japanese, a phonological one, namely inserting a boundary before the focused item to prevent downstep from applying, and second the phonetic boosting of the accent. Gussenhoven (2004: ch.5) explicitly points out that phonetic implementation may be brought under speaker's control and may be manipulated for communicative purposes. In Gussenhoven's approach, the thus arising form-meaning relations are taken to be what is universal about language, which in turn is related to the three biological codes, the Frequency Code, the Effort Code and the Production Code. But Gussenhoven (2004: 75) also points out that "meaningful control in speech production cannot always readily be distinguished from variation due to the occurrence of different phonological forms." He also concedes that in many languages the structural meanings do not differ from the universal meanings and that this is in fact the "unmarked situation" (Gussenhoven 2004: 79f.). Finally, Gussenhoven concludes that "in such grammaticalizations of the universal codes, the function is morphemic function; the form phonological" (p. 80).

The three different approaches to the phenomenon of pitch range variation make clear one thing: It is not at all clear what counts as linguistically relevant and therefore phonological in intonation and what is paralinguistic. Pitch range, it seems, is the most problematic of intonational aspects, but surely not the only one. The uncertainty, even among the strongest proponents of the 'phonological' approach, becomes obvious in Gussenhoven's statement above and in Ladd's concession that "[i]t is implausible that the global paralinguistic effects on pitch range, and what I have called metrical pitch range effects, are completely unrelated." (Ladd 2008: 308; emphasis added). Ladd goes on, referring to Bolinger (1989: ch. 9) that "resetting' the pitch range at the beginning of a new topic is too obviously related to the paralinguistic function of raising the voice for the two to be completely separate systems" (ibid.).

The approach I am taking here is to describe intonational variation where it is meaningful, whether this meaning relates to strictly linguistic categories or not. Thus I propose to include pitch height differences, such as the one related to the Anna/Manny example in the phonological description, even if it only expresses nuances of meanings, such as involvement as opposed to matter-of-fact. Thus, this
model lies somehow halfway between Ladd's and Gussenhoven's suggestions in that it concurs with the view implicit in Gussenhoven's account that downstep and pitch range variation for highlighting purposes are two different things. However, it is not clear to me how meaningful variation could be achieved by phonetic implementation. Moreover it is not immediately obvious why an early alignment of peak and a valley as opposed to later alignment is readily accepted as constituting two different phonological choices H+L* for broad focus and H*+L for corrective focus in European Portuguese (Frota 2002) while any kind of pitch span variation is invariably attributed to the phonetic implementation.

What then is the relevant difference between these two phenomena - downstep on the one side and 'paradigmatic' scaling differences on the other side? The difference is obviously a matter of prominence involving excursion size. In a downdrifting contour (e.g. Fig. 6, 11 and 13), the individual accents are normally equally prominent. If the final accent is not downstepped, the downdrifting accents do not differ essentially in the size of their excursions. Contrary to that, different excursion sizes also produce differences in prominence. While the contour in Figure 32b involves manipulation of the pitch span (Ladd 1996: 260), i.e. the vertical excursions of the accents differ in size, the contour in (33a) can be interpreted as a gradual change of 'pitch level' (Ladd 1996: 260) with the vertical excursions of the accents remaining approximately the same. Downdrift involves a - frequently quite monotonic - repetition of configurationally alike accents at approximately the same level of prominence, whereas in case of the Anna/Manny experiment, not only two different accent shapes are involved, but also two degrees of prominence. Although experiments with artificial contours (Rietveld & Gussenhoven 1985; Ladd & Morton 1997) have shown that perceived prominence correlates with pitch height, everything else being equal, it seems that things are in fact more complicated. Based on other experimental findings, Gussenhoven (2004: 85) suggests that the listener's impression of stronger prominence "results from an estimate of the pitch span in relation to an estimated pitch register." Furthermore, weight differences are not necessarily expressed by pitch phenomena. Consider the case of low monotonal accents, for instance, total downstep. What experimental work can do is abstract away from different factors and establish some facts on a single dimension – e.g. pitch range – *everything else being equal*. But in real language, everything else is almost never equal. We must therefore not forget the various other factors that influence our perception of prominence, such as phonetic factors like intensity and time dimensions as well as positional effects. In the realm of intonation however, we can cautiously say that pitch range phenomena may exhibit two different types, the linear modification of register lines involving the syntagmatic iteration of segments with equally sized tonal spaces as in the downdrift case (Fig. 33a) or the vertical paradigmatic expansion or compression of tonal space as in Figure (32b).
The reluctance to accept pitch range variation as structural, hence 'phonological' is based on two major assumptions that lie at the heart of the AM approach, the principled divide between the linguistic and the paralinguistic in intonation and the resulting view that accentuation is categorical and that all variation that is gradient belongs to the realm of paralinguistics while discrete distinctions are linguistic. The other axiomatic assumption is that there is no preplanning and pitch levels are determined in a strictly local way. As repeatedly pointed out in this chapter, the view I take here is that between what is paralinguistic and what is linguistic in intonation cannot be neatly distinguished. I will make the case here of linguistically meaningful differences in the degree of accentuation, which of course implies that accentuation is gradient, not all-or-none. I have defended that view above in section 2.1.3. Evidence comes from languages such as Egyptian Arabic, in which less prominent accents occur in the place of deaccented material in English, as in the early narrow FOCUS case illustrated in Figure 33. As an aside, I would like to point out that complete deaccentuation has also been doubted for English. Pierrehumbert (1980: 223) reports the existence of “echo accents” after the nuclear accent. Bolinger (1986: 127) notes that to say that there is no prominence at all in terminal level contours “is an overstatement” and that the longer such a contour is, the more likely prominences are to occur. I take these observations as evidence for the present approach that firstly assumes gradient accentuation and secondly assumes a gradient difference between a ‘deaccenting language’ such as English and a ‘non-deaccenting language’ such as EA. In fact, the difference is that EA more relies on pitch more severely to mark the lexical word accents than a prototypical stress-accent language like English. This is why we find no or very little deaccenting in EA. As a corollary, EA and English are not categorically different in their employment of accentual prominence for linguistic purposes, which has a bearing on focus prosody as we will see in the following chapters. However, as the perceptual difference between the prominence of a strong accent and following deaccented items is greater than between a prominent and a less prominent accent, a deaccenting language is more likely to grammaticize accent position.
Fig. 33: Different pitch range extensions due to pragmatic factors. In the pitch track on the left the portion in between the horizontal dotted lines is characterized by a wide pitch range and in the pitch track on the right, the pitch range of the segmentally identical portion is compressed. The wide pitch range is associated with FOCUS constituents whereas the compressed pitch range is associated with post-FOCAL constituents.

As the EA example in Figure 34 shows, narrow FOCUS induces tonal space compression right after the focus accent, a fact that causes all post-focal accents to exhibit much smaller pitch excursions. Post-FOCAL accents are thus much less prominent than the focus accent. The EA facts are thus reminiscent of the situation in Chinese as described by Xu and his colleagues (Xu Xu 1999; Xu & Xu 2005). Interestingly, as Figure 34 indicates, it is not the phonetic boost that renders the focus accent more prominent, but rather the compressed pitch span of the following accents. Figure 34 also shows that the compression of pitch range is not only brought about by lower peaks, but also by a higher value of the final low, i.e. by narrowing the 'grid' from above and from below. A quantitative study, investigating these phonetic cues will be presented in Chapter 4.3.1.1. It will be argued in the next subsection and in more detail in ch. 4 that this relative size of tonal space is a means for backgrounding textual material. It is thus the EA equivalent of what has been called deaccentuation in descriptions of the intonational systems of West Germanic languages. In the present account complete deaccentuation is regarded as a special case of downtoning.

If the above assumption is correct, this constitutes evidence against an unwarranted distinction between the tonal and other phonetic cues to prominence made in AM theory. I therefore reject the assumption that FOCUS is marked linguistically by accentuation in English, but not in Egyptian Arabic. It has been argued that there is no prosodic reflex of information focus in EA except phonetic pitch range variation for contrastive focus (Hellmuth 2006a, 2010), which should be interpreted as paralinguistic, i.e. emphasis. A theory that rests upon the belief that emphasis and focal highlighting may be neatly kept apart, therefore has to make these assumptions, whereas the view taken here acknowledges the interrelatedness of what belongs to language and what belongs to paralanguage in intonation and thus is more flexible in dealing with these problems.
Deaccenting and downtoning

The contrast between figure and ground may be minimized resulting in a flat contour with only a very small excursion or even complete absence of the H-tone. These compressed accentual and occasionally even ‘accentless’ phrases may occur for example at the beginning of intonation units and are used for downtoning presupposed constituents, as shown above in Figure 35. By contrast, as we have already seen in the discussion of total downstep (2.2.4.3), monotonal accents do not have to be less prominent. In such cases, other phonetic cues have to step in where a tonal correlate of prominence is lacking. We have already seen that duration plays a major role in utterance-final position. In Chapter 4, I will argue that intensity is important, too.

The gradient view on accentuation implies that it really is a matter of relative prominence and thus among other factors - also a matter of relative height and excursion size of accents, which accent is perceived as the strongest one. Egyptian Arabic intonation contours show various examples of that meaningful prominence and pitch range differences. In EA, function words are typically realized without an accent and tonally encliticized to the preceding accent or assigned a linking contour, both strategies play them down in the utterance so that they may function as the background for accents on the content words which are more important to the message to be conveyed. Similarly, content words that are less important may receive a less prominent accent or sometimes be deaccented altogether. Accentual prominence will be thoroughly discussed in Chapter 4.3, suffice here to say that differences in accentual prominence also indicate differences in semantic and pragmatic weight. Figures 34 and 35 may illustrate this claim. The contour in Figure 34 depicts an answer to the question Where did you see Maryam? sufna marjam fi bina:jit it-tiligizjo:n (saw.w.e Maryam in building the-television ‘We saw Maryam in the television building.’). The reduced accent on the verb is due to the givenness of that concept and, crucially, to the action being presupposed, whereas the equally given name of the person in question, Maryam, receives full prominence to indicate its topical status, and so does the prepositional phrase which serves as the comment on the topic.

Fig. 34: Answer to the question Where did you see Maryam?: sufna marjam / fi bina:jit it-tiligizjo:n ‘We saw Maryam at the television building.’
The second example (Fig. 35) depicts the downtoning of a verbal phrase after a narrow argument FOCUS in the sentence il-ʔakla di / TE:TA lli ʕamalit-ha (this dish / granny who she.made-it) ‘This dish – GRANNY is the one who made it.’, which was given as an answer to the question "Who made this dish?" The downtoned accent on ʕamalit-ha (she.made-it) ‘she made it’ is the equivalent to English deaccentuation. Although the contour is virtually flat, an accent is clearly audible, but this accent is also clearly less prominent than the accent on granny which is rapidly falling, the fall being continued over the relative particle illi. Thus, the whole construction of a syntactic cleft accompanied by prosodic marking serves as a signal to narrow focus.

![Graph](image.png)

Fig. 35: Downtoning of the verbal phrase a narrow argument focus on te:ta ‘granny’ by compression of pitch range of the final accent; il-ʔakla di / TE:TA lli ʕamalit-ha (this dish / GRANNY REL she.made-it) ‘This dish – GRANNY is the one who made it.’

2.2.6 Phrasing

The reader will have missed explicit reference to different types of phrases as for example identified in the Prosodic Hierarchy (Selkirk 1984; Nespør & Vogel 1986). Phrasing is perhaps the most elusive of all prosodic phenomena and its definition is unavoidably circular. In many models phrases above foot level are defined on a syntactic basis, although since the advent of optimality theory mismatches between prosodic and morphosyntactic domains are more readily acknowledged, as violations of the alignment between the two types of domains are tolerated. The extrinsically defined phrases are then considered to be the domain for pitch accent distribution and the occurrence of boundary tones. On the other hand, phrases are also identified by exactly these phenomena. The occurrence of a specific tone is made the basis for the identification of a phrase end, such as the occurrence of the so-called phrase tone or an especially high or low tonal configuration may lead to the assumption of a boundary tone and thereby to the identification of an intonation phrase. But theories differ to what they assume to be edge tones and what they assume to belong to the pitch accent. One striking example is the difference between ToBI models that would analyse the EA default accent as a rising accent L+H* (e.g. Hellmuth 2006, et passim) plus a low phrase tone L, while the ToDI model would analyse it as falling accent H*L with a low trailing tone (e.g. Rastegar-El Zarka 1997). Both theories have their arguments, which will not be discussed any further here. Especially, as the issue has already been dealt with in
section 2.2.3 and the solution, of representing the EA default pitch accent as basically tri-tonal has been argued to be a way to avoid the problems that arise with either of the two models.

It seems that the problem involved in phrasing can never be totally avoided. In the present account I will assume that phrasing and melody are a priori independent phenomena, although they definitely interact. Coming back to the idea of rhythm and melody as conceptually independent components of tune, I take phrasing to be a correlate of rhythm. Low-level rhythm, as it has been referred to above, works on the level of the syllable and will not concern us here, we will only be concerned with higher level post-lexical rhythm on that is instantiated by accents. In this section I will look at three domains that are especially relevant to EA intonation, the intonation phrase, the prosodic phrase and the prosodic word.

To begin with, it is easy to demonstrate that morphosyntactic phrases and intonation phrases do not always coincide. Examples like the one in (26) are legion in the corpus. Similarly, connectors that may be viewed as syntactically proclitic (which is also reflected in writing), especially wa "and" and the determiner l- being are frequently separated from their syntactic hosts and incorporated in the preceding phrase, the same is true for various prepositions (26). Rastegar-El Zarka (1997: 200) noted the frequent occurrence of non-assimilation of the definite marker l- to word initial coronal consonants at the end of phrasal prosodic constituents. Watson (2002: 61f.) also notes that the syntactically proclitic determiner is encliticized to the preceding prosodic word. This phenomenon has its own logic and plausibility, as the elements in question occupying an end position before a pause function as cataphoric linkers to what is going to come in the next chunk of speech. If they are accented they are a means of focussing the separately phrased constituent as in example (28).

(28) Mido_01_M1_04
wasˤal sala / ?if-faˤtˤ
arrive.3SG.M on / DEF-beach
LH L°H / H′L%
'Mido arrived at the beach'

Of course, in natural language, phrasing is also employed to help the hearer identify the morphosyntactic constituents. Phrasing as a parsing aid is especially effective, where phrase boundaries depart from strict rhythm. On the other hand, in certain genres that lie at the edge of natural language, such as poetry and song (and for that matter nursery rhymes), the principle of syntactic transparency may be overridden in favour of a more strict rhythm. I will use example (5) of section 2.1.2, which is repeated here as (29) for convenience, to illustrate the methodology adopted in the present work in identifying prosodic phrases and intonation phrases. At the end of the section I will turn to the problematic, but all the more interesting notion of prosodic word.
If we have a look at the nursery rhyme (29) again, we will find that its articulation will preferentially abide by strict rhythm. We will assume a rendition that divides the rhyme into two equally long halves that coincide with the subject phrase and the predicate phrase, respectively. All through the whole rhyme trochaic feet are formed with only the penultimate one being dactylic. To get isochrony right in the rhyme there are numerous phonological adaptations.

(29)

// x - x - / x - x - / x - x - / x - x - / x - x - / x - x - //

hadi badi / mister M. / the-Badgadi // took.away-it and added-it / all.of.it on this

‘Hadi-badi Mr. Muhammad from Bagdad took away and put everything on this (side)’

The decision to posit two intonational phrases, indicated by the double slash, each one divided into prosodic phrases, indicated by a single slash, basically rests on perceptual cues supported by acoustic ones. The strength of the break is different. The first decision for the identification of an intonation phrase is always perceptual. It is fairly easy to identify a boundary when there is a real pause between two phrases. But this, of course, is not always the case. Sometimes there is no pause, but still there seems to be an audible break within a contour. As a first step I identified a break by listening to a stretch of speech and decided whether it is kind of finished or not, i.e. whether it could have been followed by pause, as it stands, or not. For the most part, this procedure was sufficient, when annotating the corpus of spontaneous speech. When there was doubt or generally in annotating the experiments, the acoustic signal was consulted in a second step to identify phonetic boundary cues, such as the longer duration of the last segment and syllable, and to a lesser degree of the stressed and penultimate syllables (Beckman & Edwards 1990; Jong & Zawaydeh 1999). Other cues are optionally a break in the tonal contour, such as a change of direction, the higher realization or a later alignment of an H tone, and the lower realization of a L tone, or F0-reset, i.e. suspension of declination and a reset of the baseline (Vaissière 1983: 57; Peters et al. 2005: 145). Other phonetic factors are glottalization or segmental phonological processes such as epenthesis, assimilations and reductions of all kind that may point to the existence or non-existence of a boundary (see also). This bundle of phonetic cues (Peters et al. 2005: ibid. may jointly serve to determine the depth of a break. However, the decision, where a phrase boundary actually occurs, is frequently difficult to take. Especially in spontaneous data, ambiguities are rarely resolved as noted by Brown et al. (Brown et al. 1980: ch. 3). These authors (p. 43) suggest that ambiguities or unclarity seems to be tolerated by the interlocuters and understanding may be reached according to Grice's (1975) conversational maxims in a co-operative process. Thus, different analysts will typically make different decisions as to where they assume phrase boundaries to occur, and even the same analyst's decisions will not be consistent when taken at different occasions (Wichmann 2000: 12).
As will be shown in the next subsection, IPs may also exhibit boundary tones. Yet, a high rise or a low fall alone is not necessarily indicative of an IP boundary. Although an abrupt change in the F0 normally points to the existence of a boundary, the independence of tonal and metrical component is also supported by the opposite fact. It can be observed that a tonal movement may be continued over an IP boundary, even if it is accompanied by a silent pause.

Another type of cue that has generally been assumed for the existence of a boundary is the appearance or blocking of phonological processes, such elisions, assimilations and the like (Nespor & Vogel 1986). As noted by Kenstowicz (1980), syllabification processes occur across word boundaries within a phrase. Watson (2002) also identifies the phonological phrase as the domain for phrasal epenthesis. Hellmuth's (2004) findings suggest that epenthesis is possible over major phonological phrase boundaries.31 This concurs with the results of the present study which suggests that epenthesis and phrasal syncope are not blocked by rhythmic prosodic structuring, only by the occurrence of an IP boundary.

The above example of the nursery rhyme shows the elision of the vowel in the first syllable of the name muhammad for rhythmic reasons or the elision of the vowel of wiwu ‘and’ and its encliticization of the glide to the preceding syllable, yielding [fæːluː] for example. The example in Figure 35 gives evidence of the elision of the high vowel /i/ in the first and the last syllables of the word binjaːjit (building) which in context becomes fi-bnaj-it-tilivizjoː:n applying syncope of unstressed vowels in monomoraic syllables (Watson 2002: 70ff.).

The upshot of the foregoing discussion is that phrasing is in fact a realisational phenomenon and therefore highly elusive. It is very much dependent on speech rate and idiosyncratic speech style. I will therefore not refer to phrases as the domains within which certain phonological processes will typically occur and as the syntactically identified metrical structures that are the domains of pitch accent association, as it is common practice in AM models (e.g. Hellmuth 2011). In the model presented here, phrasing is regarded as a phenomenon orthogonal to pitch contours, and phonological phenomena, such as sandhi phenomena, reductions or glottalizations are taken to be types of junctural features. For Egyptian Arabic, Hellmuth (Hellmuth 2011) proposed five hierarchically ordered prosodic categories: prosodic word, minor phonological phrase (largely equivalent to an accentual phrase), major phonological phrase, intonation phrase and utterance. She investigates the following phonetic cues that indicate the position of a break and assumes that the combination of at least two of them constitutes a boundary: local pitch reset, final lowering, pre-boundary lengthening, failure of epenthesis, pause or a phrase tone (H-/L-), whereby a high phrase tone is taken to be a boundary to a

31 It has to be noted, however, that Watson (2002) does not define the domain of her phonological phrase, it also is the only phrasal domain she posits. Watson's phonological phrase may therefore be equivalent to Hellmuth's intonation phrase.
major phonological phrase and a low one to an intonation phrase. The examples she cites suggest that a major phrase break is also indicated by an F0 reset of the first peak (that is a peak with a higher F0 than the preceding one).\footnote{Hellmuth distinguishes between reset and suspension of downdrift which implies a peak at approximately the same height as the preceding peak.} Based on these criteria, her results suggest that EA employs unusually long major phrases. She also assumes a constraint that phrases must comprise at least two accents.\footnote{Note, however, that Hellmuth’s speech data are predominantly constructed read sentences and therefore highly artificial language. While such experimental results say something on where boundaries are preferentially observed in read speech, they say nothing about where breaks occur in spontaneous speech. Thus they are of marginal importance to the objective of the present work.}

In the present model, I assume two levels of phrasing, \textit{prosodic phrases} (PP) and \textit{intonation phrases} (IP). As already noted, the identification of boundaries is predominantly done perceptionally. Pre-boundary lengthening and the non-occurrence of certain sandhi processes, such as vowel epenthesis and vowel deletion, are taken to be the only obligatory acoustic correlates of IPs. Tonal cues are not necessarily indicative of a boundary. The reasons for this are both theoretical and empirical. As laid out at length in the introductive sections, I assume the \textit{a priori} independence of the tonal and the rhythmic components of intonation and I consider phrasing to be a means of parsing out rhythmic chunks from the speech flow. The assumption that certain tonal formations necessarily indicate a boundary of some sort is contradicted by empirical facts. What is analyzed as a phrase or boundary tone in Standard AM models – e.g. a high rise or a low fall, the late alignment of an H tone – may occur in positions that are neither compatible with a break in morpho-syntactic terms, e.g. within a construct phrase or any other type of NP, nor is there a perceptually identifiable break in such cases. Consequently, the tonal behavior that is being accounted for by phrasing, phrase and boundary tones in most AM accounts is modeled using scaling and alignment features in the present account. This procedure warrants the detection of tonal patterns in within-phrase positions that might be overlooked in a model that takes phrasing as prior to the association of tones. A further problem with different types of boundaries and the prosodic constituents they are supposed to delimit is the gradient depth of the break, which is reflected in the phonetically oriented annotation system ToBI that provides different levels of break indices. To avoid this problem I only assume one type of break that may exhibit special types of terminals and edge phenomena, namely that of an IP. If a stretch of speech is perceptually complete, whether followed by a pause or not, this stretch is assumed to form an independent IP. If there is not enough evidence for such a break, no boundary will be assumed.

The \textit{prosodic phrase} is considered a purely \textit{rhythmic} unit. That means that the listener may perceive rhythmic units within a longer stretch of speech without perceiving these chunks to be independent IPs. When we listen to a speech signal and stop the rendition abruptly at the potential end of such phrases, the auditory impression is not that of a complete IP. It seems that such phrases could not be followed by a real pause without additional lengthening. There will typically be no edge phenomena
such as the non-application of epenthesis or vowel deletion at and around such boundaries. The salient acoustic feature that is indicative of prosodic phrase boundaries is F0 reset, which means that the peak of the first accent of a phrase is realized at the approximately the same height as or higher than the last peak of the preceding phrase. In section 2.1.1 the theoretical foundations for the assumption of this prosodic entity have been laid out, the example in Figure 36 serves as an illustration of an IP divided into three PPs.

![Figure 36](image-url)

**Fig. 36:** A downdrifting intonation contour subdivided into three prosodic phrases: //taqu:mu biha-l // /ʔana l-ʔa5la / fi-l-ʔansa:n// (it conducts it / the super-ego / in the human being) ‘that is accomplished by the super-ego of a human being’.

The example is taken from a Rastegar-El Zarka (1997: 178), where I have observed a tendency for PPs to contain two accents. This phenomenon is not to be taken as a minimum requirement, but rather as a rhythmic preference. The data investigated in Rastegar-El Zarka (1997) belong to a rather formal register, which is sometimes referred to as Educated Spoken Arabic (Badawi 1973; Mitchell 1984). The formal register and the slow speech rate that is associated with it promote the use of rhythmicized speech (especially in Arabic). As I have argued earlier (El Zarka (2005), Egyptian Arabic is a *strict rhythm* language, in which rhythm may sometimes be valued higher than grammatical structure. This claim is also supported by the conclusion Heliel (1976) draws on the results of experimental data: "In Arabic [i.e. EA, DE], at least, the temporal organization of spoken language seems to take place in terms of speech production units which are fairly independent of the morphological or syntactic structure of the utterances." (p. 317). In spontaneous speech, such rhythmic patterns are, however, very rare. Thus, the PP will not play a role in the present study.

However, the assumption of a certain prosodic unit as the obligatory domain for phonological phenomena proves to be at least highly problematic, as has also been recently shown in typological work on word domains (Bickel et al.). This work suggests that prosodic word domains are by no means universal and one and the same language may show multiple prosodic word domains for different phonological processes. In EA, the prosodic word has been assumed to include stem and affixes as the basis to identify the position of word stress (Mitchell 1993; Watson 2002: 61). At the
same time, Hellmuth (2007) suggests that the prosodic word is the domain for pitch accent assignment. However, the accentual domain (i.e. the domain covered by one accent) is not co-extensive with the prosodic word that is used for the identification of the metrically strong syllable on word level. The examples in (31) and (32) show a mismatch between the accent domain and the prosodic word. In (31a) the accentual domain is larger than the prosodic word, while it is co-extensive with it in (31b). On the other hand, the example in (32) shows four IPs, each of which contains one prosodic word. However the number of accents is different in the four phrases, i.e. an accentual domain may in fact be smaller than one prosodic word. In the following examples, the prosodic words are indicated by square brackets and the accent domains by curly brackets.

(30)  
\[
\begin{align*}
\text{il-mi'za \ bita:'it kama:l} \\
\text{the-goat POSS K.}
\end{align*}
\]

\[
\text{‘Kamal’s goat’}
\]

a) \{[il-mi'za] [bita:'it]} \{ka{ma:l}\} or \{[il-mi'za] [bita:'it] [ka}\ \{ma:l\}\}

b) \{[il-mi'za] [bita:'it] [ka{ma:l}]\}

(31)  
\[
\begin{align*}
\text{ka-ʔi\lamijji:n} / \text{ka-muθaqqaʃi:n} / \text{ka-muʃakiri:n} / \text{ka-masʔuʃi:n} \\
\text{as-journalists} / \text{as-intellectuals} / \text{as-thinkers} / \text{as-responsible.PL}
\end{align*}
\]

\[
\text{‘as journalists, as intellectuals, as thinkers, as the ones responsible...’}
\]

\[
\begin{align*}
\text{ka-[ʔiʃ]} \ \{la:mij\} \ \{ji:n\} / \text{ka-[mu{θaqqa} \ {fi:n}] / ka-[mu{fakki} \\
\{ri:n\}]} / \\
\text{ka-[{masʔu} \ {li:n}]} \\
\end{align*}
\]

Example (32) shows an accent on the particle *ka* ‘as, like’. Such particles and other sub-minimal function words\(^{34}\) have been claimed to be "unstressable" clitics, cliticized to the following content word (Watson 2002: 93) and are supposed never to be accented by Hellmuth (2007: 307). Proceeding from this assumption, Hellmuth argues that when accented, function words are promoted to prosodic word status. Otherwise they are not incorporated in the prosodic word, but, following a proposal by Selkirk (1996), are to be regarded as free clitics. Evidence to the contrary comes from naturalistic data, the MSA data investigated in Rastegar-El Zarka (1997) as well as the colloquial data of the present study, which show a fair number of accented functional elements, closely knit to a following lexical item as many of the previously cited examples have shown.

The content words in (32) also give clear evidence that double accentuations may even arise in long monomorphemic words. To posit two prosodic words in these cases is highly counter-intuitive. In other words, we may say that accents, in general, operate on the level of *actual utterances*, seeking out the lexically or rhythmically specified potential positions. Thus accents may partly be rhythmically governed, but are basically employed for pragmatic purposes. It therefore makes no sense to assume different prosodic domains for pitch accent distribution as in Hellmuth (2007), because the actual realization of pitch accents, if a matter of rhythm which is a phrasal phenomenon. However, accent is

\(^{34}\) The minimal word in EA is obligatorily bimoraic (Hayes (1995); Watson (2002)).
employed for some highlighting function, in which case it makes direct reference to the semantic-pragmatic structure of the utterance. An approach that distinguishes between tonal and rhythmic prosodic units does not run into the same empirical problems. Furthermore, the assumption that accents are assigned directly to semantic and functional categories instead of prosodic ones also allows to predict some metrically unexpected, but actually occurring accentuations, and metrical-tonal mismatches such as early and late peaks or the accentuation of metrically weak syllables are given a plausible explanation.

2.2.6.2 Boundary tones

The importance of terminal contours in intonation is common knowledge, not only in linguistics, but of linguistically untrained speakers. As already noted, AM theories make a fundamental distinction between edge tones and pitch accents. In the view I take here, in principle, a tone is a tone, and a contour is a contour, whether it arises in the middle of a phrase or at the end. To illustrate this claim, we may again return to (31), the pitch track of which is given in Figure 37.

The utterances in Figures 37a and 37b are segmentally identical, they are, however, answers to two different questions, replicating Jackendoff’s Fred ate the beans example35. The utterance in Figure 37a is an answer to the question ‘What did Kamal’s goat eat?’ and Figure 37b answers the question ‘Who ate the beans?’ As we can see the number and identity of tones does not differ in the two prosodic rendition, but what clearly differs is the alignment of the tones. If Kamal’s got is FOCAL (Fig. 37b), L2 is aligned much earlier, namely at the end of the content words which gives the accents a falling shape and turns them into a closing accents. In Figure 37a, a default accent is associated with the first word (mišča ‘goat’) and a leading accent with the second word (Kamal) marking the constituent as a topic. Note that an IP boundary is only present after the Kamal in (Fig. 37b), the first accent is tonally identical to the second one. This obviates the analysis of the second accent as involving a boundary tone. Moreover, in the topical constituent, the first accent also exhibits an L2. Attributing this to a boundary or phrase tone we would have to stipulate a phrase boundary at the end of the possessive marker bita:ɔ-it separating the first part of the semantically and syntactically closely knit possessive construction from the second one.

35 The example is from the translation task in the Questionnaire for information structure (QUIS) (Skopeteas (2006)).
A last argument against analysing the occurrence of the low target as a boundary phenomenon can be seen in the contour shown in Figure 38. This is a read sentence without any preceding question, read in a neutral fashion, without putting special weight on any specific constituent - the kind of sentence characterized above as a ‘neutral citation form’. If we analysed the contour as consisting of three prosodic phrases we would have to explain the occurrence of downstep (downdrift) within the whole phrase which should be blocked by the prosodic phrase boundary.

To sum up the argumentation above, the two possible analyses that could be offered under a more standard AM view would be the assumption of a boundary whenever an L tone arises after the accentual rise that covers the accentual syllable or the assumption of a focal L tone. Both analyses are flawed. The option of a focal tone can immediately be dismissed as the L tone also occurs in the topical constituent. The boundary analysis has more to recommend it. But this analysis would have to predict differently sized phrases for every observable articulation (comprising one syllable, one word, one word plus clitics and so forth).

However, there are cases of high boundary tones in EA. In Chapters 4 and 5 we will see many examples of high boundary tones that can clearly be identified as an independent pitch rise on the final, metrically weak syllable of a word. The high boundary tone enhance the leading quality of an accent, yielding a high rising contour that is especially frequent in the speech of young people. An example can be seen in (32) (Figure 39).
(32) 28ARZ-A01GNT-00F1300

A:

\textit{\textit{jand-ik faRA:sha ?}}

at-you butterfly

\textit{LH^4H%}

‘Do you have a butterfly?’

B:

\textit{?AJwa}

yes

\textit{L^*H%}

‘Yes.’

Fig. 39: Two IPs ending in high boundary tones.

In the present model, phrasing is thus viewed as orthogonal to the tonal configurations. A boundary may however enhance a certain contour, in the case of a leading contour, by making it even higher.

2.2.7 Overlap and ambiguity

Like every model of intonation, this model also has its problems with categorizing configurations that are somehow intermediate between the posited categories due to the gradient variation of both, pitch and prominence. Some accents are clearly of the closing type as for example the ones in Figure 37b, and others are clearly of the leading type such as an accent with a very late peak alignment. Both types of accent can be enhanced by a higher or lower scaling of the H or L tones, respectively, or in case of the leading accent by a high boundary tone. The leading and closing properties may still be enhanced by a following break as can be seen in the last accent of the rising contour in Figure 37a above. But if we consider the first accent in Figure 37a, although its perceptual difference from the first accent in panel (b) is out of the question, the ‘leading’ quality of such accents cannot always be unambiguously established. I assume that the F0 value at the end of the semantic unit it is associated with is decisive for a contour to be perceived as rising or as falling. Another cue may be the velocity of the fall. These questions will be investigated in a production experiment described in Chapter 4.4.1.1. The results of this experiment give a first indication of the validity of the categories assumed. However, to verify
categorizations from the listener’s point of view and to identify the boundaries between them, we would need to conduct perception experiments that will exceed the scope of the present study.

One problem is that the realization of the targets is not totally under speaker’s control, but also subject to phonological factors, such as the amount of segmental/syllabic material available between two accent peaks and the position of the accented syllable within the word. Note that we have mentioned above that in case the material is not enough to perform a certain tonal movement, there are the possibilities of compression and truncation, which are not only dependent on the language or dialect in question, but also seem differ between speakers of the same variety. That means that a grey zone between the clear prototypes of the leading and the closing categories remains.
3. On the interaction of prosody and information structure

As already mentioned, the treatment of the information structural interpretation of intonation in prosodic research to date has been primarily concerned with focus (or FOCUS)\(^1\) marking. It is generally assumed that focus is marked by the strongest prominence or the nucleus, which is also regarded as a mere structural position in some approaches, of an utterance which more often than not is equivalent to the final accent. The preoccupation with nuclear accent placement in the literature on intonation is to be understood from the dominant role West-Germanic languages played in the early investigations of intonational patterns. Only comparably recently more and more in-depth phonetic studies on other languages have been carried out. This broadening of the data base has for example shown that there are languages that do not all that easily deaccent given information (cf. Cruttenden 2006 for a crosslinguistic survey) and whose intonation patterns – contrary to what has been found for English, German or Dutch – exhibit a dense pitch accent distribution, hence in these languages the above illustrated wave pattern abounds. This group comprises genetically unrelated languages such as Romance languages, North Germanic languages such as Swedish (Gårding 1983, 1998) and Danish (Thorsen 1980, 1983; Grønnum 1998), Estonian (Asu & Nolan 2007; and references therein) and Egyptian Arabic (EA) (Rifaat 1991; Rastegar-El Zarka 1997; Hellmuth 2006).

The basic assumption on which the present model rests is the idea that prosody and text are essentially independent of each other and are brought only together in the process of tune-text association as illustrated in Chapter 2. It is through this mechanism that linguistic meaning may be expressed prosodically. It has been claimed that in certain languages, accent is not as readily utilized as a means of marking the focus of a sentence. For instance, Ladd (1996: 178–179) suggested that it is not possible in Italian to place the main accent on a focused possessive pronoun with accompanying deaccenting of its nominal head. Instead, as Ladd observes, a morphosyntactic strategy of right dislocation of the noun constituent is used. The perfectly natural English sentence, given in (5) thus does not correspond to the barely acceptable literal translation in (6), but rather to the sentence in (7) in which the syntactic operation of right dislocation of the noun is utilized. The example is an authentic one said to a child by his mother after having finished bathing the child’s baby brother.

(5) Now I’ll run YOUR bath.
(6) Adesso faccio scorrere il TUO bagno. (?)
(7) Adesso faccio scorrere il TUO, di bagno.  (cited after Ladd 1996: 179)

---

\(^1\) In this literature review the term focus will be used in the sense it is used in the different studies at hand, which might be either in the meaning of focus or FOCUS as defined in the present model.
The dislocated noun is pronounced with a tag intonation. This allows a main accent on the narrowly focused pronoun without compromising the ban against deaccentuation within a phrase. Ladd therefore suggests that there is a distinction between “deaccenting” and “non-deaccenting” languages. Based on similar observations from Catalan, (Vallduví 1992) proposed a typological distinction between plastic accent and non-plastic accent languages, referring to languages such as English and German as plastic accent languages that exemplify a considerable freedom of nuclear accent placement, whereas many Romance languages fall into the category of non-plastic accent languages which are characterized by edge-marking nuclear (i.e. final) accents.

I have proposed earlier that this observed behaviour is ultimately linked to the prosodic properties of the languages in question (El Zarka 2005). I suggested that the prosodic structures of these languages abide by relatively strict rhythm instantiated by the high frequency of pitch accents in an intonation phrase. Strict rhythm and the general precedence of prosodic preferences, which I will come back to at the end of this chapter, can also be held responsible for the lack of deaccenting which would lead to long stretches of accentless syllables counteracting the rhythmic speech flow. Thus, the distinction between “deaccenting” and “non-deaccenting” languages is ultimately grounded in a distinction between strict rhythm and non-strict rhythm. Egyptian Arabic, being a prototypical example of a strict rhythm language, is a case in point. Hellmuth (Hellmuth 2005 and subsequent work) observed that EA resists deaccenting which is obviously true, although natural data clearly show that deaccenting is not impossible, it is only not favoured by the phonology of the language (see Chapter 4.3 for the empirical investigation). As predicted by Vallduví, EA, like the Romance languages, clearly makes heavy use of syntactic means in expressing informational categories. EA may thus be grouped with Italian among the flexible syntax languages as opposed to flexible accent languages like English to use two terms from Lambrecht (1994: 320). In the light of the natural and experimental data that are investigated in this work, Vallduví’s claim, however, clearly seems to be too strong. The data of EA rather suggest that prosodic marking of information structure, even though it may not be as powerful an instrument as it is in English, nevertheless accompanies the syntactic strategies, by virtue of the iconic basis of the prosodic features employed and due to the primary function of intonation to express emotion and attitude – which cannot be completely ‘turned off’.

It is certainly true that some trade-off between morphosyntactic and prosodic strategies exists, a claim that has repeatedly been stated for the signalling of different question types with and without morphosyntactic marking. Grønnum (1998), for example, showed that polar questions in Danish which are syntactically marked by inverse word order and wh-questions that are unambiguously marked by a question word have a steeper, more declarative-like intonation contour than ‘declarative’ questions that are signalled by intonation alone. Such questions exhibit the total suspension of downstep. A certain trade-off between morphosyntax and prosody most probably exists regardless of
whether a language is resorting to deaccenting or not. It has, for instance, been claimed by Lambrecht & Michaelis (1998), that WH-words that are held to constitute the focus of information questions in English, are not accented. This has been interpreted as a trade-off between accent and syntax in English (Lambrecht 1994: 285). Conversely, (Haan 2001) has provided experimental evidence that in Dutch information questions the WH-word always carries an accent while Chen’s study (2006) using natural speech data of the same language only ascertains a predominance of accented question words (77%). These few examples suggest that prosody obviously does not have the same degree of ‘grammaticity’ as e.g. word order or morpho-lexical marking. While the use of a question word is an indispensable means of signalling an information question, its position being determined by the syntax of a language, prosody or intonation is only a secondary instrument of signalling linguistic categories and we can at most determine certain tendencies for its grammatical use. It is, of course, well-known that intonation may turn a declarative utterance into a question, as stated for Danish above. This option is also available in English or German. There always remains, however, a certain flavour of assertion to the putative question. It could also be argued that such questions are not true questions, but rather incredulous statements uttered with a questioning attitude. This questioning attitude is mostly preserved in polar questions that are already unambiguously marked by syntactic inversion in English and German. This again supports the Bolingerian view that there is no categorical difference between the attitudinal and the grammatical in intonation, but that they are only two extremes of a continuum. Looking at Egyptian Arabic will give us an example of a strong conventionalization of the questioning rise at the end of an utterance that obviously by time has made it possible for the language to dispense with an overt question word which is present in other varieties of Arabic such as the Classical or Modern Standard Arabic hal or Moroccan Arabic waʃ. Contrary to that Egyptian Arabic may solely rely on intonation in signalling sentence mode in polar questions.2

The view held here challenges Vallduví’s suggestion of a dichotomy between languages that resort to prosodic marking of information structure and languages that do not. Although the basic observation implicit in Vallduví’s claim that a correlation between the non-deaccenting property or strict rhythm and a heavy functional load of morphosyntactic marking certainly exists, I will argue for a model that integrates syntax and prosody and that considers prosody as basically always present, being a largely involuntary epiphenomenon of general cognitive and emotional processes, albeit with the potential to conventionalization and ultimately perhaps grammaticization. This approach does not deny the vital role of intonation in speech perception and hence in the decoding process of the speech signal to uncover meaning, but it does not assume a biunique relation of grammatical or informational and

---

2 It has to be noted, though, that polar question may be expressed using a personal pronoun, huwwa (3.SG.M) or hijja (3.SG.F) as a question marker sentence initially.
intonational categories. In that sense the proposal made here is essentially in the spirit of Bolinger’s work.

3.1. Accent as a marker of focus

3.1.1 Terminological problems: accent and focus, two names for the same thing or two different concepts?

Before we turn to the question of how the interface between information structure and accentuation has been accounted for in various approaches, a terminological problem has to be clarified. When speaking about ‘focus’ different research traditions do not refer to identical concepts. One basic difference relates to the perspective from which the phenomenon is looked at in the individual research paradigms. In most syntactic and semantic work on information structure focus is viewed either as a semantic-pragmatic concept (e.g. Lambrecht 1994) or a syntactic feature as in the generative tradition following (Chomsky 1970) which is formally expressed by a phonological category, namely nuclear accent. In these models focus refers to a functional domain which is marked by a phonological unit, whereby the domain and the phonological unit, i.e. the accent, are not necessarily co-extensive. Based on Halliday’s remark that focus marks a part or the whole of a “message block” (Halliday 1967: 204), these models developed the idea of one accent being the focus exponent of a focus domain.

On the other hand, there are models that start out from a phonological basis of focus. In Halliday’s (1967) focus is phonological “highlighting”. To my knowledge, this tradition starts with (Bolinger 1954) who introduced the notion of information point (Bolinger 1954: 153) and Halliday (1967) who uses the term information focus for the main element of an information unit. Halliday posits as many information units as there are foci, leaving it up to the speaker how many information foci he chooses to convey. Thus, Halliday’s concept of focus is co-extensive with the phonological tonic (nucleus). The one-to-one correspondence of focus and accent is also fundamental in the approaches of Bolinger (e.g. Bolinger 1972b, 1986, 1989 and Chafe 1974, 1976, 1994).

Some lines of research on intonation still follow that tradition in equating focus and emphasis. For example, Hirst (1998), referring to Sperber and Wilson’s Relevance Theory (Sperber & Wilson 1986), explicitly denies the need of positing a linguistic category focus and regard the use of the term focus instead of emphasis essentially as a “trivial question of vocabulary (or fashion)” (p. 66). But also some of the structure-based accounts (a term coined by Gussenhoven (1985) that will be explained below) at least partly use the term focus in a sense that bears strong similarity to the use in the approaches that view focus and accent as essentially two sides of the same coin. One of these is Selkirk’s Pitch-
Accent-First Theory (1984). Selkirk followed Bolinger in the belief that accents are directly assigned to word-level constituents. According to her Basic Focus Rule, “a constituent to which a pitch accent is assigned is a focus” (Selkirk 1984: 207). Starting from this focal point, the focus spreads according to syntactic and semantic principles to cover a higher focus domain that dominates the originally assigned focus. Selkirk thus conceives focus as a recursive category. But also Gussenhoven’s use of the term focus, although defined as a semantic feature, is in fact based on the one-to-one correspondence of focus and accent (this point will become clearer in the discussion of his model, cf. Section 3.1.4.6).

I have provided this very short and eclectic exposition of different approaches, which, of course, does not do justice to the theoretical assumptions that lie behind these viewpoints, to show that, especially where it comes to intonational or prosodic matters the term focus not always designates an information structural category as outlined in chapter 1. A short review of some of the theoretical frameworks and their fundamental assumptions will be given in section 3.1.4.

3.1.2 The focus-to-accent conundrum

What has occupied linguists dealing with intonation and prosody for decades is the question of where the accents go, given a certain sentence form. In the beginning of the 1970s there existed two approaches whose difference is intimately related to the notion of predictability. The two ‘focus-to-accent’ approaches that dominated the field were (i) that accent is predictable from the syntax of a sentence and (ii) that accent is not predictable, but rather a question of the speaker’s choice who will put accents here and there according to what he wishes to put in the foreground or to ‘focus’ on (cf. Gussenhoven 1983: 377f.). This distinction can also be understood as a distinction between a ‘syntactic’ and a ‘semantic’ approach (Ladd 1980: 70–80). The first approach developed by the early generative school in the early 1970s (Chomsky & Halle 1968; Bresnan 1971, 1972; Jackendoff 1972; Berman & Szamosi 1972; Lakoff 1972) held that, given a certain syntactic structure, the sentence accent or nuclear accent is placed on the basis of that structure by rule. This assumption relies on the concept of normal stress which according to Ladd (1996: 160) goes back “at least to Newman (1946)”. Normal stress, a notion that was never properly defined (Gussenhoven 2008: 84), was regarded as a ‘neutral’ accent pattern with no particular meaning which is assigned by a phonological rule to the surface structure of a sentence. This phonological rule, the Nuclear Stress Rule (NSR), was explicitly formulated by Chomsky and Halle (1968) and assigns one primary stress to the last stressed vowel of the last lexical item of a sentence. Of course, this rule was not successful in deriving all attested accent patterns that did not exhibit the strongest prominence on the last lexical item of a sentence. These
marked patterns were subsumed under the notion of ‘contrastive’ or ‘emphatic’ stress and, not being amenable to a rule-governed account, did not find the interest of the generative theoreticians.

The normal-stress view was severely criticized by Bolinger (1972b), who forcefully argued against the dependency of accent upon syntactic structure and the distinction between ‘normal’ and ‘contrastive’ stress (i.e. accent), claiming that accents are the direct expression of the speaker’s intent and interest and that every attempt to predict their location is futile. His convincing argumentation has been heeded in later approaches that developed out of the conviction that accent is predictable from linguistic structure, although this structure is no longer seen to have a syntactic, but rather a semantic-pragmatic basis that takes into account contextual factors.

The early syntactic approach may be regarded as obsolete today (cf. Ladd 1996), but it had a strong impact on the further theoretical development of the structure-based accounts prevalent in mainstream linguistic thought today by making a conceptual distinction between focus and accent in denying the one-to-one correspondence implicit in Bolinger’s claims of a direct relationship between the two concepts. A new era may be said to have started with Susan Schmerling’s thesis (1976) that incorporated Bolinger’s critique of the normal stress view and the distinction between ‘normal’ and ‘contrastive’ stress made by the generative scholars. Important work by Ladd (1980, Ladd 1983) and Gussenhoven (Gussenhoven 1983) that developed a unified account of sentence accentuation based on linguistic structure followed suit. The debate between the two approaches is far from being settled, only that the dichotomy has shifted from ‘syntactic’ vs. ‘semantic’ to ‘structure-based’ vs. the ‘highlighting-based’ focus-to-accent approaches (cf. Ladd 1996: ch.5 for an overview).

3.1.3 ‘Broad’ vs. ‘narrow’ focus

As Ladd (1980) pointed out, the syntactic approaches laid the foundation for distinguishing between the two notions of ‘broad’ and ‘narrow’ focus, two terms coined by Ladd (1980: 74f.) to account for the differently sized focus domains that correspond to the old concepts of ‘normal’ and ‘contrastive/emphatic’ stress. Under this view, normal stress is the one accentuation pattern that leaves the width of the focus unspecified, while narrow focus means focus on a particular constituent. If we assume that focus is signaled by accent, it is evident that in the latter case there is a direct relation between accent and focus – whatever is accented is also focused. In broad focus cases, however, the relation is generally assumed to be an indirect one, accent placement being based on language-specific structural principles, such as Gussenhoven’s (1983) Sentence Accent Assignment Rule (SAAR).

---

3 The thesis was approbated in 1974, the published version of the thesis appeared 1976 which will also be the one referred to in this book.

118
structure-based account, focus on larger constituents or whole utterances may thus be signalled by a single nuclear accent which has scope over the whole focus domain. Moreover, in such accounts broad focus utterances are viewed to be the “‘unmarked’ or ‘default’ patterns” (Ladd 1996: 164).

Being useful for prosodic purposes, the broad-narrow-focus distinction also has to be related to information structural categories if a treatment of the interface between the two linguistic components is aimed at. In trying to do so, we find that this relation is not always clear. While the correspondence between narrow focus and information structural categories, at least at first glance, seems comparatively straightforward, the relationship of broad focus to pragmatic categories is far from clear. It is commonly assumed that broad focus, being the equivalent of normal stress, is the ‘neutral’ articulation of a sentence and thus signals focus over the whole utterance. Of course, ‘neutral articulation’ is not a type of linguistic or pragmatic category, as it cannot be related to any context. But Ladd’s formulation has been interpreted as focus over the whole sentence domain as typical of out-of-the-blue utterances. However, if I understand it correctly, Ladd’s conception of focus domain commences from accent placement and determines its scope, which is the lexical constituent (or even only part of it) in the case of narrow focus and anything larger than one lexical constituent in the case of broad focus. This becomes evident in his discussion of Halliday’s example

(8) John painted the shed yesterday.

when he says that “the focus can be the shed, or painted the shed or the whole sentence. Thus (8) could be used to reply to a range of questions like “What’s new, What did John do, What did John paint yesterday, etc. By contrast, other possible accent placements narrow the focus...” (Ladd 1980: 74, italics in the orig.). The same characterization of broad focus is more explicitly given in Ladd (1996: 162-163) in the following example:

(9)

\[
\text{I didn’t give him} \quad \{ \begin{array}{l}
\text{a dollar} \\
\text{fifty centimes} \\
\text{my notebook} \\
\text{your camera} \\
\text{the car keys} \\
\text{a sandwich} \\
\text{a lot of money} \\
\end{array} \} \quad \text{I gave him five francs.}
\]

According to Ladd, ‘five francs’ here is a case of broad focus. The choice of this example illustrates the fact that broad focus can be anything from ‘narrow’ argument focus to ‘broad’ sentence focus, if we want to match it with the information structural categories established in chapter 1. The above example also shows that even narrow focus cannot be unambiguously matched with an informational
category like argument focus, for instance. It only suggests that focus and accent coincide and that the accented syllable only projects a domain as large as the word it belongs to. Another terminological problem is the frequent confusion of narrow and ‘contrastive’ focus in the literature, going back to the terminological use of the term ‘contrastive’ for a marked focus position in the early syntactic approaches. To avoid such problems, a sharp distinction is made in this study between the different types of functional categories identified in chapter 1, and broad/narrow focus when speaking about the extension of the formal domains.

3.1.4 ‘Structure based’ versus ‘highlighting-based’ focus-to-accent approaches

In this section a brief overview of some milestones in the development of the focus-to-accent idea will be given, stating the different positions as regards the fundamental question whether focus is directly expressed by accent or whether this relationship is rather an indirect one mediated by focus structure. In the above section the difference between the two major theoretical paradigms concerning the interface of prosody and information structure has already been pointed out. Until today, there is fundamental disagreement about the grammaticality of accent placement (as there is disagreement about the grammatical status of information structure). Those who deny the grammatical status of accent and intonation in general believe that its linguistic use can be explained in terms of more general natural principles like interest and arousal, being intimately related to gesture sharing with it the manifestation of “the opposition of up and down” (Bolinger 1986: 221). This view of intonation was most prominently represented by Dwight Bolinger who may be considered the most important proponent of the ‘highlighting’ approach. As already mentioned, the modern structure-based approaches have developed out of the Bolinger’s and Schmerling’s (1976) critique of an account that derives sentence accents by rule from syntactic constituents and particularly the concept of normal stress which is derived by Chomsky and Halle’s *Nuclear Stress Rule*. While the early structural accounts of normal stress were based on syntactic structure without recourse to semantic and pragmatic meaning, later accounts of focus accentuation (Schmerling 1976; Ladd 1980; Gussenhoven 1983), resorted to a semantically based definition of focus domains for the mechanistic accentuation rules to apply. The new structure-based accounts incorporated Bolinger’s insights into a strictly linguistic theory of accentuation, trying to reconcile the hitherto incompatible views of the speaker’s free choice to highlight whatever he considers appropriate to put into the foreground and the notion of ‘normal stress’ or neutral accentuation. In the remainder of this subsection, I will only review those accounts that are more directly relevant to the position taken here. I therefore will not deal with the various strictly formal syntactic accounts. For an overview confer Winkler (1997).
3.1.4.1 Bolinger: “Accent is predictable (if you’re a mind reader)”

In his famous rebuttal of the Chomsky-Halle (1968) Nuclear Stress Rule and its modifications (Bresnan 1971, 1972; Lakoff 1972; Berman & Szamoszi 1972), Bolinger argues against the view that the position of the strongest accent in a sentence is a matter of structure (Bolinger 1972b). He cites a number of counterexamples to the NSR and shows that it is in fact predictability and relative semantic weight of the lexical item and not syntactic structure that is responsible for the accent position. Bolinger cites a number of sentence pairs that could have the main accent on either the noun or the verb, depending on the speaker’s decision which of them to emphasize (cf. the examples in (10)):

(10)

a) I have a point to emphasize.
b) I have a point to make.
c) I can’t finish in an hour – there are simply too many topics to elucidate.
d) I can’t finish in an hour – there are simply too many topics to cover.

Bolinger is careful to state that the actual occurrence of the accent as suggested in the examples cannot be predicted with any certainty, but that the speaker will probably choose the semantically richer verb to be accented, because he “has already made up his mind that the operation rather than the thing is the point of information focus” (Bolinger 1972b: 634) when making the lexical choice for “elucidate” instead of “cover”. Bolinger further argues that “[t]he error of attributing to syntax what belongs to semantics comes from concentrating on the commonplace.” (ibid.) He points out that in many phrases with noun accentuation the verb is highly predictable, as is the case in the above collocation of ‘topic’ and ‘cover’ or ‘point’ and ‘make’. Bolinger has often been criticized for his explanation of the strongest prominence in terms of predictability (Schmerling 1976; Gussenhoven 1983). But in fact, Bolinger seems to regard lack of accent because of predictability only as a special case of ‘backgrounding’ of information that is of less interest. He does not hold that the verb has to be fully predictable from the noun in such cases. The notion of ‘relativity’ is crucial here. It is not the weight of the lexical item itself, but its weight in the context of the utterance. According to Bolinger, this also explains the accents on so-called ‘empty’ nouns like ‘thing’ or ‘place’ with concomitant deaccenting of the semantically richer verbs (p. 637) as in:

(11) I have so many things to take care of.

Bolinger distinguishes between two types of functions the accents fulfil: informing and impressing (1986: 74). Of course, the two functions are not completely separable, but, as Bolinger notes, there are accents that, whether informative or not, that affect the utterance as a whole which he calls ‘accents of power’ (Bolinger 1986: 74-88). This phenomenon includes climactic accents (an accent shift to the
“exclamatory early accents”, i.e. accents for the emotional highlighting of highly charged words such as ‘terribly’, ‘so’, or ‘absolutely’ that may carry the strongest accent of an utterance in an unexpected, early position, and ‘cumulative accents’ which are also often called ‘secondary accents’ in the literature. These secondary accents may not only go on function words, but they may also occur within a content word that already carries an accent to the left of the latter. In sum, these accents give the whole utterance a stronger impact, especially when they stand very close to each other, they bestow the utterance a chanting rhythm. Bolinger gives the example of Á compléte fáilúre. or Pút thát dówn. (Bolinger 1986: 85). The informing type of accent is called ‘accent of interest’ (Bolinger 1986: ch.7).

A given word, say, is accented to show its informativeness, which in turn usually reflects its newness, its unexpectedness, its special contribution toward answering the question that usually lies back of the spoken sentence. (Bolinger 1986: 89)

The questions Bolinger has in mind are, of course, wh-question that are typically used to ask for FOCUS constituents. For Bolinger intonation is as an independent largely iconic system, closely related to gesture, and motivated by the emotions of the speaker rather than as a grammaticized linguistic component (Bolinger 1986 et passim). Although he does not deny that intonation may serve to distinguish grammatical categories or encode certain linguistic functions, he emphasizes that it is “not designed to do so. What it is designed to do is distinguish between something ‘not new’ [...] and something ‘new’, or newly asserted [...]” (Bolinger 1983: 113). What is new and not new is signaled by the presence or absence of accents. In accordance with the strictly iconic relationship between accent and information he assumes. He is also not interested in relating these information points to a larger structure, whether syntactic or pragmatic, but rather regards interpretations concerning the meaning of the whole utterance as pragmatic inferences.

3.1.4.2 Halliday: Information structure is realized by intonation

Similar to Bolinger, Halliday conceptualizes focus also in terms of single information points rather than in terms of a structural domain of some size. For him, information focus “is one kind of emphasis” and “reflects the speaker's decision as to where the main burden of the message lies” (Halliday 1967: 204). In Halliday’s model, every information unit has one point of information, the information focus, that corresponds to an accent (the tonic in Halliday’s terms) which may be simple or compound. An information unit usually only contains one accent, the tonic. Halliday only identifies one type of information that contains a primary accent followed by a secondary one - this is the compound unit mentioned above. While conceding that an information unit in the unmarked case extends over a whole clause (p. 201), he emphasizes the autonomous nature of information structure. Information units are not defined syntactically but represent a distinct dimension of structural organisation, which
is then mapped on to syntactic constituent structure (p. 200). Contrary to Bolinger, Halliday thus acknowledges some kind of structure. As mentioned above, he also alludes to something like broad focus in observing that focus may assign the function ‘new’ to the whole domain it is part of (p. 207), i.e. in (8) the accent on *shed* may signal that the whole sentence is ‘new’ and uttered out-of-the-blue, for example answering the question “What happened?” But Halliday’s definition of information unit and information focus is to some extent circular. Although he assumes the primacy of the information unit as a unit of information structure, he does not specify any rules for its formation. It becomes obvious that his views on information and focus do not differ grossly from Bolinger’s when he states that the speaker has the freedom to decide into how many information units, i.e. quanta of information, he wishes to divide his utterance and which constituents they are mapped on in syntactic structure. In Halliday’s system the information unit is a phonological domain which in the unmarked case is realized by *tonality*, i.e. the placement of “the tonic [nucleus], on the last accented syllable of the item under focus” (p. 207). But as the information unit is never defined functionally, but rather phonologically being co-extensive with the tone-unit, it rather seems to be dependent on the presence of an information focus, i.e. an accent, whose exact location is not specified. This becomes much clearer in Halliday’s later work (cf. Halliday & Matthiessen 2004: 14-17, 91f.) than in his “Notes on transitivity and theme” (1967/68). In the unmarked case, the tonic falls on the last lexical item, but in the marked case it may fall on any other location (Halliday 1967: 202; Halliday & Matthiessen 2004: 91f.). The information unit then can be thought of as projecting from the point of information on to a certain domain which is limited by rhythmic consideration, and only vaguely matched with syntactic or pragmatic categories. Halliday’s dictum that information structure, in his system the highlighting of new or newsworthy information is realized phonologically is of course only meant by him to hold for English.

3.1.4.3 Chafe: Consciousness and information flow

The approach taken by Wallace Chafe (Chafe 1974, Chafe 1976, Chafe 1987, Chafe 1994) can justifiably be called strictly ‘cognitive’. He approaches the prosody of information from a perceptual and functional perspective. The central notion in Chafe’s model is that of consciousness. The basic assumption about how certain ideas⁴ are represented in the consciousness of a speaker and a hearer is responsible for the way it is encoded in the prosodic pattern.

Three main concepts are relevant for the issue of accentuation in Chafe’s approach: *activation* and *activation cost*, contrastive focus and the one idea constraint and its relation to the intonation unit. In

⁴ Note that for Chafe idea may denote a referent just as well as an event or a state, i.e. it may relate to all types of content words, nouns, adjectives and verbs.
his later work (Chafe 1987, 1994), Chafe differentiates between different types of what he previously referred to as ‘given/new’ (Chafe 1976). He identifies three different activation states concerning the representation of certain ideas in the interlocutor’s minds: active, semi-active and inactive (e.g. Chafe 1994: 53-56). The main point is that if an idea is already active at a certain point of a conversation it can be regarded as given information. Conversely, what is characterized as ‘new’ information is something ‘newly activated’ at this particular time. The third cognitive state, something being ‘semi-active’ in the consciousness of a speaker/hearer, renders this idea accessible at the particular point in the conversation. Chafe assumes that these different activation states correspond to different activation costs. Accessible and new information are more ‘costly’ in Chafe’s terms as far as the cognitive effort is concerned. This higher cognitive effort is reflected phonologically as stronger accent, while given information which does not need to be activated therefore also requires less cognitive effort and is rendered by ‘attenuated pronunciation’. Chafe assumes that accessible and new information is more or less expressed in the same way (1994: 75). In that sense, Chafe’s original statement about the linguistic expression of activation cost is still valid, albeit with language-specific differences. Chapter 4.3.1 will be devoted to this issue in EA.

The principal linguistic effects of the given-new distinction, in English and perhaps all languages, reduce to the fact that given information is conveyed in a weaker and more attenuated manner than new information. This attenuation is likely to be reflected in two principal ways: given information is pronounced with lower pitch and weaker stress than new, and it is subject to pronominalization. (Chafe 1976: 31; emphasis added)

The second concept that makes use of the term focus is contrastiveness. Contrary to Bolinger, Chafe argues that contrastiveness is indeed a linguistically relevant category, claiming that “contrastive sentences are qualitatively different from those that simply supply new information” (Chafe 1976: 34). The expression of contrastiveness is said to be “stronger prominence” (Chafe 1976: 36) or “primary accent” (Chafe 1994: 76). Importantly, this category of contrastiveness is independent of activation state and activation cost; it rather cross-cuts the given-new distinction. Given the fact that given items can be assigned a strong accent makes it impossible to infer the activation status of a referent, for example, from prosody alone. While in English an unaccented or weakly accented argument hints at the givenness of the referent, an accented argument does not necessarily imply its being new.

From what has been said so far, it becomes obvious that Chafe’s category of contrastiveness at least partially overlaps with what has been called focus by most other linguists reviewed in this section. Chafe himself uses the term ‘focus of contrast’ (Chafe 1976, 1994: 78), but in his 1994 book he points out that the term suffers from lack of precision. Interestingly, he also considers his notion of contrastive focus as being essentially different from information focus and leaves the question open, avoiding the use of the term focus in the first place (Chafe 1994: 76). We will see in Chapter 4 that the most recent accounts (Féry & Samek-Lodovici 2006; Katz & Selkirk 2011) take a very similar position.
The third concept mentioned at the beginning of this section relates to another aspect of how information flow is expressed prosodically. Not unlike Bolinger, but contrary to the other authors reviewed in this section, Chafe does not refer to any kind of (pragmatic) domains or any kind of structure when talking about the relationship between information and accentuation. He therefore also does not put up any rules for accent placement. The domains which he calls *intonation units* are strictly phonological, but as Chafe (1994: 57) notes, phrasing “operates in happy synchrony with basic functional segmentations of discourse”. Although Chafe uses the term *primary accent* (Chafe 1994), he does not follow the British intonationist tradition in arbitrarily limiting the intonation unit to just one primary accent, the nucleus. Nevertheless he notes that the amount of new ideas that is normally expressed in one intonation unit is heavily restricted. In fact, in his data, which is natural, totally spontaneous conversational speech, this amount is usually restricted to one idea per intonation unit. Due to this *one-idea constraint* (Chafe 1994: ch.9), an intonation unit will also more often than not only have one primary accent. It is clear that the one-idea constraint will not hold for read speech, for example; but if I understand it correctly, this is not what Chafe wishes to claim. If we think of constraints as something that may be violated if other considerations are overriding it, this constraint seems to give a very good empirical fit for spoken language.

### 3.1.4.4 Schmerling: Topic-comment structure

Schmerling (1976) represents an important step further in the direction of what also is the aim of the present study, which is relating prosodic structure to informational categories. Like Bolinger, she heavily criticizes the normal stress view and the concept of ‘neutral accentuation’ it is based on. Schmerling ((1976: 39) identifies two major tenets that form the foundation of Bolinger’s views. The first one is the lack of predictability of accent location and the dependence of the latter on the context, and the speaker’s decision which parts of the utterance to put in the foreground. The second one is the assumption that this highlighting always reflects relative semantic weight and informativeness. Schmerling abides by the first of these tenets, but rejects the second one. She presents two ever since often cited real-life examples of structurally identical sentences (12, 13) that exhibit different accentuations and argues that Bolinger’s ‘relative semantic weight’ falls short of explaining the difference. She reports that one morning when coming down the stairs of her parents’ home, her mother told her the latest news which was that president Truman had died (12). At that time, the American president Truman had been hospitalized for quite a while and reports of his health condition
had been on the media all the time. A few weeks later, her husband told her another piece of news, namely that the former president Johnson had died (13):

(12) Trůman díed.
(13) Jóhnson díed.

Schmerling notes that Johnson’s health had not been under discussion and thus this news came rather unexpectedly (Schmerling 1976: 41ff.) She states that Bolinger’s theory makes the wrong predictions because the unexpected and newsworthy item is the one that is accented in Truman’s case while Johnson’s death was, according to Schmerling, a much less expected thing to happen and was therefore more apt to be accented if Bolinger were right. Schmerling concludes that this difference must be due to something else than informativeness or relative semantic weight and suggests that predicate accentuation reflects the topic-comment structure of the sentence (p. 93). This observation serves as the basis for her Principle IV (p. 94).

(14) Schmerling’s Principle IV:
In a topic-comment utterance, stress both the topic and the comment.

Schmerling’s account thus for the first time identifies some kind of pragmatically defined structure that serves as a basis for the observed occurrence of accents in the sentences. Her Principle IV does not, however, specify the location of the main stress or accent. This is provided for by her Principle III (p. 86):

(15) Schmerling’s Principle III:
Given a sequence of stresses which are equal and greater than other stresses within the intonational unit, the last such stress will be more prominent than the others.

What is crucial in that respect is the fact that Schmerling draws a sharp distinction between phonological and pragmatic principles. Schmerling’s third principle in fact is the equivalent of the Nuclear Stress Rule. Like the NSR it refers to the surface structure of the sentence and only applies after other principles have been applied.

The other two principles Schmerling presents in her seminal thesis have also been fundamental in the subsequent theoretical developments by Ladd (1980) and Gussenhoven (1983) and other work based on these early approaches. Principle I (p. 75) takes care of what has later on most often been referred to as the ‘deaccenting of givenness’ (Ladd 1980, 1996; Cruttenden 2006). But Schmerling uses different terms. She interchangeably resorts to descriptions like “taken for granted” and “insignificant” (p. 75). In her view it is not the mere fact that deaccented material is ‘old information’ or ‘given’, but rather that the speaker regards it as already given or taken for granted at the moment of the utterance.

---

5 I have preserved Schmerling’s original notation of accent degrees here, a fact that has been ignored in most quotations of these examples in the literature but that is not irrelevant.

126
Schmerling’s Principle I:
Certain items in an utterance are treated by the speaker as relatively “insignificant” and fail to be assigned stress.

Schmerling (1976: 78ff.) presents some evidence to show that “taken for granted” is not the same as presupposed, she rather assumes that this distinction “cross-cuts the distinction between what is and what is not presupposed.” (p. 78) Thus Schmerling’s view on that issue is again closer to Bolinger’s ideas of non-predictability and speaker’s choice than to the structurally oriented accounts like, for example, Ladd (1980) which will be presented in the next section. These truly structure-based approaches resort to ‘mechanistic rules’ that have a stronger sense of compulsoriness compared to ‘principles’ which allow a more flexible implementation. Consequently, it is not possible to predict the deaccenting of certain items in Schmerling’s approach from the contextual facts, neither from textual nor situational givenness, but it is rather the individual decision of the speaker that determines it, which in turn cannot be accounted for by structural rules. In Chapter 4.3, I will argue for a distinction between between two types of pragmatic properties: givenness in the sense of familiarity (Prince 1981) or accessibility (Chafe 1974) on the one hand and background as the opposite of focus, which is in fact Schmerling’s “taken for granted”.

The last of Schmerling’s principles (Schmerling 1976: 82) formulates a generalization on the grounds of the observed facts of accentuation in English concerning the accentability of arguments and predicates which turns up in one way or another in all subsequent accounts of sentence accentuation.

Schmerling’s Principle II
The verb receives lower stress than the subject and the direct object, if there is one; in other words, predicates receive lower stress than their arguments, irrespective of their linear position in surface structure.

This generalization, even if it holds for English, German, and Dutch, is not explanative. In her view, the accent contrast in the minimal pair cited under (12) and (13) is “rule-governed”. While Schmerling admits that she is “not able at this time to propose a specific treatment of these cases in a grammar, she goes on to say that the recognition of the structural properties of these sentences “is the requisite preliminary step toward doing this” (p. 90). Schmerling’s approach was discussed at some length here because of its relevance to the development of the later structure-based approaches and to my own ideas to be laid out in Section 3.3 below. What is of special interest to the present approach is Schmerling’s theoretical distinction of “pragmatically based principles referring to discourse function, and phonological principles related to phrasing and tempo” (Schmerling 1976:81). The present approach follows Schmerling in this distinction.
Building on Bolinger’s and Schmerling’s critique of the syntactic approaches to sentence accent, Ladd (1980) offers what may be regarded as the first structure-based account that incorporates the insights of the Bolinger-Schmerling semantic theory of sentence accentuation. Ladd accepts their arguments concerning the semantic basis of accentuation and Schmerling’s criticism of the notion of ‘normal stress’ as the accent pattern of contextless utterances. But he refuses Bolinger’s concept of focus as ‘point’ of information which implies a one-to-one correspondence between focus and accent. Although agreeing with Schmerling that the Chomsky-Halle-version of normal stress is, particularly the notion of ‘contextlessness’ is untenable, he disagrees with her conclusion to give up the whole idea of a neutral or ‘normal’ accentuation pattern (Ladd 1980: 74, 1996: ch. 5.1.1). Ladd believes that the solution lies in the the integration of the concept of focus as a structural basis for accentuation rules (Ladd 1980:74). As already noted above, he points out that most of the accent locations possible in English sentences have been hitherto dismissed as cases of emphasis or ‘contrastive’ stress and therefore no attempts have been made to derive them by rule, but that one accent placement exists “that leaves the focus broad or unspecified” (ibid.) which he thinks to be equivalent to the ‘normal stress’ pattern (p. 75). Instead of speaking of contrastive and normal stress, Ladd thus suggests that we are rather facing cases of ‘marked’ and ‘unmarked’ accentuation or ‘narrow’ and ‘broad’ focus, respectively. He thus supplants the old syntactic basis of accent placement with the ‘semantic’ concept of focus and by this lays the foundation for the development of a theory of ‘accent as a marker of focus structure’. His account is also able to give a unified account of accent placement, whether under narrow or broad focus conditions:

Instead of saying that normal stress goes on the rightmost accentable item in the sentence, we will say that accent – in general – goes on the rightmost accentable item of the focus constituent. If the focus constituent is the whole sentence, we get ‘normal stress’; if not, we get a narrow focus on the constituent identified by the placement of the accent. (Ladd 1980: 77; emphasis in the orig.)

It has to be noted, though, that Ladd (1980) does not provide an account of what a focus constituent is. He also accepts Schmerling’s basic distinction between nouns and verbs and incorporates it into his “hierarchy of accentability” (Ladd 1980: 85) according to which content words are more readily accentable than function words and nouns are more readily accentable than verbs. He then offers a revised version of his focus rule (cf. the quotation above) that replaces “the rightmost accentable item” with “the most accentable” item. Ladd thus reduces the mechanical character of his preliminary rule and rather leaves accent placement to the interaction of position and accentability, where accentability is not regarded as a dichotomy, but as a gradient and a relational concept that is to be seen “relative within the focus constituent” (Ladd 1980: 86).
Related to the notion of accentability is the notion of ‘deaccenting’ as conceived by Ladd. He views deaccenting as “a lowering of the degree of accentability of an item or constituent” (p. 87). Ladd is thus able to derive marked accent locations where the accent does not fall on the final accentable item or syllable of the domain because in cases where the focus nevertheless has to be regarded as broad. A case in point is the following examples of deaccented final nouns (Ladd 1980:87; 1996:175):

(18) A: Has John read Slaughterhouse-Five?
   B: No, John doesn’t read books

(19) A: I found an article for you in a German journal.
   B: I don’t read German.

Ladd (1980) argues that the location of the accent is determined by ‘default’. The last lexical items ‘books’ and ‘German’ that should be accented due to the Focus Rule are in fact deaccented because of their givenness, i.e. their predictability from discourse. The important thing about this assumption is that, contrary to the ‘highlighting’ view, accent placement is viewed here as dependent on structure. The accent on ‘read’ is not regarded as reflecting the importance of ‘reading’ in that special case, as Bolinger would have it, but as a case of ‘default’; ‘read’ only receives the accent because it is the accentable item contiguous to the one that would have been accented by rule. These assumptions enable Ladd to reconcile the two opposite views in the literature at that time which he believes to have each “a piece of the truth”. Crucially, Ladd does not do away with the notion of ‘normal stress’ like Schmerling did. He says that every native speaker has clear intuitions about normal stress and that the notion should not be dismissed all that easily. He, however, ties together normal stress and the semantically based notion of focus or focus domain and thus is able to derive marked accent patterns that have hitherto been dealt with under the heading of contrastive stress and unmarked accent patterns by the same mechanism. This is possible as Ladd’s deaccenting does not only shift the accent to the left, but also to the right which also successfully derives marked compound accents (Ladd 1980: 88) or a shift in lexical stress as in Bolinger’s classic example This whisky wasn’t ExPorted, it was DePorted (Bolinger 1961: 83; Ladd 1996: 177f.). Deaccenting in English may apply for reasons of contextual ‘givenness’ or the semantic ‘emptiness’ of the item under consideration. Thus indefinite pronouns such as something, anything etc. are usually deaccented in English, just as is the category of semantically ‘empty’ content words, e.g. person, thing, stuff etc. (Ladd 1980: 92ff.; 1996: 179ff.)

Ladd’s (1996: ch. 5.2) arguments for the structure-based approach show that while accent placement in narrow focus cases can readily be accounted for by a universalist model, it is broad focus in which the whole utterance or a longer sequence is in focus, that, according to him, makes an indirect – structurally determined – relation between accent and focus necessary. Moreover, as has already been

---

6 In his book Intonational Phonology (1996), Ladd changes that term due to the different meaning this term has assumed in the course of time (p.293, n.2).
pointed out above, languages differ in their susceptibility to deaccenting and thus also as far as nucleus placement is concerned – a fact that, according to Ladd, does not speak in favour of a universalist approach. But what if we do not need a focus accent to mark the focus (i.e. FOCUS in our terms) at all?

One very important aspect of Ladd’s approach is the coupling of pragmatic facts of accentuation with the then emerging metrical stress theory (Liberman & Prince 1977). In Ladd’s account, deaccenting of an item is viewed “as a relative weakening of its hierarchical rhythmic position” (Ladd 1980: 56; emphasis in the orig.). This position, of course, is tied to the idea that sentence stress is a prominence-related or rhythmic phenomenon rather than a tonal one, conveyed by (pitch) accents. Ladd uses the tree notation to represent the rhythmic structure of a sentence which implies that this structure builds on syntactic constituent structure. His views are thus also compatible with the line of research that develops the idea of focus as a syntactic. The following example taken from Ladd (2008: 270) illustrates the different accentuation with ‘normal stress’ (20a) and ‘default accent’ (20b) as in example (19) above as a “reversal of relative strength in a metrical tree” (emphasis in the orig.).

(20) Ladd’s ex. (18)

The assumption of the relational nature of accents or rather stress positions has two major theoretical consequences. The first one is that the whole focus-to-accent issue is referred to metrical structure which relieves the melodic component, i.e. intonation proper, from the functional load of focus marking. Strictly speaking, for Ladd accent is not itself the marker of focus, but rather relative prominence. Only, this prominent rhythmic position is in most cases also ‘assigned’ a pitch accent. By the conceptual distinction between accent and prominence on the one hand and their relation to focus on the other hand, Ladd, like Schmerling before him, recognizes focus as a pragmatic phenomenon distinct from its phonological realization and acknowledging accents on constituents that are not in focus (cf. Ladd 1996: 223ff.). The second theoretical position implied by Ladd, which is not adopted in the present study, is the primacy of rhythmic over intonational structure and the assumption that it is only the metrical position determines the place of pitch accents. The details of this view have been laid out in Chapter 2.

Ladd’s (1980) theoretical work can be regarded as a further step in the investigation of the interface between information structure or rather only focus and intonation carried out from a prosodic perspective. But the focus category it assumes is not well defined. Such a definition is given by
Gussenhoven in his very influential paper on accentuation in English and Dutch (Gussenhoven 1983). In his review of the article, Ladd (1983) basically adopted Gussenhoven’s approach which will be the discussed in the next section.

3.1.4.6 Gussenhoven: Focus domains over semantic constituents and rule-governed accent placement

Like Ladd, Gussenhoven (1983: 378f.) goes a long way defending the idea that accent is predictable from linguistic structure. He criticizes the Bolinger-Schmerling view for the implicit demand that a linguistic theory should offer predictions as to the choice a speaker will make under given circumstances and says that this is “tantamount to wanting to predict what people are going to say” or in the same way whether a speaker is going to opt for a passive construction or not (p. 379). He admits that a theory aiming for that will only be able to make probabilistic predictions. What to his mind is not probabilistic, but can be predicted with great accuracy, is the position of the accents once the speaker has made his choice from “a set of linguistic primes” (ibid.). On this view, the speaker has a considerable freedom in choosing what he wishes to put in focus; but once this decision has been made, the rest follows automatically, in the sense that accentuation is totally determined by rule from the chosen linguistic structure. As Ladd put it, Gussenhoven conceives the speaker to make his choice “at a more abstract level, in the assignment of a semantic feature [+ focus] to semantic constituents” (Ladd 1983a:158). This also means that for Gussenhoven focus is really a semantic phenomenon belonging to the realm of linguistics. Identifying focus as a formal linguistic category is in fact drawing a dividing line between the linguistic system and pragmatics (Gussenhoven 1983: 379).

As regards the definition of focus, Gussenhoven follows work in the generative tradition (Chomsky 1970; Jackendoff 1972) defining it formally as a “binary variable which obligatorily marks all or part of a sentence as [+focus]”, leaving it “semantically undefined” (Gussenhoven 1983: 380). Nevertheless, Gussenhoven comments on the ‘meaning’ of focus and characterizes it as “the speaker’s declared contribution to the conversation, while [–focus] constitutes his cognitive starting point. Thus, his [±focus] feature does relate to discourse, e.g. the specific information status of the individual referents. However, as Gussenhoven rightly observes, these contextual conditions do not fully determine the focus structure of the utterance. Gussenhoven clearly distinguishes between the formal marking of the linguistic category of focus and the reasons why a speaker chooses any given option (ibid.).

Following Schmerling (1976) in the distinction between arguments and predicates in terms of their accentability, Gussenhoven makes this distinction the basis of his focus domain formation rules over which his sentence accent assignment rule (the SAAR) operates (Gussenhoven 1983: 391). A focus
*domain* as defined by Gussenhoven consists of “one or more constituents whose [+focus] status can be signalled by a single accent” (Gussenhoven 1983: 391). The semantic constituents identified for the purpose of domain formation are arguments (A), predicates (P), and modifiers (M)⁷, which last category basically contains adverbials. Two basic assumptions are made: the first one is that items that are considered [−focus] are generally unaccented, and the second one integrates the empirically observable fact of English that nouns are more readily accented than verbs. Gussenhoven’s domain formation rules account for these observations in stipulating that, if marked as [+focus], a predicate will merge into one focus domain with an adjacent argument, while arguments do not merge, but project a focus domain of their own and so do modifiers, if they are focused. This leaves us with different pragmatic inferences to be made from unaccented arguments and unaccented predicates. While the referent of an unaccented Argument is necessarily to be considered as ‘given,’ or [−focus] in Gussenhoven’s terms, an unaccented predicate does not tell us anything about its information status.

The preferred focus domains are thus [AP] or [PA], upon which the SAAR operates putting an accent on the Argument, yielding [A*P]⁸ and [PA*]. A typical all focus SVO sentence thus usually consists of two focus domains, namely [A*] [PA*]. But to be able to deal with sentences that have an accent on every lexical item, Gussenhoven’s domain formation rule allow for the option that a semantic constituent may form its own focus domain.

Gussenhoven’s Sentence Accent Assignment Rule (SAAR)

Domain assignment: 
\[ P(X)A \rightarrow [P(X)A] \]
\[ A(X) P \rightarrow [A(X)P] \]
\[ Y \rightarrow [Y] \]

Accent assignment: [ ] → [*]. In AP/PA accent A.

(A = Argument, P = predicate, X and Y stand for any of these, or for a modifier; underlining indicates focus and * stands for accent.)

Expressed in prose, the SAAR could be characterized as assigning an accent to every predicate, argument, and modifier, if focused, except predicates that are adjacent to their arguments (Gussenhoven 1992: 84).

Gussenhoven’s account has been taken to task by Lambrecht (1994: 317) for its difficulties in accounting for double accent sentences with accents on the argument and the predicate as shown in example (16) cited from Faber (1987):

(21)

(a) TRESPASSERS will be PROSECUTED.⁹
(b) JOHN’s WORKING.
(c) JOHN PROTESTED.

---

⁷ The original term in Gussenhoven (1983) was ‘Condition’. The term was changed in Gussenhoven (1992).
⁸ The asterisk is used to signal an accent in Gussenhoven (1983).
⁹ I keep with Lambrecht’s use of capital letters for accented words here.
Lambrecht notes that while Gussenhoven’s rules are able to account formally for the accent patterns, they do not explain why the option of two focus domains (the Y option) is chosen and how a single argument constituent can constitute a focus domain. Lambrecht’s position is that a referent can only be in focus if it is construed as a predicate (p. 318).

While Lambrecht’s critique is not fully warranted, at least as far as (17a) and (17b) are concerned, it is illuminating in one important respect. It shows that Gussenhoven’s focus domain and Lambrecht’s focus domain are in fact not the same thing. We will return to that main issue at the end of this Section, where I will argue that different approaches in the literature explicitly or implicitly rest on different definitions of focus and that the solution lies in assuming two a priori independent concepts. That his focus is clearly different from Lambrecht’s identically named concept becomes obvious when we have a closer look at what Lambrecht (and many others) would classify as an accented topic, which in Gussenhoven’s account is given various explanations, but is always analysed as [+focus]. Gussenhoven, who explicitly rejects Schmerling’s assumption of an accent-bearing topic, is thus forced to put different syntactic and semantic constraints on the wellformedness of sentences with an accented subject and an accented predicate. An explanation in terms of a topic-comment articulation à la Schmerling would, however, be sufficient to account for all cases. In my opinion, Gussenhoven’s rejection of Schmerling’s topic-comment category is exactly the crux with his account. Distinguishing between preposed elements that are topics and those that are foci (i.e. FOCI) makes it possible to explain why only (18a) is felicitous in a context like “Who do you like?” while both (18b) and (18c) could be an answer to “What about John?”

(22)
(a) JOHN I like.
(b) JOHN I LIKE.
(c) John I LIKE.

The SAAR is a language-specific rule which basically holds for other West-Germanic languages, namely Dutch and German, as well but clearly fails in accounting for the accentuation facts of many other languages, such as languages Romance phylum or many varieties of Arabic that do not make the same distinction between predicates and arguments. It thereby treats the principles of accentuation for these languages as fundamentally different from the principles valid for other intonation languages. Contrary to Ladd (1980), who also rejects an approach in terms of universal principles, opting for structure-based and language-specific accounts of the intonational systems of individual languages, Gussenhoven does away with metrical structure and the notion of normal stress and any variety of the old NSR which are both replaced by the provisions of the SAAR.

---

10 Without going into detail here, we may note that Gussenhoven puts an important semantic condition on the formation of [A*P], namely that the sentence should have an ‘eventive’ reading, which cannot be assumed for (17a) and (17b).
Working within the framework of generative grammar, specifically Chomsky’s *Extended Standard Theory* (Chomsky, 1981), Selkirk (1984) offers a book-length treatment of the syntax-phonology interface. Central to this approach is the assumption that the relation between the surface syntactic and the surface phonological representations of a sentence involves a set of mapping rules between the two components. Even though these assumptions seem to differ from the basic theoretical assumptions Gussenhoven makes, who stresses that his account is based on semantic, not on syntactic structure, both theories cover essentially the same empirical ground and, following Schmerling’s original suggestion, make appeal to the *argument structure* of a sentence for *focus projection* (Selkirk 1995: 555). The major difference could be characterized by the different derivational directions their models take. While Gussenhoven’s model may be termed a top-down model regarding the syntactic derivation that identifies a focus domain assigning the focus feature to a higher syntactic node first and then looking down the syntactic tree to identify the *focus exponent* and assign it an accent, Selkirk’s model goes the other way, freely assigning a focus to a lexical item and *projecting* the focus domain from the innermost focus constituent within syntactic constraints in a bottom-up manner.

This goal is achieved by two basic rules which I cite from the more recent refined version (Selkirk 1995: 555) of the theory. The first rule assigns the basic *focus* to any individual word of a sentence.

(23) Selkirk’s Basic Focus Rule  
An accented word is F-marked

The second rule then projects the basic focus to a higher constituent referring to the predicate-argument structure and identifying the syntactic wellformedness conditions that have to be obeyed.

(24) Selkirk’s Focus Projection  
(a) F-marking on the head of a phrase licenses the F-marking of the phrase.  
(b) F-marking of an internal argument argument of the head licenses the F-marking of the head.

Contrary to all the other accounts, Selkirk’s model thus assumes hierarchical focus domains and pitch accent assignment that is in principle as free as in Bolinger’s approach. Even though she distinguishes between focus as a syntactic feature and pitch accent, her use of the focus concept is a hybrid one, referring to a syntactically defined domain with a pragmatic function on the one hand, and to the highlighting of particular elements by pitch accents on the other hand.

Principle (b) of the focus projection rule accounts for the frequently observed fact that in English objects which mostly occupy a position closer to or at the right edge of a sentence are accented while verbs go without an accent. In the Gussenhoven model this is provided for by the accent assignment rule of the SAAR. As Gussenhoven makes special provision for the formation of AP-domains in the
case of event sentences his rules also account for short subject accented sentences with mostly intransitive or unaccusative unaccented verbs, such as Johnson died. In the early version, Selkirk’s model did not account for these cases. In the 1995 version, however, Selkirk assumes that, if the subject and the verb are new, the verb is generated within the verb phrase. When the subject is moved the predicate remains in focus (F-marked in Selkirk’s terms) on the provision that “the F-marking of a constituent licenses the F-marking of its trace” (Selkirk 1995: 559). Even though she denies the notion of ‘normal stress’ just like Schmerling or Gussenhoven, Selkirk preserves the NSR -that assigns a phrase accent to a syntactic constituent - as a “default principle, applicable only when ‘all else is equal’ (1995: 563).

We claim that phrase stress is irrelevant to meaning or pragmatics, that it is simply the presence or absence of a pitch accent on a word that is taken into account in interpreting a sentence for semantic or pragmatic purposes, and thus that principles like the NSR are, in a sense to be made precise, "exceptionless" and truly "phonological". (Selkirk 1984:145, emphasis added)

Unlike Gussenhoven (1983), Selkirk makes use of metrical structure. The interesting aspect of the approach is the assumption that metrical rules are purely phonological and distinct from focus accents. Selkirk’s model exploits the option of the metrical grid as suggested by (Liberman & Prince 1977) and especially (Prince 1983). The metrical grid is built on the syllabic composition of its constituent structure. Phrasal prominence emerges from the interaction of pitch accent as a choice made by the grammar and grid euphony rules with the former determining the range of possible rhythmic patterns of an utterance (Selkirk 1984: 144). As the Principle of Rhythmic Alternation which is responsible for the grid euphony rules (cf. Chapter 2.1.2) could not account for some empirical cases of accent retractions and protractions discussed by Bolinger (1981, 1986), she introduces the Phrase Edge Prominence rule that takes care of the derivation of edge demarcating accents (Selkirk 1995: 565), but rejects Bolinger’s explanation of an optimal pitch contour for this phenomenon. Whatever the phonetic factors may be that determine the occurrence of edge prominences, phonologically speaking they are a manifestation of the demarcative function of accent. As already mentioned in Chapter 2, the approach taken in this work also makes a principled distinction between purely prosodic processes and meaningful accent locations that might even contradict the natural rhythmic patterns. The main difference to Selkirk’s account is the rejection of the NSR as a phonological end rule pertaining to syntactic constituents. I rather assume that there is a principle of Phrase Edge Prominence that optionally applies to prosodic constituents in a strictly phonological manner without any reference to syntax.

11 This assumption allows Selkirk to account for a range of other problematic cases of the focus-to-accent issue. Of course, in a functional framework, such ‘explanations’ are unsatisfying, but somehow they only serve to express in a formal way what Gussenhoven’s PA-formation does for eventive sentences, namely account for the fact that the subject and verb are integrated in one domain (Fuchs 1980; Jacobs 1993, 1999), expressing ‘one idea’ (Chafe 1994) or a non-bipartite structure (Sasse (1987). Theticity will eventually be dealt with in Chapter 6.
Selkirk, however, also acknowledges that intonational phrasing does not necessarily stand in one-to-one correspondence with syntactic phrasing. Intonational phrases are only thought to be restricted by semantic considerations, obeying the requirement that one intonational phrase must correspond to a “sense unit” (Selkirk 1984: 290). However later on she concedes that such a requirement is “difficult to implement” (1995: 567), observing that “the principles governing intonational phrasing are not well understood” (ibid.). In her modular model an underlying phonological representation is derived from syntactic surface structure in two stages: (i) the mapping of the syntactic surface structure into the intonated surface structure including intonational phrases, the pitch contour and the individual focus or pitch accents and (ii) the construction of the metrical grid (i.e., the stress pattern) on the basis of that intonated surface structure. Selkirk (1995: 565) observes that the different principles and rules (in her example Phrase Edge Prominence and the NSR) could be viewed as constraints that could be violated in favour of the satisfaction of another constraint.

In sum, contrary to the stress-first approaches of (Liberman 1979) and Ladd (1980), Selkirk’s approach is closer to Bolinger’s views and to the position taken here, in that she does not derive accents of ‘interest’ or ‘information’ in a bottom up manner from the syllabic and higher prosodic structure of an utterance. Giving this strategy an interpretation within the theoretical framework of the present approach, we could say that she keeps the semantico-prosodic and pragma-prosodic aspects of accentuation distinct from the natural prosodic processes of rhythm and phrasing. And this is exactly the position taken in the present work.

3.1.4.8 Lambrecht: A unified functional account of sentence accentuation

Lambrecht (1994) devotes a chapter of his book to the prosodic coding of information structure. His account is not developed from an intonational perspective like most other accounts reviewed in this section. His functional model is explicitly approached from a pragmatic viewpoint. What distinguishes his approach from the ones reviewed so far is the fact that he provides a principled account for informational categories which determine the placement of sentence accents. That is, the different accents in Lambrecht’s account are all functionally motivated. Lambrecht’s approach can be counted as structure-based, only that the structure he refers to is not syntactic or semantic, but pragmatic in nature.

Lambrecht follows Ladd (1980: 213) in the belief that focus is signaled solely by accent location (1994: 239), thus distinguishing between focus and ‘emphasis’ of which he says that it is expressed by other intonational means such as wider pitch range and higher loudness. Another important difference between Lambrecht’s approach and the ones of the phonologists and Chafe, for instance, is the fact that it deals with all of ‘information structure’ and not only focus structure. Given his view on
information structure as part of the grammar of a language, Lambrecht consequently also regards the prosody of information structure as an issue of grammar and not of universal principles that simply highlight whatever is to be put in the foreground. In accordance with his contention that information structure is signaled by accent location alone, he thus lines up with the proponents of the structure-based approach. Although acknowledging the mainly iconic motivation of prosody (p. 242), he resorts to grammatical rules as well, saying that the relationship between highlighting by accent and the communicative function can “at best be partially iconic” (ibid., emphasis in the orig.), given the fact, as conceived by Lambrecht, that the sentence accent falls on a single syllable while the domain it signals as focused extends over a much larger stretch of the utterance. Citing the example of onomatopoetic expressions like cockadoodledoo that differ between languages without losing their sound-symbolic shape (cf. kikerikii in German, kukukuuku in EA and kukulukuu in Persian), Lambrecht states that in prosody too it is necessary to “walk the constrained way” through the grammar of a language to arrive at the language specific expressions of, in that case, informational categories. For him, “[s]entence accentuation may be iconic in its foundations, but it is filtered through the machinery of grammar.” (p. 243)

The clear-cut distinction Lambrecht draws between focus and highlighting allows him to relate accent also to other informational categories, such as topic. Lambrecht’s General Phrase Accent Principle, building on Halliday’s accentuation rule and Ladd’s (1980) Revised Focus Rule, the latter’s ‘focus domain’ with a more general ‘pragmatic entity’ (p. 247):

(25) Lambrecht’s General Phrase Accent Principle:
A phrasal accent marks the right boundary of a syntactic domain expressing a pragmatically construed portion of a proposition.

Lambrecht is careful to point out that ‘syntactic domain’ is not coterminous with syntactic constituent. The formulation “pragmatically construed portion of a proposition” is essential to refer to topical accents which in Lambrecht’s proposal are not simply ignored, but motivated in their own right. To justify accents on topical constituents, Lambrecht adopts Chafe’s concept of activation states discussed above and in Chapter 1. Generally following Chafe in the idea that fully active referents do not need to be activated by an accent, he claims that this lack of accent, called “attenuated pronunciation” by Chafe (1987: 26), is a ‘grammatical’ correlate of the cognitive category of ‘activeness’ (Lambrecht 1994: 95). An accent on a topic constituent may be one of activation. But as Lambrecht correctly notes, expressions referring to discourse referents may also be accented if the referent is active in the interlocutors’ consciousness. He states the case of accented pronouns. If a discourse referent is referred to with a pronoun, he doubtlessly has been activated before. Lambrecht concludes that there must be yet another function of the accent and suggests this function to be the establishment of the pragmatic role a discourse referent plays with respect to a proposition. Thus the sentence accent that marks the focus domain by rule and the topic accent that falls within the topic domain, both mark the linguistic expressions they are assigned to for their respective functions as
topical and focal discourse referents. The whole notion of ‘activation’ is thus redefined in Lambrecht’s account. It is no longer understood to simply “conjure up a representation of [a referent] in the mind of the addressee but to establish a relation between it and a proposition.” (Lambrecht 1994: 324).

In line with the tenets of construction grammar, Lambrecht (1994: 318) analyses the different patterns associated with different focus structures as separate formal categories which he calls prosodic constructions. The predicate focus structure is regarded as the unmarked pragmatic construction and consequently exhibits unmarked prosody, i.e. the sentence accent shows up at the right edge of the domain. But, as Lambrecht points out, a predicate focus may also have the sentence accent in any other position, provided it falls within the predicate. The sentence accent is shifted to the left if the material following it has already been activated. Lambrecht thus incorporates Ladd’s insights about ‘deaccenting’ and ‘default’ accentuation and argues that the focus domain may contain activated, non-focal elements which are by definition accentless in English.

A predicate focus construction, being the unmarked focus articulation, is vague and has more than one interpretation or readings which to be drawn from conversational implicatures and not derived by grammatical rules of focus construction (p. 296). Prosody alone, as Lambrecht notes, cannot differentiate between these different readings. To differentiate them formally, it is necessary to resort to morphosyntactic marking like word order, for example. All else being equal, the predicate focus construction will be interpreted as the unmarked pragmatic articulation of a topic-comment sentence. If, however, this interpretation is to be excluded, the predicate must be marked by absence of prominence. Just as predicate focus is the unmarked focus construction, sentence focus and argument focus are regarded as marked. In an interesting treatment of the longstanding problem of subject accented (thetic) sentences he suggests that these are prosodically encoded by a construction he terms prosodic inversion. Arguing that this prosodic construction functions on a par with the syntactic inversion construction, which is common to some Romance languages to mark sentence focus, he compares Italian with English in that respect, showing that while Italian uses word order, English resorts to prosody for the same end. They, in his view, both use unmarked word order/prosody to signal unmarked predicate focus, and, quite iconically, make use of an inversion construction to signal the marked type of sentence focus (p. 319):

(26) Lambrecht’s prosodic and syntactic inversion

<table>
<thead>
<tr>
<th></th>
<th>Predicate Focus</th>
<th>Sentence Focus</th>
</tr>
</thead>
<tbody>
<tr>
<td>Italian:</td>
<td>SP</td>
<td>PS</td>
</tr>
<tr>
<td>English:</td>
<td>SP</td>
<td>SP</td>
</tr>
</tbody>
</table>

(S = subject, P=predicate, bold face = accented)

Lambrecht’s proposal thus addresses an observation also made by Ladd (1996: 167) that there are “consistent cross-language differences in patterns of sentence accentuation”, i.e. there are languages that readily allow accent retraction from the ‘normal’ final position like English and others that do not
or not that readily allow this. It also concurs with Vallduví’s distinction between ‘plastic-accent’ and ‘non-plastic accent’ languages.

Lambrecht’s explanation, appealing though it is, shares with other structure-based accounts the weakness of neglecting more general explanations for the occurring accent patterns and the striking similarities of accent patterns across languages, whether they be of the flexible-accent (plastic accent) type or the flexible-syntax (non-plastic accent) type. This issue will be taken up in Section 3.3 and thoroughly discussed for EA in Chapters 4 and 6, where I will offer a different, multi-layered and multi-causal account of the phenomena at hand that, as I believe, holds for the empirical prosodic facts in both types of languages.

In general, Lambrecht’s account has the virtue over the ‘intonationist’ approaches in that it thoroughly defines the pragmatic structure the prosodic patterns are related to. However, due to its preoccupation with the one main accent in a domain (which it shares with other structure-based accounts) and its concentration on the marking of discourse referents, his analysis is too crude to account for accents other than topic accent and nuclear focus accent. Although he mentions ‘secondary accents’ he does not deal with them systematically. By contrast, Gussenhoven’s analysis accounts for the not so rare cases of ‘an accent on every content word’ through the flexible rules of domain formation, but his in his account all accents are ‘focus accents’, and there is no place for topic accents.

3.1.5 Summary and conclusion

In this section I have reviewed some of the most important focus-to-accent approaches that have been suggested in the literature. I have only reviewed the older, original work on this issue, concentrating on the prosodic approaches (Ladd, Gussenhoven, Selkirk) and the functionalist approaches to information structure (Halliday, Chafe, Schmerling, Lambrecht). However, the ideas expressed there have found their way into the more recent models and reappear in various combinations.

The main purpose of this eclectic overview was to trace back the relationship of focus and accent or accentuation in the history of research in the information structure-prosody interface and to outline the development of the structure-based focus-to-accent approach, the prevalent theory today. As already noted above, the comparison of the models suggests that focus is not a uniform category. It rather seems that there are two different, although interrelated, concepts at stake that are assumed in the different approaches. One is focus as pragmatic propositional category (whether syntactically defined or not) and one is focus as a point of information which is usually equated with accent. Bolinger’s concept of focus is explicitly related to this second focus concept while Lambrecht’s focus clearly belongs to the first category. Taking a closer look at Halliday’s model, we find that he differentiates
between two layers that both are merged in many more recent models of information structural (Lambrecht 1994; Vallduví 1992 and others) – his information structure that operates on the phonological level and deals with information focus and his thematic structure that operates on the syntactic level. An example may illustrate the difference of the approaches. While in Lambrecht’s model the comment of a topic-comment sentence is the focus of the proposition. The topic-comment (or theme-rheme division) in Halliday’s account has a priori nothing to do with information focus. The comment and the information unit may only accidentally coincide in case the comment only has one accent on the final content word and the topic is accentless. As Halliday (1967: 205) notes, “in the unmarked case the focus of information will fall on something other than the theme”.

Gussenhoven’s model makes recourse to both types of focus – the propositional domain which is marked out by the variable [+focus] and the focus domains and their heads the (focus) accents, which are the quasi-equivalent of Halliday’s information unit and information focus his. However, note that Gussenhoven, following Höhle (1982), assumes discontinuous focus domains domains (1999: 49ff.), implying that his concept of the propositional domain differs from the one as defined by Lambrecht.

Finally, Selkirk (1984, 1995) is most explicit in assuming pitch accents that are foci to be freely assigned to word-level constituents and focus domains that may be projected by these foci according to syntactic rules. Thus, in Selkirk’s account focus structure hierarchical with focus (i.e. accent) on the lower domain and focus (i.e. propositional constituent) on the higher domain. In a paper comparing these two basic focus projection rules presented in the literature, (Gussenhoven 1999) comes to the conclusion that focus projection has to be restricted to the kind of domains derivable by the SAAR, which is tantamount to saying that in Gussenhoven’s model there is no “vertical focus projection”. This latter term was introduced by Büring (2006) who also suggested that vertical focus projection up the syntactic tree should be dispensed with. There has been a wealth of other formal proposals in the literature on the prosody-focus interface that have been put forward in the past decades and especially in the past ten to fifteen years that I will refer to in later chapters when discussing the EA facts. For the time being, this little survey may suffice to present the corollary drawn for the present study.

As already laid out in some detail, the approach taken here differentiates between focus in the Bolingerian sense that has nothing to do with a pragmatic relation of ‘focus’ as assumed by Lambrecht. I follow Bolinger in the assumption that focus is a point of information signaled by highlighting (i.e. by an accent) the occurrence of which is not based on structure. I, however, draw a conceptual distinction between accent and focus, assuming that not every accent is a focus accent. On the other hand, the present model also incorporates ‘focus’ in Lambrecht’s sense as a pragmatically structured proposition, the rhematic FOCUS. These two are a priori independent. It will, however, be assumed that a narrow FOCUS usually contains a focus. In spontaneous speech, however, not every FOCUS domain is signaled by a focus. A focus accent may occur on topics as well, but focus has a
strong affinity to FOCUS. Thus, it is not to be expected that in a single proposition, the topic is characterized by a focus and the FOCUS is not. In this respect, the present model is in line with recent proposals in the generative literature (Féry & Samek-Lodovici 2006; Selkirk 2008; Katz 2011) in assuming that default accents are a matter of phonology while focus accents are a correlate of especially highlighted semantic contents (contrastive focus in the respective models; cf. Chapter 4.5.4 for a discussion).

3.2 Tonal marking of topic and focus

The preoccupation with accent placement as a means of encoding information structural categories has its origin in the prominent role languages such as English, Dutch and German have played in the investigation of prosody in general. If we were to look at the phenomenon from the perspective of many other languages, the idea of equating one accent with the one signaling a certain focus structure would probably never occur to us.

Another intonational feature that is relevant to information structure is the tonal component of prosody, which has been given less attention as far as its relation to information is concerned. The correlation between rising F0 and topic or theme on the one hand and falling F0 and comment or rheme on the other has been noted in early work dealing with the communicative meaning of intonation ([Navarro Tomás 1974 #985 [1944]; Bolinger 1958; Brazil 1975] and later on also in some generative studies (Jackendoff 1972; Gussenhoven 1984\(^{12}\); Uhmann 1991; Féry 1993,Büring 1997, and others). As a topic usually precedes the FOCUS, the unmarked tune in a declarative sentence in the West-Germanic languages is the hat pattern. Hirst and de Christo (1998: 20) note that “[i]n most languages the falling nucleus is generally prepared by a rising pitch occurring on the first stressed syllable of the unit” which they refer to as pitch “onset” which together with the falling ‘offset’ results in the hat pattern if no other accents intervene.

In the following subsection we are going to look in some detail at three approaches that have dealt with tones as correlates of information.

\(^{12}\) Note that in Gussenhoven's model the rising accents are also focus accents. His characterization of the fall-rise as selection, however, shows that its function is assumed to be the same as what other authors call contrastive topic.

141
3.2.1 Bolinger: thematic and rhematic accents

We have seen that for Bolinger accents directly express the emotional involvement of the speaker and his attitude towards what he is saying. On the one hand, this fact is related to the occurrence of accents as markers of ‘interest’ or ‘power’. On the other hand, Bolinger also recognizes tonal movements as markers of information structural categories, such as topic/theme and comment/rheme. He invokes the metaphor of a question-answer pair to illustrate the tonal characterization of these categories:

There’s a part that lays the ground-work, that asks the question, that relates to what we already know or can guess, and a part that adds the figure to the ground, that answers the question, that supplies what was not already known. The first part is called the theme and the second part the rheme. (Bolinger 1986: 47, emphasis in the orig.)

Bolinger (1958, 1986, 1989, et passim) identifies three types of accents, his ‘profiles’ which he arbitrarily calls A, B, and C. Accent or profile A is the basis citation (rise-fall) that is correlated with the rhematic part of a sentence, being the “assertive profile par excellence” (Bolinger 1986: 164, emphasis in the orig.), while accent B, which is predominantly rising is, among other things, used for thematic parts. Semantically, it differs from A in terms of “connectedness-separateness” (Bolinger 1986: 166) and thus gives rise to a whole range of semantic inferences. The third type, profile C, is described as the “mirror image” of A (1986: 149), i.e. a fall-rise. Being a kind of “anti-accen” (Bolinger 1958: 143) or a “reverse accent” (Bolinger 1972a), it is quite iconically used to play down the linguistic items it is associated with. Profiles B and C are therefore well suited to encode given items. Bolinger gives examples of the thematic use of profile C as well, such as the following pair of sentences (27a, b), taken from Bolinger (1986:182):

(27)

(a) Your bróther broke his lég. (A+A)

(b) Your bróther broke his lég. (AC+A)

Bolinger notes that, while an A accent on ‘brother’ would be typical to the answer to the question What happened?, the AC profile is used for “theme particularization” and would thus be a felicitous answer to the question What happened to my brother?

It is the combination of a B and an A accent that forms the hat pattern. Bolinger (1986: 50) notes that “[t]he favored, typical, ‘unmarked’ shape of an utterance containing a theme and a rheme is to have
two main accents, thematic and rhematic, one toward the beginning and one toward the end.” Importantly, this pattern is much more widely used than strictly for the pragmatic encoding of topic (theme) - comment (rheme), e.g. in citation forms of single words, such as *ónomatopoétic* (ibid.) and many other constructions. In Chapter 2.1.4 we have seen that EA also makes frequent use of the hat pattern and a variant of it, the roof pattern, although its use on a topic-comment sentence is the exception due to the accents that usually intervene between the topic and the final accent of the comment. However, the EA facts give evidence for the more general validity of Bolinger’s profiles and his dichotomy of *connectedness-separateness* that will be discussed in more detail in Chapter 4.5.5 and applied to the EA data throughout this study.

Bolinger’s A and B accents have been taken up by Jackendoff ([Jackendoff 1972 #1010 /yearonly /nopar]) in his treatment of information structure. Like Bolinger, Jackendoff associates the B accent with the topic and the A accent with focus. Jackendoff’s work was especially important in the development of formal linguistic accounts (e.g. Büring 1997). The important aspect of his work was that he dealt with the phonological realization of topics. His proposal will not be discussed in more detail here, because the phonological aspect that is at issue here has already been described in the discussion of Bolinger’s work.

### 3.2.2 Brazil: referring and proclaiming tones

A theory of intonation that is specifically related to discourse functions, is David Brazil’s work (Brazil 1975, 1978, 1997). Like Bolinger and Chafe, Brazil purposely refrains from matching intonational categories with syntactic structure and explicitly denies their reliance on any kind of linguistic decisions. For him intonational meaning is exclusively pragmatic, the intonational choice is dependent on the “context of interaction” (Brazil 1997: 25). A speaker, in making such a choice, is at the moment of the utterance, making assumptions about the state of understanding between him and his interlocutor.

The intonational choice is in the first place related to the category of ‘tone’. Brazil identifies five tones along with their communicative values: fall and rise-fall, rise and fall-rise and level tone. The basic dichotomy is between two pairs, the (rise)-fall and (fall)-rise. These pairs are assigned two opposing communicative values, *proclaiming* and *referring*, respectively. Brazil (1997: 68f.) ascribes to the proclaiming ((rising)-falling) tone the function of signifying something “freshly introduced into the conversation” and to the referring ((falling)-rising) one the signification of something “already in play” or “what we are talking about” (p. 68, *emphasis in the orig.*). The primary choice a speaker

---

13 The following is based on Brazil (1997) which is a slightly amended version of a book originally published in 1985.
makes is between proclaiming and referring, the second decision pertains to the exact tonal shape these tones are given. He regards the fall and the fall-rise as the unmarked or at least more frequent members of each pair, suggesting that a fall may be substituted by a rise-fall and a fall-rise by a simple rise (Brazil 1997: 82f.). The third tone, the level tone, is regarded to indicate that the linguistic items thus pronounced have no communicative significance in the speaker’s or the hearer’s world (p. 133). That is, while referring and proclaiming have clear discourse functions, the level tone has none. This is illustrated in the following diagram representing the speaker’s decisions as to which tone to choose for a certain tone unit:

(28)

When we compare these tonal categories to Bolinger’s profiles, three main differences between the can be detected: i) Bolinger’s categories are not exclusively based on discourse functions, but are rather assigned much more general functions, ii) Brazil recognizes a level tone on prominent syllables that is not compatible with Bolinger’s accents which are themselves defined by pitch obtrusions iii) Brazil’s classification of the tones is a binary one matching his two main pragmatic categories given/referring and new/proclaiming (if we ignore the neutral tone for that purpose) which is then subdivided into subcategories, while Bolinger’s is a ternary categorization, based on the physical properties of the accent and not on its communicative value. As far as meaning and the sound-meaning relation is concerned, there are, to be sure, striking similarities between the approaches. Abstracting away from the differences, we can compare Bolinger’s B and C accents to Brazil’s referring tone and Bolinger’s A accent to Brazil’s proclaiming tone, both on the basis of the melodic contours they exhibit as well as the pragmatic functions they fulfill.

3.2.3 Gussenhoven: addition, selection and relevance testing

The last model to be reviewed here is Gussenhoven (1984 [1983]). As Gussenhoven himself notes (1984: 252), his model is in spirit very much akin to Brazil’s suggestions. One aspect of the latter that Gussenhoven refutes is the idea that rises are the marked or intensified versions of fall-rises. He gives many examples where these two tones are not readily commutable or, if so, yield different semantic
interpretations. Instead, he proposes a tripartite categorization of tones to which, like Brazil, he attributes quite specific meanings. The tones identified by Gussenhoven are: the fall with the semantic function of ‘addition’, the fall-rise with the meaning of ‘selection’ and finally the rise to which he ascribes the meaning of ‘(relevance) testing’. It is noteworthy that Gussenhoven (1984: 201) views these tones as “manipulations” of the nuclear tone. In accordance with his definition of focus (cf. 3.1.4.6), these manipulations must be regarded as pertaining to the main contribution of the utterance, or in Gussenhoven’s terms to the “Variable”. The three manipulations of the nucleus modify this variable by either 1) ‘adding’ it to the background using the fall, or 2) ‘selecting’ it from the background using a fall-rise or 3) ‘not committing oneself’ as to whether the Variable belongs to the background or not, which intention is expressed by the use of a rise. Clearly, his notion of ‘selection’ is largely equivalent with what Brazil calls ‘referring’ and his notion of ‘addition’ corresponds to Brazil’s ‘proclaiming’.

As has already been pointed out in section 3.1.4.6, it is the notion of ‘background’ that is problematic from an information structural perspective in Gussenhoven’s analysis. There seems to be a sense of contradiction in these definitions which come as a consequence of Gussenhoven’s rejection of the topic category. For him an accented topic automatically becomes a focus. Given the semantic characterization of ‘focus’ as the “declared contribution to the conversation” (Gussenhoven 1983: 383), “selection from background” cannot be a focus. However, if we understand his focus as a two-dimensional concept, as suggested above, the analysis becomes transparent. We could still assume a topic-comment, i.e. TOPIC-FOCUS structure (cf. ch. 5) and interpret the highlighted elements as foci.

One aspect of Gussenhoven’s treatment of the tonal make-up of the nuclear tones is of special interest to the position taken here. Gussenhoven regards the tonal choices as an amendment of accent placement. In his view, the location of the accents only signifies the “distribution” of focus; but given that there is no one-to-one correspondence between the accented items and focus, the accent locations are compatible with different focus distributions and thus yield underspecified and therefore ambiguous surface structures. The resulting variation in the realization of the sentence accents is due to the fact that the speakers choose among the different nuclear tones which, at least in my interpretation of Gussenhoven’s idea, will narrow down the range of possible interpretations. Gussenhoven (1984: 201) writes:

Now, if sentence accents, which as we have seen are there to realise the speaker’s focus options, receive a further specification in terms of the FØ movement that nuclear tones consist of, it is natural to assume that the semantic contribution of these tones consist in a further specification of the status of the ‘contribution’ with respect to the ‘background’. Specifying the focus contribution is not enough: the speaker must specify what relationship exists between contribution and background. It is proposed that it is this relationship that provides the common element in the semantics of the nuclear tone paradigm.
3.3 A model of pragmaprosody

Starting from the theoretical considerations discussed in the preceding sections and in chapter 2, I will outline the main tenets of the model for the prosodic realization of information structure that is used for the following empirical analysis of Egyptian Arabic, and that can, as I believe, also be useful as a conceptual framework for the description of this interface in other languages.

Based on the ideas of Natural Phonology (Stampe 1979, Donegan & Stampe 1979, Dressler 1984, Hurch 1996a, 1996b), I will lay out a preliminary model of Pragmaprosody that serves as the conceptual framework of the present study.

3.3.1 Prosodic and pragmatic preferences

The present account relies on the principal divide between prosody as a natural phenomenon and its functions as at least partly conventionalized meanings. The approach outlined here is based on the assumption that it is the interaction between prosody and pragmatics that gives rise to the form-meaning relations that are observable. On such a view, it is not the case that grammatical operations necessarily have precedence over prosody. In the line of Dressler’s Morphonology (1985, 1996) and the Hurch’s Morphoprosody (1996b), I will lay out some ideas for a model of Pragmaprosody that rests on the assumption that the actual intonation patterns of linguistic utterances are a result of the interaction of prosodic and pragmatic processes (in collaboration with components of grammar, especially syntax) that may be in conflict with each other. Thus, naturally occurring structures are made use of through the process of tune-text association to express linguistic meanings. However, it is especially the deviant or marked pattern that is better suited to encode a linguistic category. As Hurch observes, accent is used as a device of signaling a morphological category “where the accentual structure itself does not follow purely prosodic principles but – on the contrary – fulfills its function best when it is a negation of the language specific syntagmatic prosodic processes.” (Hurch 1996b: 193; emphasis added).

Take as an example the natural process of phrasing which has its physiological basis in the necessity to breathe during speaking and the limits the available airstream imposes on speech production. It has its perceptual basis in the pressure for rhythmic structuring. In the process of matching tune and text, phrases fulfill the function of parsing out meaningful units of the text and are thus more often than not coextensive with syntactic or semantic constituents. But this correlation is by no means perfect. As we have seen in chapter (2.1.3.3), the nursery rhyme (4) is not phrased according to semantic or syntactic principles, but rather follows purely prosodic ones in order to preserve a strict rhythm. To be sure, everybody has previously experienced the distorting effect of poorly matched lyrics and melody in
songs of poor artistic quality. In real speech, it seems hardly possible to keep up perfect rhythm and prosodically induced phrasing without sacrificing meaning. On the contrary, it is rather rhythm and other prosodic principles that are sacrificed for the expression of meaning.

To see how prosodic and pragmatic principles interact to establish a form-meaning function let us first look at the prosodic preferences in the next section. In the section following it I will suggest three pragmatic principles referred to in the present study. The list of preferences proposed is by no means exhaustive. I will only discuss some processes that I claim to be relevant to the empirical study of information structure in Egyptian Arabic in the following chapters of this study.

3.3.2 Prosodic preferences

a) Metrical preferences

*The Principle of Rhythmic Alternation*

This preference ensures that a stretch of speech be rhythmically structured. Such a principle has been incorporated in most work on accent and intonation (Liberman & Prince 1977; Prince 1983; Hayes 1981, 1995; Selkirk 1984; and many others).

*The Principle of End Weight*

This preference accounts for the fact that the ends of domains are frequently more prominent than other parts of the same domain. In some sense this principle is akin to the *Nuclear Stress Rule*, the main difference being that it is neither syntactically motivated, nor does it assign a nucleus to the intonation unit. Some of the accounts dealing with focus prosody also make this basic distinction between and focus accent placement and prosodically induced strengthening of the domain end such as Selkirk (1984) and Jacobs (1988). The position taken here shares with Jacobs’ proposal the idea that domain end strengthening is optional (Jacobs 1988: 117).

b) Tonal preferences

*The Principle of High-Low*

This is a preference that requires pitch to go first up and then down and not the other way round. This sounds like a trivial observation, but it accounts for the empirical fact that intonation curves are generally concave in shape and not convex. The beginnings of intonation units are normally situated around the medium frequency of a speaker’s pitch range and then the pitch rises before it comes down again. Thus, highlighting by pitch accents is usually realized by a rising pitch with the frequent association of the highest point with the stressed syllable and a subsequent fall as shown for the unmarked accent in EA (Chapter 2.2.) and less frequently by what Bolinger (1972a) called *inverted*
The articulatory motivation for this lies in what Gussenhoven (2002, 2004: 89-90) termed the Production Code. Speech production is tied to the exhalation phase that has its prosodic reflection in Lieberman’s breath group (Lieberman 1967). The fact that more effort is expended at the beginning of an utterance which declines towards the end is responsible for the observable high beginnings and low ends. The high subglottal pressure at the beginning of the breath group results in higher frequencies and the decrease of the air flow is naturally accompanied by a falling slope. The Principle of High-Low is also related to a preference for falling accentual structures suggested by Hurch (1996a, b, 2000) that argues that trochees and dactyls are more natural than iambs and anapests.

The Principle of tonal prominence

This preference accounts for the fact that accents are preferably tonal. Since Fry’s experiments in the 1950s (1955, 1958) it has been assumed that the most reliable cue to prominence is pitch. This is even true for languages that have traditionally been viewed as stress-accent languages such as English as shown by Beckman (1986) in comparing English with the pitch accent language Japanese. The prominence-lending function of pitch has been further assumed in work within the framework of the Dutch school of intonation, the so-called IPO tradition (t’Hart et al. 1990). Starting from Bolinger, the idea that accent is always signaled by pitch obtrusion has become the usual way of thinking about prominence. As we have seen in the discussion of accentuation in Chapter 2.1.1, it is also possible to view accent as being cued predominantly by duration without any pitch obtrusion (Kohler 1991). More recently, phonetic studies have been conducted to show that the old category of dynamic stress is indeed justified, showing that spectral emphasis is a more reliable acoustic correlate of linguistic stress than overall intensity (Slujter & van Heuven 1996 for Dutch, Ortega-Llebaria et al. 2007 and Ortega-Llebaria & Prieto 2006 for Spanish and Catalan). These authors claim that spectral intensity and duration enable listeners to detect linguistic word stress in the absence of pitch accent. However, other studies have found no evidence for that claim (Campbell & Beckman 1997 for English, Sadeghi 2007 for Persian). It thus seems that languages differ in this respect. Crucially, however, none of the languages that appear to have dynamic stress does not make use of pitch as a correlate of prominence. Even a tone language such as Chinese uses pitch range variation to indicate focus (cf. Xu 1999). In sum, this approach regards accent as the realization of prominence, no matter how it is coded. In any case, EA is a language in which the preference for tonal prominence is fully borne out.

3.3.3 Pragmatic preferences

The other type of natural preferences to be discussed here are pragmatic preferences concerning the information structure of an utterance. Such preferences are cognitively motivated and are immediately comprehensible using common sense. I tentatively identify three such principles that I believe to be
necessary to account for some basic facts of the encoding of information structure in language in general and the prosodic encoding in particular.

The principle of marking pragmatic properties

This principle assumes that languages preferably mark semantic contents as important or interesting by foregrounding and concepts/ideas that are ‘taken for granted’ by backgrounding. This may be done by various linguistic means, among them word order. Maybe the most important means, however, is prosody. Languages seem to differ whether they tend to background semantic items only because of their information status (cf. 1.1) like English, or only if the denotations of the given lexical items are part of the presupposition, like EA (see 4.5.2).

The principle of marking pragmatic relations

This principle assumes that languages tend to mark the pragmatic relations of TOPIC and rhematic FOCUS. This may be done via morphological markers such as Japanese wa to indicate the topical function a referent fulfills within a proposition, or syntactically by topicalization, or last, but not least, prosodically by means of a rising tonal configuration connecting the topical part of an utterance to the subsequent comment.

The principle of logical sequencing

This principle relates to the fact that what is being talked about is preferably mentioned first as a starting point before it is commented on. This natural process accounts for the probably universally preferred order of topic or theme before comment or rheme. Clearly, this pragmatic preference interacts with the syntactic component of a language.

3.3.4 The interaction of prosodic and pragmatic preferences

As already noted, I assume that prosodic and pragmatic preferences interact to produce language-specific forms for the expression of information structure. In this section, I will only give some examples of possible interactions between prosodic and pragmatic preferences. Preferences may, but need not be in conflict. This is also true for the preferences of one component among each other. For example, the principle of marking information status by definition conflicts with the principle of marking informational relations in case a topic is prosodically marked by accent - as typically given referents assume topical status in a proposition, two pragmatic preferences are at work at the same time: the tendency to establish the topical relation to the rest of the proposition by accenting and the tendency to downtone given material.
Another conflict of pragmatic preferences that will be the subject of Chapters 4.5 and 5.7 is the marking of new topics, e.g. in out-of-the-blue utterances. Such topics may be regarded as part of the FOCUS domain. At the same time, they have a topic relation to the proposition they are part of. In such cases a conflict between TOPIC and FOCUS marking strategies will arise.

Preferences also may conflict across components. The one case of conflicting pragmatic and prosodic principles that I believe to be responsible for the alleged typological distinction between plastic and non-plastic accent languages as suggested by Vallduví (1992 and other work) is the tension between the Principle of marking pragmatic properties and the Principle of rhythmic alternation. According to the first of these two, a narrow FOCUS should ideally carry the only pitch accent in an utterance with the rest of the utterance being deaccented, as it is the case in English or German for example. On the other hand, the prosodic principle requires roughly equally spaced accents to occur which inevitably results in a number of accents throughout the utterance, contradicting the preferred singling out of the narrow FOCUS by assigning it the only accent in the utterance, in accordance with the Gestalt principle of figure against ground. In El Zarka (2005), I have suggested that Vallduví’s typological distinction is actually based on a typological distinction between strict-rhythm and non-strict-rhythm languages with English and German belonging to the second group and Egyptian Arabic and Italian to the first. Clearly, all languages are rhythmically structured, but it seems that this rhythm is more easily sacrificed in some languages than in others. Of course, the prosodic type of a language correlates with its syntactic type. Thus if a language has a flexible syntax, it may preserve rhythm and utilize it for other means such as the marking of individual lexical items to make them more easily identifiable in the speech flow. But if a language has an inflexible syntax as English has, it may be forced to sacrifice rhythm to be able to use prosody unequivocally as a means of expressing information structural categories. We could thus assume languages to be situated somewhere on a typological cline between strict-rhythm and non-strict rhythm, with EA, for example, more towards the one end of the cline and English more to towards the opposite end. By the same token, it may be noted that unmarked information structure, largely equivalent to broad information FOCUS shows less deviant prosody than marked information structure, such as narrow contrastive FOCUS.

To sum up, the view endorsed here lies somewhere in-between the structural and the radical focus-to-accent approach in that it assumes a language-specific prosodic structure. The actual intonational realization of an utterance is conceived of as the outcome of the interplay of various linguistic and non-linguistic factors, of a pragmatic (including attitude), semantic, syntactic, or phonological kind, and is therefore in a sense unpredictable. I follow Bolinger in the contention that intonation, although it may acquire a certain degree of arbitrariness, should be analyzed in terms of “meaningful gesture” rather than “abstract phonology” (Bolinger 1986: 140). I, however, believe that the idea of functionally motivated constructions is in line with Bolinger’s conviction that in intonation, there is no “distinction between the grammatical and the ideophonic except as they represent extremes of a scale” (p. 32).
The model proposed here is illustrated in Figure 1. It is a flexible model that does not claim to be able to predict every single pattern that is observable in a certain language, but rather tries to explain the motivation behind them. How the individual solutions in a language look like has to be investigated on the basis of thorough empirical studies. The observable patterns may be described as holistic constructions involving different linguistic features. Such an approach has the advantage of accounting for the empirical facts without postulating mechanistic rules, exceptions to which are legion. The assumption of preferences that are in competition with each other predicts variation, and the actually occurring forms can be explained by recourse to these general principles and described as constructions.\(^{14}\) In such a model, deaccentuation and downtoning are subsumed under the same principle of backgrounding semantic material that is taken for granted, subsuming semantic and pragmatic weight under one heading of ‘interest’ as suggested by Bolinger. Moreover, such a model is compatible with the existence of variability within one language. Finally and most crucially, it relates the similar behavior of so many languages to a handful of principles.

\(^{14}\) I take the view that, although Optimality Theory is actually based on similar ideas, an optimality theoretic account is a formal mechanism that only replaced the older derivational rules, claiming to be able to account for all kinds of possible empirical data as rule-based, leaving no room for idiomatic and rote-learned structures.
4 The prosody of information structure in Egyptian Arabic: pragmatic properties and pragmatic relations

Based on the foundations laid out in the previous chapters, I will outline the main assumptions about the prosodic marking of information structure in Egyptian Arabic that will be further substantiated in the subsequent chapters on topicality (ch. 5) and theticity (ch. 6), based on the prosodic model that has been presented in Chapter 2.

In Section 4.1 a very brief overview of the literature on information structure in Arabic in general and Egyptian Arabic in particular will be given. The literature overview will only cursorily reflect the work done so far and will rather concentrate on studies and claims made therein that are directly related to the present investigation. In section 4.2 I will present the corpus data the present study is based on. Section 4.3 will be dealing with the prosody of pragmatic properties, considering both information status (givenness) and focus-background structure, and Section 4.4 will be concerned with the prosody of the pragmatic relations of TOPIC and FOCUS. In the latter two sections the results of two production experiments will be presented in addition to the qualitative corpus study. As this book concentrates on topicality and theticity, I will only briefly dwell on the issue of different focus types, association with focus (Jacobs 1983; Rooth 1985, 1999; Krifka 1992, 2006) describing the use of focus sensitive particles (also, only, even) will not be touched in this book; I will also not have much to say about syntactic and morpho-lexical focus marking strategies, an issue that will only be briefly surveyed in Section 4.5.6. Section 4.5 will deal with different types of 'interactions', i.e. interactions between the different information structural components (4.5.1 and 4.5.2) and interface phenomena such as the problematic issue of nuclei and focus exponents (4.5.3), the status of contrast, its relationship to focus and its prosodic realisation (4.5.4). In Section 4.5.5 the prosodic concepts of integration and separation and their bearing on the realization of FOCUS domains exceeding a default accentual domain will be discussed. Section 4.5.6 will briefly discuss the related notions of narrow and broad FOCUS. Finally, Section 4.6 gives a summary of the main empirical results and discusses the theoretical implications emerging from the study.

4.1 Information structure in (Egyptian) Arabic: the state of the art

Studies on information structure in Egyptian Arabic and Arabic in general are not too numerous, although EA is an otherwise quite well-researched variety of Arabic. Even the recent very rich and detailed grammar by Manfred Woidich (2006) does not contain a comprehensive section on information structure, but only deals with the phenomenon cursorily. Relying on a wealth of spontaneous speech data, it does, however, discuss a considerable number of syntactic constructions
that readily lend themselves to an explanation in terms of information packaging. I am also unaware of any monograph on the information structure of any other regional variety of Arabic.

The only book-length investigation of information structure in Arabic, albeit not in EA, is Moutaouakil (1989). Moutaouakil’s study is aimed at explaining different syntactic constructions in Classical Arabic, a language that presumably was the literary koiné more than 1000 years ago, which implies that it most probably was not even used for ordinary conversation at that or any other time. Based on Moutaouakil (1989), a standard analysis within a formal framework of focus\(^1\) in Arabic focal structures is the much cited article by Ouhalla (1993).

There exist several smaller studies on different aspects and different varieties of Arabic (e.g. Mughazy 2010, Holes 2010, Owens et al. 2010). An insightful discussion of aspects of topic-comment structures under the heading of copular sentences is offered by Eisele (1999).

To date, the most comprehensive description of information structure that deals with spoken language, based on four different spoken Arabic varieties, is a chapter dedicated to that issue in Brustad’s book *The Syntax of Spoken Arabic* (Brustad 2000).

The picture that emerges is that Arabic in general is a language with a wealth of syntactic structures that are pragmatically motivated. Not only can a sentence start with virtually any constituent encoded as a topic, there are also many different topic constructions which all seem to be extensively used. Focus is also encoded syntactically, either by fronting a focussed element or by the abundant use of various cleft constructions. All studies mentioned only cover the syntactic aspects of information structure, however, and do not consider prosodic aspects in a systematic way.

As regards prosodic studies, a small pilot study by Norlin (1989) has to be mentioned. Within a generative framework, Sam Hellmuth has published a number of studies dealing with the prosody of information status, focus and theticity in EA (Hellmuth 2005, 2006a, 2006b, 2010a, 2010b). My own research has partly been published in El Zarka (El Zarka 2011a, 2011b, 2012, 2013). I will not summarize the different proposals here, but rather take issue with claims made in the mentioned studies in the course of the discussion of the relevant questions in the following sections.

---

\(^1\) When discussing other work on information structure, the terms *focus* and *topic* is not to be understood according to the definition it is given in the present study, but as a cover term for the different definitions that are employed in the individual works.
4.2. The corpus: methods and data

In this study, great importance is attached to the claim that the investigation of grammatical structures be based on naturally occurring language. In addition, the study also relies on experimental data to control the information structure of the utterances and to get a larger amount of comparable data covering different types of contexts, which is a prerequisite for the investigation of prosody.

Data from spontaneous discourse have been subject to qualitative investigation and the data from the production experiments have additionally been submitted to statistical analysis. Prosodic analyses were carried out on the basis of auditory and acoustic analysis using PRAAT (Boersma & Weenink 2012). For the acoustic analyses, the standards of recent research methodology in the field of intonational phonology have been followed. Thus, the study is aimed at bringing together two major lines of methodology, the descriptive methods of the typological linguist and fieldworker and the rather technical approach and practice of the experimental phonologist doing acoustic phonetics. Based on such considerations, the corpus chosen for the investigation is composed of the following data:

- a corpus of data elicited on the basis of a questionnaire (QUIS) developed by the D2 project, SFB 632 (Skopeteas 2006),
- Egyptian narratives (spoken and read speech),
- other narratives (including a narrating task based on a picture book, “The frog story” by Mercer Mayer),
- natural conversations and expository monologues (recipes and other texts)
- Egyptian television talk shows from (Sabaḥ il-xe:r ja masˤr ‘Good morning, Egypt’ and Miḥwar)
- a play written in Egyptian Arabic by Tarek Eltayeb (Al-Asanse:r) (read speech)
- a compilation of short stories written in Egyptian Arabic by Khaled Al Khamissi (Taxi) (read speech)
- production experiments to elicit topic, focus and contrastive focus

The corpus of the D2 project of the SFB 632 contains experimental tasks that have been conducted by Sam Hellmuth in 2007 in Cairo with 10 young Egyptians from the capital city and a translation task that has been conducted by Doulagy Hanna, an Egyptian student of German-Arabic translation at the department of Translation Studies in Graz under my supervision. A variety of sentences were translated and subsequently read by two female Egyptian students and one male Egyptian trainee teacher. The corpus consists of 29 experiments and the translation task that were designed to elicit structures conveying different information structural functions, such as broad/narrow focus, contrastive focus, givenness and several types of topics. The translations and the experiments were particularly useful in the initial phase of the investigation to gain preliminary insights in the field and

---

2 As an associate researcher of the D2 project I was responsible for the transcription and annotation of the whole corpus and the supervision of the translations from English to EA.
establish some basic structures. They have, however, to be used with caution. As is well known, translations are prone to be heavily influenced by the linguistic structures of the source language.\(^3\) Some of the experiments are also problematic in some respect, but I will discuss these problems where relevant.

My own corpus comprises several hours of recorded data which was all transcribed, annotated (for the most part) and archived in an xml-format in EXMARaLDA, a database developed by the Sonderforschungsbereich 538 "Mehrsprachigkeit" (www.exmaralda.org; Schmidt & Wörner 2009). This database was chosen because the QUIS data had also been archived in this database. As EXMARaLDA lacks a search tool that offers the possibility of multiple queries, the data were also transferred to the language documentation software ELAN, developed at the Max Planck Institute of Psycholinguistics in Nijmegen, Netherlands (http://tla.mpi.nl/tools/tla-tools/elan/). ELAN offers complex search options. In addition to some pilot experiments, one production experiment was conducted with 10 speakers from Cairo and Alexandria, all of whom had been living in Graz for a short period at the time the experiment was recorded. In this experiment I used stage communicative events, giving the speakers some basic information about the contents of the conversation as a prop beforehand. The aim was to elicit thetic utterances; the experiment will be discussed in Chapter 6. To elicit different types of FOCUS, another two production experiments using read speech were conducted with 6 speakers (see 4.3.1.1 and 4.4.1.1). The read speech data (narratives and drama) were recorded with a young female speaker from Alexandria and the narratives with four speakers (three female and one male) in Graz. During a field trip to Alexandria in 2010 I recorded the conversational data and the monologues with four female speakers. To be able to base the research on a large amount of data which can be subsequently used – by myself or by fellow researchers – I could rely on the help of student research assistants. Fortunately, some of my students had experience with digitising and transcribing data. They had already been subcontracted for the preparation of the QUIS data.

Annotation of the data was carried out roughly along the lines developed by the SFB 632 (Dipper et al. 2007). The data of the production experiments were manually labelled by myself in PRAAT and subjected to quantitative analysis using the statistics programme SPSS. For the statistical analysis I relied on technical assistance. The basic annotation included a broad phonetic (almost phonemic) transcription including marks for intonation phrase boundaries (tier 1), a morphemic glossing (tier 2), the annotation of clause structure (tier 3) and of parts of speech and syntactic function (tier 4). Furthermore, the data was subject to prosodic analysis on three different levels: tone level (H and L tones) (tier 5), tonal category (leading, linking, closing) (tier 6) and occasional description of other prosodic cues (tier 7), such as duration and intensity. Regarding information structure, I annotated

\(^3\) It is worth mentioning that the translator had received a sound training in the methods of translation and thus avoided the pitfalls of transferring linguistic structures from one language to the other.

156
information status (tier 8), using a crude taxonomy of given - accessible - new. To operationalize the
annotation it was decided that a denotation counts as given if mentioned within the previous five
intonation phrases (cf. Baumann & Riester 2010). Strictly speaking, information structure (topic,
TOPIC, focus, FOCUS) was not 'annotated', but had to be interpreted on the basis of the context and
the formal cues in the speech signal itself (word order, prosody etc.). These interpretations were
annotated in two tiers (tier 9 and 10) to register interactions between the individual concepts, the
second tier containing foci of interest as well as occurrences of contrast. The whole annotation was
linked to the audio file in EXMARaLDA; a screen shot of one file is presented in Figure 1.

Fig 1: Screen shot of a corpus file in EXMARaLDA.

4.3. The prosody of pragmatic properties

In chapter 3 I discussed in some detail different proposals concerning the prosodic encoding of focus
and partly of givenness, i.e. the information status of referents, in the literature. It has been claimed,
for instance, that in (British and American) English repeated entities are obligatorily deaccented. The
observation that in English given entities, whether previously stated or inferred, are deaccented has
already been made as early as 1781 by Walker and Bell & Bell (1879) (cf. Cruttenden 2006: 314). In
Ladd's examples (14, 15), cited in chapter 3, the concepts denoted by the deaccented lexical items are
in fact only implied. The inner argument ‘books' in 'réad books’ is concept given (Baumann 2006: 119)
by the mentioning of 'Slaughterhouse Five', being a hyperonym of the latter. In such cases the marked
shift of the main accent to the verb is obligatory. Halliday's system (Halliday 1967) also incorporates a

---

4 I should like to thank Utz Maas for this very important observation that helped me sort out things conceptually.
provision for his 'tonic' accent (i.e. 'nucleus') - which he takes to mark new information - to fall on the final lexical item in the unmarked case or before in the marked case. He states that in the latter case the early accent placement may be caused by the fact that the last item (in fact all items after the accent) denote given information. The same has been claimed to be true for Dutch (Nooteboom & Kruyt 1987). As Ladd (Ladd 1996) observes, however, even English has some varieties that do not necessarily deaccent given items. Ladd (1996: 176) quotes an example from Indian English given by Gumperz (1982): If you don’t give me that CIGARETTE I will have to buy a CIGARETTE. In the same way, deaccenting is not freely allowed in Romance languages which resort to morphosyntactic strategies instead to achieve similar effects (Ladd 1996: 178-179). Another familiar case known from the study of English is the fact that verbs are frequently deaccented, while arguments (esp. direct objects) carry an accent (Schmerling 1976; Gussenhoven 1983, inter alia). Ladd therefore draws the conclusion that languages differ in their treatment of given entities with respect to deaccenting and also whether they differentiate between arguments and predicates. He suggests that

there is a difference between deaccenting and non-deaccenting languages – or more precisely, between (a) languages that permit, prefer, or virtually require the deaccenting of repeated and otherwise given material, and (b) languages in which such deaccenting is dispreferred or syntactically restricted […] (Ladd 1996: 179)

Syntactic restrictions are reported for Italian which is said to allow deaccentuation of clauses and simple NPs (Avesani et al. 1995; Hirschberg & Avesani 1997; D’Imperio 1997, 2001; Farnetani & Zmarich 1997), but avoids deaccenting within NPs (Ladd 1996: 179). To test this hypothesis, Swerts, Avesani & Krahmer (2002) conducted an experiment to compare the accentuation strategies within Italian and Dutch noun phrases from an acoustic, perceptual and functional point of view. The data were elicited in a dialogue game played by Dutch and Italian speakers using repeated occurrences of nouns and adjectives (geometrical figures and colours, such as a ‘blue square’, a 'black triangle’ etc.) in different combinations. The results showed that Dutch speakers accented new and contrastive information and deaccented given information while Italian speakers accented all nouns and adjectives irrespective of their information status. Similar results are reported by Avesani & Vayra (2005) who used a different type of experimental task apt to elicit the spontaneous production of repeated items and also co-referring expressions. Their results show that the vast majority of repeated items in Italian experimental dialogues were accented (93% of accented words out of 290 tokens). They further argue that deaccenting in Italian is not related to information status at all. Rather the rare instances of deaccentuation fall into four categories, the largest of which is post-focal position (50 %, i.e. 9 out of 19 occurrences).

A couple of other studies have also shown that information status is not enough to trigger deaccentuation. Specifically, Terken & Hirschberg (1994) investigated the contribution of syntactic function and surface position between utterances on the occurrence of deaccentuation in English. The results of their study indicate that givenness is not a sufficient condition for deaccentuation, observing
that given entities may well be accented. Instead they argue that the persistence of surface position and syntactic function contribute to the actual accessibility of an entity to the listener and thus plays an important role in triggering deaccentuation. Interesting results are reported by Bard & Aylett (1999) for English spontaneous speech. These authors found that deaccenting of repeated mentions was far less frequent than reported in the literature, in approximately 75% of the cases there was no change of the accent (also not in accent type) across repetitions. This result was not really unexpected by the authors whose aim was to see whether Terken & Hirschberg’s claim that deaccenting is not only a matter of givenness, but influenced by other factors would also hold for spontaneous speech. They point out that spontaneous dialogues do not offer that kind of parallel structures found in experimental designs. Instead repetitions and co-referring expressions are expected to occur in different functions and positions and thus should not be conducive to the occurrence of deaccenting. Bard & Aylett’s results are in line with other research on spontaneous data (e.g. Brown et al. 1980; cf. also Fuchs 1984). Interestingly, Bard & Aylett make the important observation that given entities, even if not deaccented, are produced with less articulatory detail, i.e. shorter and “less intelligible” than their first mentions – the same result had been obtained by Fowler & Housum (1987) in a production and perception experiment with English speakers.

Another important fact about the realization of deaccentuation concerns its occurrence within a prosodic unit, especially with respect to the focal accent (cf. 1.1 and 4.5.2). It has been noted that while pre-focal given items may exhibit accents for rhythmical reasons, they are necessarily deaccented after an early focus in English. Although Italian is said to resist deaccentuation, a number of studies have found such deaccentuation in post-focal position (Farnetani & Zmarich 1997; D’Imperio 1997, 2001). It has to be noted, however, that Farnetani & Zmarich for instance, interpret as deaccented what is in fact not always completely deaccented but rather highly compressed. The authors note that “[p]ost-focal fall patterns […] have lower F0 onset values, smaller F0 excursions and rates”, than focal falls, the differences were statistically significant, the post-focal words, however, did not exhibit shorter durations, which the authors contribute to the wish to preserve lexical stresses. In the Neapolitan variety of Italian, D’Imperio (1997, 2001) also observed a structural difference between focal and post-focal accents with a difference between declaratives and polar questions. D’Imperio’s data clearly show that in declaratives the post-focal material is tonally completely flat while questions have a much reduced final accent. The upshot of the Italian results is that, at least as far as focus structure is concerned, the differences between a ‘non-deaccenting’ and a ‘deaccenting’ language are much less than expected on Ladd’s assumption. The difference can rather be ascribed to the fact that the ‘non-deaccenting’ languages are reluctant to suppress lexical prominences altogether. I have proposed (2005) that this fact may be due to different rhythmic preferences in the given languages – i.e. some languages prefer rhythmic realizations that are instantiated by regular prominences (which is
largely equivalent to accents as argued in Chapter 2.1.1) while other languages attach less importance to rhythmicity.

So far, the studies mentioned were dealing with one or maximally two languages for comparison. In a typological review of a considerable number of European languages (plus Tunisian Arabic), Cruttenden (2006) tackles the question of whether deaccenting of given information is a universal phenomenon. He comes to the conclusion that some languages resist deaccenting. On a cline from deaccenting to what he calls "re-accenting", German and English are situated on the 'deaccenting'-end and Spanish on the other extreme. French, Tunisian Arabic and Swedish are close to Spanish, but the sample contained slightly more cases of deaccenting in these languages. This property of high accent distribution, as it stands, is shared by many genetically unrelated languages such as those of the Romance language family, the Skandinavian languages (Gårding 1983; Thorsen (Grønnum) 1983; Estonian (Asu & Nolan 2007, and references therein) and Egyptian Arabic (Rifaat 1991, Rastegar-El Zarka 1997, Hellmuth 2006b). It thus seems that the strong deaccenting tendency of the West-Germanic languages is the exception rather than being the rule cross-linguistically. The question whether deaccenting is a universal phenomenon does not seem to be the right question to ask, the more promising question rather is whether Chafe's suspicion that what is given might be universally expressed in a "weaker and more attenuated manner" (Chafe 1976: 31) is right.

Some studies have also responded to the fact that given information is not necessarily deaccented. We have also seen that givenness is not a homogenous phenomenon, but may involve 'degrees' yielding at least a three-way distinction between given, accessible and new (cf. 1.1.1 and 4.1.1). While Brown (1983) concluded from an experiment on English that a more fine-grained givenness taxonomy, such as the one developed by Prince (1981), is necessary to account for morpho-lexical and syntactic facts, intonational marking only requires the distinction given – new. Some studies found a correlation between categories of information status and accent type (Pierrehumbert & Hirschberg 1990; Baumann 2006; Grice & Baumann 2006; Baumann & Riester 2010). Neglecting the differences in detail, these studies point to the existence of a givenness scale from new to given: H* > !H* > L* > no accent. The other group of experimental studies that is of special relevance to the position defended here is concerned with the scaling of pitch accents (Féry & Kügler 2008, Féry & Ishihara 2010). As these experiments suggest, givenness and focus (which in the present frame work correspond to the pragmatic properties of information status and focus) are expressed by relative pitch height. Interestingly, it is not assumed that given information is necessarily deaccented, but rather that pitch range manipulation is not simply a "yes or no matter" (Féry & Ishihara 2010: 52). I will return to these studies when comparing them to the results of the EA study in sections 4.2.1 and 4.2.2.
We have seen that the results of different experiments conducted for the investigation of deaccentuation in different languages are not homogenous, not even within one language. The lessons to learn from this are several:

a) As Bard & Aylett (1999) also point out in the conclusion to their study, the predictions from laboratory experiments do not necessarily generalize to spontaneous speech.

b) Givenness is not the only factor that is relevant to deaccentuation (or weaker articulation in general). Especially the occurrence of an early focus accent in narrow FOCUS cases plays an important role in the behaviour of the accents.

c) Languages in fact differ regarding deaccentuation. There are languages, such as English, Dutch and German, that more readily deaccent content words, while other languages, such as those of the Romance family or Arabic varieties, definitely Egyptian Arabic, resist deaccenting.

d) Deaccenting, i.e. the non-occurrence of pitch accent, is not the only way to articulate linguistic material in a weaker manner. To put it more generally, pitch is not the only option we have in prosody. Rather weaker articulation relates to durational and other acoustic effects concerning segmental properties as well as intensity that turn expressions more or less intelligible.

e) If we acknowledge the necessity to investigate also other phonetic cues involved in prosodic articulations, this amounts to acknowledging that the prosodic coding of information structural categories such as information status is a gradient, rather than a categorical matter. Hence, the assumption of deaccenting as a categorical means of prosodic expression is an unwarranted assumption.5

The history of prosodic treatments of given/new or focus/background is, however, based on the idea that languages differ in whether they use accent placement as a device of marking a category new or focus, or not. On a par with Ladd’s distinction between deaccenting and non-deaccenting languages, Vallduví (1992) proposed a typological distinction between plastic accent and non-plastic accent languages, referring to languages such as English and German as plastic accent languages and to some Romance languages, e.g. Catalan, as non-plastic accent languages. While the first type exemplifies a considerable freedom of main accent placement, the second is characterized by edge-marking nuclear (i.e. final) accents. In the present study I argue that this hypothesis is too strong. If the prevalent notion of accent vs. no accent is abandoned and a gradient concept of accentuation is acknowledged instead, Chafe’s basic insight may still be valid. In the section, I will show that in EA attenuation, i.e. accenrtual subordination, may be found instead of deaccenting to fulfil the same functions as the latter.

5 In chapter 6 I will elaborate on the argumentation concerning gradient vs. categorical accentuation.
4.3.1 The prosody of pragmatic properties in EA

In this section we will be concerned with the question how pragmatic properties are reflected in the prosody of EA. This question, however, involves two sub-questions: (i) How is the information status of referents reflected in prosody? and (ii) How is focus reflected in prosody? As we have seen in the overview above, not all researchers make that distinction. But it will be shown that such a distinction is in fact of major importance.

The prosody of information status and focus in EA has been addressed by Norlin (1989), Rastegar-El Zarka (1997) and especially by Hellmuth (2005, 2006b, 2006a and 2010b). While Norlin's short paper is a pilot study with only one speaker and Rastegar-El Zarka (1997) only briefly addresses focus as one aspect of the intonation of Egyptian MSA, Hellmuth's quantitative studies rely on experimental data, elicited especially for the investigation of the issue. In the following, I will take issue with the results of Hellmuth's studies and the conclusions drawn from them.

To test whether given information was deaccented in EA, Hellmuth (2005) replicated the experiment conducted by Swerts et al. (2002) on English and Italian with speakers of EA (Cairene) and extended the methodology to examine the hypothesis that EA does not mark givenness by deaccentuation, neither within, nor across phrases. Hellmuth reports that both hypotheses were confirmed. The design comprised production and perception experiments to test whether the discourse function (i.e. givenness of the lexical item or the phrase) could be retrieved by EA listeners. The production data unambiguously show no deaccenting in and of phrases (with the exception of one speaker whose limited use of deaccentuation Hellmuth ascribes to the speaker's near-native command of English). Unfortunately, the production data could not be used for investigating gradient marking of information status, such as pitch height differences (Hellmuth 2005: 109f.), but the function test results suggest that speakers do not identify a previous context from the prosodic shape of an utterance. This parallels Swerts et al.’s (2002) results for Italian and is contrary to the results of their English experiment.

In another experiment, reported in Hellmuth (2010b), EA speakers read out a short dialogue which was used to examine the prosodic reflexes of information focus and contrastive focus. The dialogues contained target words that were either contrastively focused or supposed to be new (which is equal to information focus in Hellmuth’s framework). The results indicate that EA uses pitch range manipulation to mark the difference between a contrastively focused item and the adjacent accents by widening the pitch excursion of the contrastive accent and narrowing the pitch range of following accents. Thus Hellmuth concludes that contrastive focus has gradient, hence non-phonological, prosodic reflexes. Regarding information focus, she reports neither categorical nor gradient prosodic reflexes. There are two problems concerning the comparability of Hellmuth’s experiment with other studies investigating information focus, however. First, what she takes to be a case of information
focus would be a topic many frameworks, also in the one adopted for the present study. Judging on the material given in Hellmuth 2010b: 172), one case of the sample could be analysed as a topic shift, while in the rest of the cases the target word is in fact even a continuous topic (ma:ma ‘Mum’ continuing the ‘entity topic’ (cf. Chapter 5) of the preceding sentence ‘my mother’). The second problem concerns the presence of a new element after a contrastively focused one. In this position it is contrasted with a given element. The presence of a contrastive focus could at least in principle even out a potential difference between new and given targets. But interestingly, Hellmuth also reports no difference between the two targets if there is no contrastive focus preceding them. This suggests that information status in fact has no prosodic reflexes in EA.

The investigation of my own EA corpus supports Hellmuth’s conclusions as far as deaccenting is concerned. This holds true for the experimental data from the QUIS questionnaire (Skopeteas et al. 2006) as well as the spontaneous monologues and dialogues and the written data. Regularly (but not necessarily) accentless are functional categories such as prepositions (1), pronouns (1) and also fillers and coverbs (2) that add an aspectual meaning to the full lexical verb following it or situate the event in time.

(1) (Mido_01_F0_02)

\begin{verbatim}
ana  h-a-si:b-ik
1SG  FUT-1SG-leave-2SG.F
L.  LH-  L-
\end{verbatim}
'I will leave you, froggie,…

\begin{verbatim}
hina fi-l-śilba l-ʔi:zaːz
here in-DEF-box DEF-glass
°H L- LH  °L^H
\end{verbatim}
'here in the glass box’

(2) (FAGR_F3_03)

\begin{verbatim}
a.  wu-rāḥ-it waʔš-aːː / kullṣ-na wʔš-na
and-go-3SG.F fall.PTCP.SG.F all-1PL fall-1PL
L- LHL- LH- L^H
'And then she fell down - we all fell down'
\end{verbatim}

\begin{verbatim}
b.  jaʃni kan mawqif s'aʃb
FILL be.3SG.M situation difficult
L- LΗ^L
'so it was a difficult situation'
\end{verbatim}

While deaccenting is commonly viewed as categorical and hence phonological, pitch excursion is commonly viewed to be a gradient phonetic cue. But if we accept the idea that deaccenting is just an extreme case of reducing prominence by the use of a compressed pitch range (Xu & Xu 2005), we find that EA makes pervasive use of prominence differences that are exploited for informational purposes. Importantly, such differences are not expressed by tonal features alone. Duration, spectral characteristics and intensity are other important acoustic correlates of prominence. However, pitch
height relations in combination with excursion size are a reliable cue to prominence, at least if the accent is not in utterance final position.

The QUIS corpus contains some experiments that investigate whether givenness induces deaccenting of the given item. One such experiment is a picture task that was designed to manipulate the discourse status of a referent (animal) and of one or two of its properties (colour, size, number). Similarly to the results of Swerts et al. (2002) for Italian and Hellmuth (2005) for EA, the QUIS data shows no deaccenting of the repeated items. If we, however, re-examine the data from the perspective of 'attenuation' instead of deaccenting, we find that there are differences between new and given items increasing the prominence or salience of the former and diminishing the prominence of the latter. These differences pertain to various prosodic features, i.e. tonal contour, phrasing, duration/tempo, intensity and segmental properties which turn given information less intelligible than new information. New information, on the other hand, is articulated with great accuracy. Let us first consider the prosody of repetitions as in (3) (pitch tracks are provided in Figure 2). The example shows two successive utterances from a picture experiment (7 Birthday Party) in the QUIS corpus. The task involves two informants who receive different cards, one of them containing three animals and three presents with arrows indicating which present goes to which animal, the other informant has a card showing only the corresponding animals or the corresponding presents. The speaker of example (3) was instructed to answer a question about what the individual animals get (juice boxes of different colour in this case). (3a) is the first answer to the question. The word order is given-new putting the FOCUS part at the end of the utterance. In the new (FOCAL) part, the noun and the adjective are phrased separately. The adjective phrase, syntactically a subordinate clause, tonally integrates the subject topic of this clause by a linking contour and the strongest prominence of the whole utterance lies on the expression of the colour, ʔaħmar 'red', which expresses the most informative semantic content. Note that although the juice box (ʕilbit ʕasˤi:r) is new and thus expressed as a full lexical indefinite NP and that it belongs to the FOCUS of the utterance, the real (contrasting) information lies in the colour. We have to keep in mind that the speaker has in front of her three juice boxes that only differ in their colours. This fact is even clearer in (3b) where the expression 'juice box the colour of which (is)' is repeated forming one phrase. What is furthermore interesting is the existence of other cues than phrasing that are responsible for prosodically backgrounding the repeated element. Note that the contour is tonally relatively flat exhibiting the natural declination towards the following strong accent. Most notably, intensity is much weaker than in the rest of the utterance which is indicated by the waveform above the pitch track in Figure (2b). Again the colour expression (ʔazraʔ 'blue') is
phrased separately also exhibiting a very high rise making the accent as the most prominent of the whole utterance.  

(3)  

a.  

ha-ni-ddi li-l-ʔarnab il-ʔahmar hidijja a:: /  
FUT-1PL-give to-DEF-rabbit DEF-red present HESIT  
'We will give the red rabbit a present ah…’  
Šībit ʕasˤiːr... / lon-ha ʔahmar  
box  juice colour-3SG.F red  
H  L^H  L-  LH^H%  
a RED juice box’ (lit. a juice box the colour of which is red)  

b.  

wi-ha-ni-ddi li-l-ʔarnab il-ʔaxdˤar /  
and-FUT-1PL-give to-DEF-rabbit DEF-green  
'We will give the red rabbit a present ah…’  
Šībit ʕasˤiːr lon-ha / ʔazraʔ  
box  juice colour-3SG.F blue  
H (H-)  ^H%  
‘And we will give the green hare a BLUE juice box’ (lit. a juice box the colour of which is blue)  

Fig. 2: pitch tracks of two successive utterances. In panel (a) Šībit ʕasˤiːr is the first mention, while in panel (b) it is repeated. Both instances are part of the FOCUS domain.  

While in (3) the repeated element was part of the FOCUS domain, the utterance in (4) is an example from the same experiment with a repetition within the presuppositional part of the utterance. Contrary to the situation in (3), in (4) the present - which syntactically is the DO of the sentence - is the given (i.e. repeated) element and the act of giving this present is the presupposed information while only the recipient is new information. Again, we see that the FOCAL part is divided into two phrases one for the NP and another for the AP with the main prominence on the colour expression as before. What

---

6 The accent is transcribed with a high boundary tone H% here. For the time being, we will not address the details of the tonal contour which will be the subject of the next subsection.
interests us here is the behaviour of the given material that shows a similar, albeit not identical, tonal contour as the one of the given element in (3b) (see Figure 2). The main difference lies in the fact that there are many bitonal accents realized in a compressed pitch range but with the full intensity typical for beginnings when subglottal pressure is still high. Nevertheless, the stretch of speech is perceived as less prominent than the final two accents. We may also say that the given part is backgrounded in several ways: firstly by compressing the pitch range and thus reducing prominence, and secondly by speech tempo that renders the linguistic material less prominent and less intelligible (note that the duration of the heavy object plus verb approximately equals the duration of the comparatively light PP), and finally by the less articulate pronunciation of the segments which induces all kinds of reductions and lenitions, such as the shortening and centralizing of vowels, the simplification of geminates, the intervocalic voicing of voiceless consonants or assimilations, as in [Sale:ha] for /Sale:ha/, [hanid:ha] for /hanid:ha/ and [xa:ma:at] for /xa:ma:at/.

(4)

\begin{verbatim}
| DEF-cake | DEF-brown | REL on-3SG.F five candle-PL |
| LH | L°H | L H- ^L^H |
\end{verbatim}

The brown cake that has five candles on it,…

\begin{verbatim}
| FUT-1PL=give-3SG.F to-DEF-cat | DEF-green |
| L°H | L°H | L-L°H^H% |
\end{verbatim}

we will give it to the… GREEN cat’
(lit.: ‘to the cat…. the green[one])

Fig. 3: Pitch track of an utterance containing repeated material within the presupposition.

Another experiment of the QUIS corpus was designed to induce narrow focus within a noun phrase and post-focal deaccenting as the experiments by Swerts et al. (2002) and Hellmuth (2005). As already noted, this experiment also showed no deaccenting supporting the results of Hellmuth’s study.
Interestingly, the spontaneous speech corpus, however, does show such cases. The example given in (5) and illustrated in Figure 4 is a classic case of deaccenting given lexical expressions known from West-Germanic languages. The second occurrence of zija:da 'increase, overflow' is part of the NP wazn zija:da 'overweight' which is mentioned in contrast with sinn zija:da 'old age'. Again both NPs are fully FOCAL, but at the same time, the nouns sinn 'age' and wazn 'weight' are in contrast. This contrast is expressed by a focus (not FOCUS) on wazn with a rare incidence of deaccenting, brought about by a flat tonal contour and low intensity (compare the waveform of the two utterances and note that the first one exhibits two incidences of strong amplitudes corresponding to the stressed vowels while the second one lacks such an excursion). We will return to this example below and discuss it in some more detail.

(5) Knee_MG1_12

\text{is-sinn} \text{ il-zija:da} / \text{w-il-wazn} \text{ iz-zija:da}

DEF-age DEF-abundance and-DEF-weight DEF-abundance

‘old age…overweight’…

Fig. 4: Pitch tracks of the two NPs \text{is-sinn il-zija:da} (DEF-age DEF-increase) ‘high age’ (panel a) and \text{w-il-wazn iz-zija:da} (DEF-weight DEF-increase) ‘overweight (high weight)’ (panel 2), the second one uttered shortly after the first one.

Another example in the corpus is the following quotation from a talk where the interviewee (6), the young Egyptian feminist sociologist and blogger Marwa Rakha is being interviewed by, the famous actor Hussein Fahmy. The discussion is about Marwa’s statement that she would make her husband a minister if he wished to become one, as an example of how husband and wife should work side by side to build their future. The witty interviewer of course did not miss the opportunity of remarking on her rather infelicitous wording that she would make her husband a minister, to support the underlying assumption that Marwa’s way of thinking and behaving is odd and inappropriate in the conservative society of Egypt. The immediate pretext of the following utterance was Marwa’s statement that if a potential husband would approach her saying that he would like to be a minister her answer would be the following:

167
(6) Mar_HF_FG3_01
\[ tˤab \ tˤaːla \ - \ a-hutˤ\tˤ maghuːd-i \ ṣala maghuːd-ak / \]
\[ ok \ \text{come.IMP.SG.M} \ 1SG-put \ \text{effort-1SG} \ \text{on} \ \text{effort-2SG.M} \]
'Ok, let’s put my effort and your effort together…
\[ w-a-šmil-ak \ waziːr \]
\[ \text{and-1SG-make-2SG.M} \ \text{minister} \]
\[ LHL \ (!HL) \]
and I’ll MAKE you a minister’

The main prominence here is clearly on the verb and the expression \textit{waziːr} is downtoned as shown by the slightly falling F0 track and the low intensity values especially in the final devoiced syllable (Figure 5a). The interviewer then mockingly comments saying that she was the one that would make her husband a minister without his doing. Upon \textit{Marwa}'s protest that she only meant to support him, he comments on her first utterance repeating it with basically the same intonation pattern, with full a flat intonation on the second word (Figure 5b).

(7) Mar_HF_MM6_01
\[ bāssə \ kilmit \ h-a-šmil-ak \ waziːr \ jašni \]
\[ \text{but} \ \text{word} \ \text{FUT-1SG-make-2SG.M} \ \text{minister} \ \text{DM} \]
\[ LHL- \]
'But the expression ‘I’ll MAKE you a minister, I mean…’

Fig. 5: Twice the same narrow focus utterance spoken by two different speakers in a sequence.

Let us now reconsider the facts discussed so far. As already suggested by all previous studies on the intonation of EA, there is very little deaccenting of full content words and regular deaccenting or accentlessness of function words. But does this mean that information structure has no prosodic reflexes at all? The investigation of the data, especially the spontaneous data suggests that this is not true. In the following, we will try to clarify three things:

i. Where does downtoning occur in relation to information structural categories?

ii. What are the prosodic features employed?

iii. Which prosodic features occur in which positions and is there a constant relationship between a certain prosodic feature and an information structural phenomenon?
In the experiments that were designed to elicit given lexical items in an NP (Swerts et al. 2002, Hellmuth 2005, and part of the QUIS data) the target words always belonged to the FOCUS domain as defined in chapter 1. If an informant looks at a picture and is asked to describe what s/he sees, the referents mentioned will always be in FOCUS. It is here where the languages obviously use different strategies. While English and German deaccent the repeated entity, Italian and EA do not. It could well be that EA speakers do not make a difference between whether a referent or a property has already been mentioned or not as long it is an important part of the information to be transferred. Obviously a red triangle and a red square are two different referents. Furthermore they are answers to two different questions describing two different pictures. Let us now consider the above data where downtoning occurs: In (3) the referents in question also form part of the FOCUS domain, but the task here is slightly different. The informants have one card each and all referents to be named are contained in the same picture and also elicited by the same question. This means that the informant has to decide which of three almost identical things that only differ in one single property should be given to which animal. This constellation makes it more likely that the differing property will receive the highest degree of attention and therefore receive the strongest prominence. We could also say that the context evokes a certain degree of contrast between the differing properties. A possible answer to the question underlying the answers in (3) would also be something like 'All three hares are given juice boxes, the red hare gets the red one, the green hare gets the green one and so forth'. As we will see below, this is an instance of focus, clearly involving a limited set of alternatives. To put the focus to the foreground, surrounding items are therefore backgrounded. In this case the backgrounding or, prosodically speaking, the downtoning applies to the pre-focal element and manifests itself (i) in a flat tonal contour, i.e. a linking contour, that forms one integrated tonal domain with the focal accent and optionally (ii) in lower intensity. Furthermore, (iii) the whole downtoned constituent is not interrupted by a phrase break. A phrase break before the focus accent is frequent, but does not necessarily apply.

The second instance of a prosodic reflex of given and new items was shown in (4). In this case, however, the repeated information (a referent and properties of that referent) are not part of the FOCUS domain, they belong to the presupposed information. We find the following prosodic features associated with it: (i) a horizontal, albeit not totally flat tonal contour, i.e. a linking contour with clear prominences, but no marked drop in intensity and (ii) no phrase breaks inside the constituent and (iii) high speech tempo inducing segmental lenitions and reductions and lower intelligibility. It seems that if material mentioned in the immediate pretext is repeated literally within the presupposition, i.e. in what is common ground between the speaker and the hearer of an utterance, this repetition is done in one go, i.e. without phrase breaks, in a rather careless fashion, i.e. at a high speed and in a reduced pitch range. Whether given material that is part of the presupposition is downtoned or not is up to the speaker's choice and may be influenced by the context as the following example from my own experimental data shows (Figure 6). Figure 6a displays a fully-fledged accent on the verb that is
repeated from the question before while it is downtoned in panel (b) before the narrowly FOCUSED object. This goes to show that in EA, focus may induce downtoning of the preceding accents as well (see below for details on that issue). Figure 6 also illustrates the major role of intensity at the end of an utterance. Note that a low level shape of a final accent does not have to imply downtoning as it can be used as a substitute for a fall (see Section 4.4.2). To distinguish between occurrences of low level pitch with high prominence on one side and low level pitch with low prominence on the other side as the examples in Figure 3 and Figure 6b illustrate, intensity differences are the major cue, confer the major difference in intensity between the last word hala:wa ‘Halawa’ in Figure (6a) and (6b).

So far we have tried to answer the question posed under (i) above to see whether information status has a reflex in prosody. To give an interim summary of the EA facts so far, we have determined that givenness as in the sense of prior mentioning of situational givenness do not necessarily induce downtoning. We may thus conclude that information status itself is a necessary, but no sufficient condition for attenuated articulation. We have seen that repeated constituents are uttered in an attenuated manner if they belong to the presupposition of the utterance, while they are fully accented if they are part of the FOCUS and judged as conveying important information. This explains the results of Hellmuth’s experiment (Hellmuth 2005) that are in contrast to evidence presented here. The other instance of downtoning found in my data was in the vicinity of a focus accent expressing contrast. In both cases, however, downtoning is not obligatory, and the degree of downtoning – or even occasional deaccenting – is up to the speaker. The main conclusion to be drawn from the evidence presented above is that EA, being a non-deaccenting language, uses downtoning and generally attenuated articulation as an equivalent of deaccenting in English to express focus-background structure, especially in cases of contrast.

In the following paragraphs we will be looking at question (ii) concerning the prosodic realization of focus, background and contrast. We have already noted that examples (5) and (6) are different from the cases in (3) and (4) in several respects. First, the downtoned expression is final or post-focal.
Remember that the studies on Italian (Avesani & Vayra 2005, Farnetani & Zmarich 1997, D’Imperio 1997, 2001) also found deaccenting and downtoning in post-focal position. In the second example, the object \textit{wazi:r} is downtoned after a strong accent on the verb. If we look at the context, it is clear that the verb here is a \textit{focus of interest}, even though the FOCUS domain is the whole VP as the speaker wants to inform us about her making a minister (out of her husband) and not only about ‘making’. It is equally true, however, that it is the act of \textit{making} a minister that is the focus of interest here and our speaker transmits this information by a strong accent on the verb and a downtoned object, which itself is kind of a secondary topic of the utterance (cf. chapter 5). Although it is not the givenness of the object alone that induces downtoning, some measure of givenness is indeed a necessary prerequisite. An articulation with \textit{wazi:r} being as prominent as the verb would be more appropriate if the speaker did not know beforehand that her husband wants to become a minister. Crucially, however, full prominence of both accents does not necessarily \textit{imply} newness of the expressed concepts. Thus, the marked difference between EA and a language such as English is that downtoning in EA is not obligatory and that givenness alone is not a sufficient criterion for it to happen. Consider the following example from natural conversation. In a jointly told story, one sister (speaker F3) had started talking about a situation when the father made his children drink lemon juice, \textit{lamu:n}, to make them vomit because he suspected that they had eaten something poisonous. The other sister (speaker F0) interrupts reminding her sister that something important had happened before that incidence.

(8) Fagr_01_F3_F0_

A: \begin{verbatim}
gajb lamu:n
bring.PTCP.SG.M lemon
'He was bringing lemon juice'
\end{verbatim}

B: \begin{verbatim}
la ʔabl il-lamu:n
no before DEF-lemon
'No, before the lemon juice'
\end{verbatim}

Figure 7 shows a strong accent on the given item \textit{lamu:n} ‘lemon(juice)’. Note that the speaker might as well have opted for downtoning this element, thereby foregrounding the denotation of the preposition ʔabl ‘before’ as she would presumably have done in a context of outright contrast, for example in contradiction to \textit{ba:d il-lamu:n} ‘after the lemon(juice)’. Importantly, the strong second accent here is an instance of what Bolinger calls \textit{accent of power} or \textit{climax} to give more emphasis to the whole expression (Bolinger 1986: 75ff.). At the same time we find an accent on the preposition which mostly goes unaccented otherwise.
Contrary to that, example (5) is in fact quite similar to the targets in the experiments on given constituents within an NP (Swerts et al. 2002, Hellmuth 2005). It consists of a combination of two nouns, the second of which is used like an attributive adjective. The construction is tightly knit, functioning semantically as a quasi-compound equivalent to its English translation ‘overweight’. This semantic integration is prosodically reflected in prosodic integration (cf. Section 4.5.5). This type of integration will be called integration II in Section 4.5.5. Such integrated contours are used frequently on annexation constructions in EA. We have already determined, however, that the repetition of a lexical expression alone is not enough to induce downtoning. A closer look at the example shows that again we are dealing with an instance of focus of interest. The speaker, a physician who is being interviewed in the morning talk show sˈabaː h iƚ-xeːr ja masˈr 'Good morning, Egypt' is talking about the various risk factors that may lead to knee arthritis, first mentioning the factor of 'sex' (ʔaːn tʊbʔa səjʃida 'if it is a lady'), then 'old age' (sɪnː zɪjəːda) and finally 'overweight' (wazn zɪjəːda). Two factors here may be responsible for the perceived intonation, the first of which is finality that induces downstep and occasionally even downtoning of the second accent in an integrated phrase. It has frequently been noted in the literature that early peaks enhance the impression of finality (Swerts et al. 1994; Wichmann et al. 2000, Rastegar-El Zarka 1997 for EA). It is more likely, though, that in this case, where we have a marked drop in prominence, or even deaccenting of the second item, this is due to emphasis or a focus of interest on the word wazn 'weight' to draw the listeners attention to the fact that they cannot avoid getting older, but avoiding obesity is in their own hands, as the doctor tells his audience in the following sentences. How then is post-focal given material expressed prosodically? First, we find that it is necessarily integrated rhythmically and tonally, i.e. it is included in the same phrase as the focus and it is tonally either partially or totally downstepped (a falling contour with no peak or a flat contour). At the same time, intensity is necessarily reduced in that position. In the next subsection we will be looking at the relation between focus and FOCUS and TOPIC as identified in the introductory chapter.
Descriptive work on EA suggested that narrow focus may be expressed in situ like in English, using prosodic strategies (Gary & Gamal-Eldin 1982; Mitchell 1993). On the other hand, Hellmuth (2006a, 2009) claims that information focus shows no prosodic reflexes in EA, and contrastive focus is characterized by hyperarticulation, which she views as non-phonological, i.e. non-distinctive, according to the theoretical tenets of Standard AM phonology. Norlin’s (1989) instrumental pilot study investigated the F0 reflexes in different focus constituents (subject, verb, object), which correlated with initial, medial and final position in the sentence. Norlin reports pitch range expansion on the focused item together with post-focal compressed pitch range as compared to the neutral declarative.

Probably the clearest case of contrastive focus is corrective focus. We will thus look at corrective focus cases first. For that purpose, I will predominantly rely on the experimental data, as they obviously provide more frequent instances of this construction than spontaneous data. The QUIS corpus only contains a limited number of cases. However, especially one experiment offers a situation where the correction of only part of an NP is elicited, for example in a noun-adjective combination as in (9). The informants were shown pictures and asked specific questions. The elicited answers were supposed to be either confirmative or corrective and what was to be corrected was not always a whole constituent, but in some cases only a part of it. In (9a) the noun is corrected, while in (9b) it is the adjective that is corrected. Note that in EA, the adjective always follows the noun. From what we have seen in examples (5) and (6) and Figures (3) and (4) we may expect a marked focus accent on the noun in (9a) and on the adjective in (9b). If this expectation is borne out we may say that at least corrective focus may be prosodically expressed in situ in EA. Note that phrasal syncope also fails in the focused item - but not as a rule.

(9)

Q: ‘Is the girl taking the SMALL glass?’

A: $\text{la? il-bint}$ $\text{-b-ta-xud}$ $\text{-ik-kubba:ja}$ $\text{K-KIBI:RA}$ no DEF-girl IND-3SG.F-take DEF-glass DEF-big 
linking contour LH $^\circ$L^HL

b. Q: ‘Is the girl taking the big BOTTLE?’

A: $\text{la? il-binta}$ $\text{-b-ta-xud}$ $\text{IL-KUBBA:JA}$ $\text{k-kibi:ra}$ no (DEF-girl IND-3SG.F-take DEF-glass DEF-big LH $^\circ$H- L $^\circ$H$^\circ$L-
Fig. 8: Two utterances with corrective focus within an NP. Panel (5a) shows a corrected adjective (ik-kubba:ja k-KIBI:RA 'the BIG glass') while panel (5b) shows a corrected noun (IK-KUBBA:JA k-kibi:ra 'the big GLASS').

As the examples show, there actually is a marked difference between the constructions. As we can see in the left panel, the strongest excursion and highest scaling pertains to the final accent, which, as expected, is on the corrected adjective. As regards the pitch of the remaining parts, it is rather flat and only rising on the word kubba:ja 'glass' leading to the focal accent on the adjective kibi:ra 'big'. Furthermore, the articulation of the focal item is prolonged - cf. the significantly longer duration of the word in the left panel, also in relation to the adjacent noun. Note that the intensity values are not especially revealing in this examples. This is due to the fact that there is a natural intensity drop on the final item in falling intonation patterns, which in the present case is clearly counteracted. However, this counteraction does not result in spectacularly high intensity values and thus may easily escape visual inspection. What can be seen quite clearly, however, is that the intensity drops on the post-focal element while it is not lower in the pre-focal given parts. These facts are in line with the observation made in the previous section on the articulation of givenness. There are however cases where the intensity is also clearly reduced in pre-focal material - a fact that also goes hand in hand with the almost complete compression of pitch range which yields a totally flat contour as shown in Figure (9).

It is not at all easy in such cases to determine whether the lexical accents are completely suppressed or only reduced to a monotonal shape. But in the approach taken here, this is of no importance as long as the prominence relations are clear.
Fig. 9: Utterance with a focus on the final lexical item bint ‘girl’ in a correction context: the boy is holding (not the hand of the boy, but) the hand of the GIRL (ʔi:d il-BINT).

An interesting observation can be made in Figure 8b where we can see that the focus does not have a higher pitch than other peaks in the utterance. In fact, the pitch peak is approximately equivalent to the peaks of the pre-focal accents, nevertheless the focal accent is unambiguously perceived as the strongest one in the utterance. This illustrates that accent scaling is not the only way of making accents more or less prominent. What is of more importance is excursion size (cf. Gussenhoven 2002, 2004). A very important joint feature of the two focal accents is its shape. We will deal with this issue in 4.3.2. By way of conclusion, we may note that the prosodic realization shown so far is not the only possibility. For one thing, the spontaneous corpus shows cases where corrections are made without using a marked focal accent, suggesting that corrective focus is a kind of emphasis that may be employed or not depending on the speaker’s choice.

4.3.1.1 Experimental evidence for focus within an NP

Neither the QUIS corpus nor the spontaneous corpus, however, contain enough data to allow a quantitative analysis. For this purpose we turn to an experiment I conducted to elicit contrastive (corrective) focus on one of the two parts of an annexation construction that constitutes one NP. The design was essentially the same as the one of the QUIS experiments discussed above, only that I used written sentences instead of pictures. This was due to fact that the target words were controlled for their segmental structure, i.e. the target syllable contained only sonorant consonants (predominantly liquids and nasals). Additionally, the syllable structures of the targets were varied. This could only be achieved by using proper names of persons and places in Cairo. The sentences were presented in a powerpoint presentation with an inserted audio file. The participants could go through the sheets at their own pace. When clicking on the icon for the audio file they heard the question they were supposed to answer. The question elicited either the correction of the first or the second part of the annexation construction. Six informants took part in the experiments. Two of them were born and raised in Alexandria and four in Cairo. Their ages varied between 28 and 80. The two elderly
informants, a married couple (70 and 80 years old) were first included in the experiment because they had never lived abroad, nor did they have a good command of any foreign language. It turned out, however, that they obviously could not fully concentrate on the task. Especially the male informant produced some adventurous intonation patterns. As the following count is not a quantitative analysis in a strict sense, I decided to include those utterances that seemed appropriate to me and analysed them separately. The data set consisted of 10 sentences with target words in first position (KF1) and 10 sentences with target words in second position (KF2), spoken by 6 speakers in two repetitions, i.e. 10x6x2=120 utterances or 40 utterances per speaker. Two examples of the target words that were used in the experiment are given in (10), the lexically stressed syllable is marked in capital letters and the expected focus word is indicated in boldface.

(10)

a. MANzil lamLU:M (KF1) and MANzil lamLU:M (KF2)
   house (of) Lamloum
   'Lamloum's house'

b. biNA:jit haLA:wa (KF1) and biNA:jit haLA:wa (KF2)
   building (of) Halawa
   'Halawa building'

c. NA:di z-zaMA:lik (KF1) and NA:di z-zaMA:lik (KF2)
   club (of) Zamalik
   'Zamalik club'

The decision was taken on an auditory basis by the author in cooperation with a native speaker of EA who is not a linguist and who did not take part in the experiment. The native speaker was told that he should determine whether the stronger prominence was on the first or on the second word of the construction and also whether the utterance was an adequate answer to the question. The task turned out to be easy for him. The results are shown in Table 1, which contains only the numbers for five speakers. The productions of the elderly male informant were excluded as they were all pronounced with a stronger accent on word two and in a quite mechanical way, they sounded very unnatural. Listening to them gives you the impression that the speaker was reading the examples from the screen without paying attention to the context at all. The elderly female informant (F2, born and raised in Alexandria, lived most of her life in Cairo) performed better, but she also had some difficulty concentrating. Contrary to the productions of the old man, many of her utterances give evidence of a strategy to focus one or the other part of the constructions, only that it is not always the right part she focuses on. As the graph in Table 1 shows, the total of her utterances that were counted as 'correct' amounts to 56% of the data (18 out of 32 utterances). The remaining four speakers pattern two by two. Speaker F0 (female, Alexandria) produced 26 sentences (65%) that confirmed the hypothesis. Interestingly, most utterances that were judged as infelicitous belonged to the category of KF2. This is unexpected in the light of the assumption that EA phrases are always right dominant with the main prominence on the end of a prosodic domain (Hellmuth 2006b, et passim). But if we look at the way
many of her KF2-renditions were produced, this fact finds a plausible explanation. While three utterances showed a pronounced focus accent with a high peak, a *pointed-hat* accent, on the final word, the other five exhibited downstep on the final word. Nevertheless the accent is mostly strong enough in these cases to be counted as the more prominent one. There is a strong tendency to have downstep on the final accent in EA phrases, especially in what might be called 'neutral' declaratives. Downstep gives the statement an assertive 'matter-of-fact' overtone and increases finality, especially when the fall of the final accent is integrated in the fall of the prefinal accent (cf. chapter 2.2.4.3). As prominence differences are a rather elusive matter, it is sometimes tricky to determine the stronger accent in these cases. A similar result was obtained for speaker M1 (male, Cairo). His infelicitous renditions are a little more balanced, but similar to speaker F0 and contrary to speakers F2 and M2 he also performed much better in the KF1 cases (85%). He showed a strong tendency to use a redundancy contour (a fall-rise-fall) that stretches out over two accent domains, cf. Figure 10d) which may also have prevented a felicitous production in some cases. Finally, speakers F1 and M0 scored highest in their performance, both produced a 100% felicitous utterances in the KF1 condition and slightly less in the KF2 condition. We may thus summarize that EA speakers have strategies to mark contrastive (at least corrective) focus prosodically (in total 85% felicitous renditions). All speakers except speaker M2 shifted the main prominence within the data. With the exception of speaker F1 who always used the same strategy for each condition, i.e. strong closing accent on the first word (mostly) with deaccentuation of the second word for KF1 (Figure 10a) and (almost always) a linking contour on the first word followed by a pointed hat accent on the second word in KF2 (Figure 10b), the speakers varied their intonation patterns. The most typical patterns are depicted in Figure 10 below.

![Figure 10](image)

**Tab. 1:** Number of felicitous sentences produced by 5 speakers of EA (3 female and 2 male) out of a total of 20 sentences per condition (except F1: 19 for KF1 and F2 16 for KF1 and 16 for KF2).

**Tab. 2:** Percentage of felicitous sentences according to position, produced by 5 speakers (3 female and 2 male).
Before we look at the different realization and try to determine which factors might have lead to the auditory effect of prominence differences, I would like to mention one important common feature of all productions. All renditions of the annexation (whether KF1 or KF2) were uttered in one phrase. In fact, the annexation construction was chosen because it has been suggested in the literature that an upcoming boundary may influence the tonal shape of the accent (Prieto et al. 1995). Although I do not believe that phrasing necessarily changes accent shape, I was careful to choose a construction that will almost certainly not be split up into phrases. This should also convince those who believe in the syntactic definition of phrases that whatever the tonal variation is that occurs, it will certainly not have been induced by a boundary. The second thing to mention is that most of the examples display also a type of tonal integration (cf. chapter 2.2.4.4). In Figure 10a, the final accent is totally subordinated to the prefinal one - and even deaccented in many of the productions, especially those by speaker F1 and to a somewhat lesser extent also by speaker M1. While in Figure 10a a linking tone or total downstep is used for the second word, the examples in (Fig. 10c) and (Fig. 10e) display downtoning by two types of partial downstep. Contrary to these, the downstep contour in Figure 10f is not an instance of downtoning, but on the contrary, the example is a case of condition KF2 and the second accent is in fact the stronger one. To differentiate between the two contours, I will call the downtoning accent downglide as opposed to the downstep of the strong accent. Closer inspection reveals two major differences between the KF2 instance in (Fig. 10f) and the KF1 instance in (Fig. 10e), produced by the same speaker F0. While the KF1 example shows a drop in intensity and almost complete tonal integration in the fall of the first accent, the KF2 example displays high intensity on the second accent and a marked terraced contour. The elbow and subsequent rapid fall on the stressed syllable of word 2 are deemed responsible for the high prominence.
Fig. 10: Eight intonation patterns displaying different focus conditions within an annexation construction. The pitch tracks in panels (a), (c) and (e) display a KF1 condition with focus on the first element and downtoning of the second while the pitch tracks in (b), (d), (f), (g) and (h) display a KF2 condition with focus on the second element.

The reverse type of tonal integration is displayed in (10b) where the linking contour is associated with word 1 with following upstep of the final accent which in turns forms a pointed hat. Note that the intensity difference between the final accent in (10a) and in (10b) is also striking. The other two types of tonal integration are exemplified by the patterns in (10d) and (10h). (10d) is an instance of the redundancy contour, a fall-rise-fall over two accents, ĶL'H(L), the first accent is an inverted accent (Bolinger’s C profile) and the F0 frequently reaches its lowest point within the stressed syllable (..L*..) until the maximum is reached in the vicinity of the following stressed syllable. Inverted accents are
frequently used on given items (Jackendoff 1972; Brazil 1975, 1997; Gussenhoven 1983; Bolinger 1989; Büring 1997), they have been called “anti-assertive” by Bolinger (1958: 147). In the present framework, the inverted accent is a special variant of the leading contour\(^7\) indicating that the semantic contents of the item it is associated with constitutes some kind of starting point for what follows. The final type of integration is the well-known hat pattern which I have called total linking in chapter 2. The hat pattern, familiar from a certain type of topic-comment sentences in West-Germanic languages (Bolinger 1986, t’Hart et al. 1990) is used in EA to link up all different kinds of constituents, for instance the two parts of an annexation as in the above example. In sum, there are two types of integration, one that subordinates one of the accents under the scope of the other one as in (Fig. 10a,b,c and e) and another one that involves two full accents (Fig. 10d, f and h). Finally, there is only one pattern that does not display tonal integration. This pattern (Fig. 10g) occurred only twice in the data set and was produced by the same speaker F0. In the case of (Fig. 10g) the final accent is produced as a pointed hat which lends it more prominence, probably because the default neutral pattern involves downstep of the final accent. Table 3 displays the frequency of the different patterns by focus condition and Table 4 shows the distribution of the patterns across speakers.

\[^7\] As an aside I would like to point out that this is a good example against the categorical distinction between different accent types as they are commonly used in Standard AM frameworks today, such as LH* as opposed to L*H. The main characteristic of the redundancy contour is a fall to a very low level and subsequent rise to the H tone of the following accent. It is not crucial that the L be aligned with the stressed syllable, the alignment of this L tone only has to be somewhere within the word it is associated with.
To summarize, EA speakers use strategies to mark corrective focus similar to speakers of English. Although deaccenting after focus (Fig. 10a) is not the preferred option, the post-focal part in EA utterances is drastically reduced in prominence and mostly characterized by a flat tonal contour. The peak of the prominent accent itself is mostly scaled higher than the preceding one, which may also be downtoned in the same way as the post-focal one, although complete deaccentuation is even less likely (Fig. 10b) in that position. This is the clearest way of marking prefinal contrastive focus. The other option is a downstepped accent in post-focal position with an accompanying drop in intensity. This downglide involves the tonal integration of the accent into the falling part of the focus accent. This type of integration is the default option for function words which are unambiguously accentless. A downgliding contour, however, may retain a certain prominence. These facts clearly point to the gradience of accentual phenomena, not only of pitch range and intensity variation, but also of accent type variation, which is normally assumed to be a robust categorical distinction.

The most frequent type of final focus accent is an upstepped accent, frequently with a pointed hat, but downstep with accompanying high intensity, whether after a low tone (Fig. 10f) or after a high tone, i.e. as the second part of the hat pattern (Fig. 10h), is also common. In sum, we may observe that post-focal material is typically characterized by a linking contour (L- or L+) or by a downglide (!H’L or L%), whereas final focus most frequently involves upstep ((L)^HL), but may also be expressed by downstep (!HL), as long as this tonal contour is accompanied by high intensity.

In all the examples presented in the present section, the focus accent falls on a narrow very narrow domain, in these cases a single word to make foreground the denotation of that expression. In principle, corrections may also pertain to a single sound or morpheme, called expression focus by Krifka (2007). As suggested in Chapter 1, this type of focus is not congruent with the FOCUS of the whole proposition, as defined in a Lambrechtian sense. Rather, it affects only part of the FOCUS, namely that part that is the semantically most relevant or most informative part of the domain. All examples discussed in this section where in some sense explicitly contrastive, being either corrections or standing in syntagmatic relation to another alternative. That is, in the cases at hand the alternative...
set was limited to two possibilities only. Recall that Hellmuth's conclusion from her experiment reported in Hellmuth (2010b) was that only contrastive focus is marked by expansion of pitch range of the focused item with simultaneous compression of post-focal material. It remains to be seen if the present result presented above are only valid for contrastive focus cases or if information focus may also be articulated by a special 'focus accent'. This question will be the topic of Section 4.5.

In Chapter 2 I have related the scaling and alignment properties of an accent to the effort code. It is a truism that the effort code is relevant to the articulation of focus, hence the widespread conflation of focus and highlighting. Focus marking in EA makes no exception in this respect. Whatever the tonal articulation of focus may be - and we have seen that it may vary considerably in EA - what all articulations have in common is higher prominence of the focus part and less prominence of the backgrounded part. The autosegmental-metrical theory makes provision for the representation of prominence relations in the metrical component. In this sense, the present approach concurs with the assumption expressed by Ladd (1996: 221) that "relative metrical strength is the essential signal to focus", albeit without assuming that metrical structure is rule-based. The following representation of EA utterances tries to account for the relational nature of prominence by assuming maximally three layers of the grid. The first layer represents the syllables of the linguistic material and the second layer pertains to lexical accents that are rhythmically induced, the third level represents a focus accent. Taking as an example data from the contrastive focus experiment described in section 4.3.1.1, I suggest that downtoning with retained lower prominence and deaccenting be represented by alike.

(11)

\[
\begin{array}{cccccc}
  x & (\text{focus}) & x & (\text{focus}) \\
  x & (\text{word}) & x & (\text{word}) \\
  x & x & x & x & x & (\text{syllable}) \\
  \text{MALga? il-musinNI:N} & x & x & x & x & x & (\text{syllable}) \\
  \text{asylum} & \text{DEF-old.people} & \text{asylum} & \text{DEF-old.people} \\
  \text{'old people's HOME} & \text{'old people's HOME} \\
\end{array}
\]

4.4. The prosody of pragmatic relations

In this section we will be dealing with the notion of FOCUS as defined in chapter 1. There it has been argued that we need to differentiate between focus and FOCUS for various reasons. One such reason is the fact that different formal characteristics can be attributed to the two notions. Of course, the two notions are not totally independent. FOCUS, being the main information of a sentence, is expected to involve prominence, which in turn has been determined as the main characteristic of focus. We will look at the relationship between the two notions in Section 4.5.2 where we will also be dealing with the relationship between focus and the other pragmatic relation identified, namely TOPIC. But before we turn to that intricate interplay we first determine the main prosodic correlates of FOCUS and
TOPIC in EA. I have already discussed the literature on that topic in chapter 3. In the following, I will thus only refer to the main assumptions in the literature the present approach is based on.

The suggestion made in El Zarka (2011a) that EA intonation can be described by the assumption of three basic tonal types: leading, closing and linking tones and contours that have been shown to correlate with sentence mood, among other things, as shown in chapter 2, will be related here to the information structural notions of pragmatic relations: TOPIC and FOCUS. This proposal is in fact not new. It is in line with proposals made in the literature on intonation as early as the mid of the preceding century when the correlation between rising F0 and topic or theme on the one hand and falling F0 and comment or rhyme on the other was been noted in work dealing with the communicative meaning of intonation (Navarro Tomás 1944 [1974]; Bolinger 1958; Jackendoff 1972; Brazil 1975). The application of the three tones identified in this model closely parallels the distinction made by Brazil (1975, 1997) between “referring” (leading), “proclaiming” (closing) and “neutral” (linking) tones. Brazil goes a long way in arguing for a correlation between these tones with what is called TOPIC and FOCUS here. He shows that referring tone characterizes presupposed constituents while a proclaiming tone is associated with the rhematic part of the sentence. The suggestion is also in line with Bolinger’s notions of thematic B accent vs. rhematic A accent (Bolinger 1958, 1986). Gussenhoven's (1984) distinction between the meanings of “selection”, ascribed to the fall-rise accent (H*LH), and “addition”, ascribed to the falling accent (H*L), was also heavily influenced by Brazil's work (Gussenhoven 1984: 215).

This basic distinction has appeared over and again in different works on intonation (e.g. Uhmann 1991, Féry 1993) and in studies devoted solely to the interface of information structure and prosody (Steedman 1991, 2000; Büring 1997, 2003; Hirst & Di Cristo 1998). Most of the cited work deals with the situation in West-Germanic languages, especially English and German. In these languages, the dichotomy of rising theme or topic and falling rhyme is manifested in the hat pattern. Hirst & Di Cristo (1998: 20) note that “[i]n most languages the falling nucleus is generally prepared by a rising pitch occurring on the first stressed syllable of the unit” which they refer to as pitch “onset”. Together with the falling "offset" this rise results in the hat pattern if no other accents intervene. As Bolinger notes, this bipartite articulation of a sentence invokes the association of a question-answer pair:

There’s a part that lays the ground-work, that asks the question, that relates to what we already know or can guess, and a part that adds the figure to the ground, that answers the question, that supplies what was not already known. The first part is called the THEME and the second part the RHIME. (Bolinger 1986: 47, emphasis in the orig.).

In terms of the pragmatic relations defined in chapter 1, we may correlate high/rising pitch with topicality in the sense of aboutness and low/falling pitch with FOCALITY in the sense of assertion. Note that we are neither concerned here with the position of the accent for the moment, nor with its height but with its dynamics, i.e. its shape.
4.4.1 The prosody of pragmatic relations in EA

It has repeatedly been pointed out that in EA there is an accent on almost every content word. This obviates or at least reduces the possibility of the occurrence of the hat pattern on a sequence that includes more than two accents. Nevertheless, the suggested basic rising-falling configuration can be observed in EA as well.

I have argued in Section 2.2.5.1 that the leading-property may pertain to whole tunes or contours as well as to individual accents. Likewise, topical expression may either consist of only one lexical item associated with one accent. The least ambiguous way to mark a TOPIC expression is by a rising contour with a subsequent break (see Figure 11).

A break, however, is not necessary to achieve a perceptionally unambiguous rise, even in cases where the TOPIC expression bears only one accent. In such cases the peak of the TOPIC accent may be aligned later than in the default case. Figure 11a displays the entire sentence whose subject Hany (proper name) corresponds to the topic referent (the alignment of the tones is illustrated in Chapter 2, Figure 21). The accent here is not realized with a peak at the end of the heavy first syllable of /haani/ as in the default case, but aligns with the final vowel of the whole constituent. Alternatively, the pitch may stay at approximately the same level after the peak has been reached and form a high plateau until the end of the topic expression (Figure 11b).

In the latter tune, the TOPIC and FOCUS expressions are conjoined in a small hat pattern, which in EA is quite common when bridging two adjacent content words, but uncommon in longer stretches. Longer TOPIC expressions may also be associated with a rising sequence of accents, i.e. with an upstepping contour (Figure 12). The resulting rising trend line, comparable to the rising “grid” of the
Lund model (Gårding 1983), gives the overall impression of a rise and thus is apt to convey the leading function.

Interestingly, the rise in the example in Figure (12b) is not only associated with the argument expressing the topical referent Nagwa in a topic-comment sentence meaning ‘Nagwa has gone to the old people’s home’ but is extended to cover the predicator, thus conveying a subtle difference in the information structure of the sentence, which was a response to the question ‘Where is Nagwa?’ In this rendition the action expressed by the verb is to be understood as presupposed (cf. Chapter 5). Imagine someone asking the above question upon entering a room where he expects Nagwa to be present, but does not see her. This implies that Nagwa must have gone somewhere. Hence the mentioning ‘go’ which is a semantically weak verb, unlike ‘travel’ or ‘run’, does not have any information value itself, the only new and important information here is the place Nagwa has gone to. I thus consider the rhematic FOCUS here to be not only about Nagwa, but rather about where Nagwa has gone to, which is why the speaker chose to extend the prosodic topicality marking to the verbal constituent. Another possibility would be to mark Nagwa alone as the topic and treat ‘has gone’ as backgrounded and only as a link to the focal part in which case the expression would be given a linking tone.

TOPICS need not, however, be associated with rising prosody. A common realization is a flat tonal contour that kind of links the FOCAL part of an utterance to the preceding one. I call this type of TOPIC linking TOPIC. Cases will be discussed in Chapter 5. Unlike a leading TOPIC, a linking TOPIC is not separated from the FOCUS by a break. There are two types of linking TOPICS, a downtoned type with little or virtually no accentual excursions as in example 4 (Figure 3) above and one with pronounced accents and a horizontal trendline as in Figure (14b) below.

Regarding FOCAL intonation, I assume the standard mode of expression to be the falling contour, and again this may involve a falling trendline with default accents (default alignment of the H tone and
partial linking) or closing accents, be it only one or a series of accents. A number of examples have been illustrated so far in this chapter. Figure 12 displays a falling trendline covering the whole FOCUS constituent; the respective contours consist of three accents (Fig. 12a) or two accents (Fig. 12b). Singleton closing accents are shown in Figures 7, 8 and 9. Many of the accents occurring in the data so far exhibit pointed hats in narrow FOCI, but downstepped accents are very common, especially in final position of broad FOCUS (see the discussion in Section 4.3.1.1). In these cases the final downstepped accent results in an overall falling (partial downstep) (Fig. 12a) or low (total downstep) gesture (Fig. 11a). Downstepped final accents often exhibit a sharp fall in the last syllable, which adds to the prominence of the word, similar to the sharp rise in the prefinal accents. The impression of finality and assertiveness may be enhanced by using an early peak, either within the lexical item by alignment before the accented syllable and sometimes at the beginning of the whole constituent, also observed by Rastegar-El Zarka (1997) in Egyptian MSA, or by a gradual fall from the prefinal accented syllable through the final content word to the end. One main characteristic of the closing tone are the alignment and scaling properties of the final L tone, which is often comparably low relative to the second low target in a leading accent or in a default one. Another important characteristic is the earlier alignment. Thus the L tone after the peak is not linked to the following lexical item, but rather associated with the focused lexical item itself by being aligned with its end or even earlier, especially in narrow FOCUS cases. Example 12 illustrated in Figure 13, which was also presented in Figure 38 (Ch. 2), shows the difference between a TOPICAL (12a) and a FOCAL (12b) long subject in syntactically identical sentences. Note that there is no noticeable difference in pitch height – neither of the peak nor of the trough – between the two conditions, the sole difference being the choice of the accent. While the second low target of the first accent in TOPIC condition is aligned with the beginning of the following content word kama:l - which is the default case - the second accent is followed by a break and is therefore only rising (Fig. 13a). In the FOCUS condition, the subject is associated with two closing accent, in the first of which the closing L is aligned very early, shortly before the end of the word it is associated with (miʃza ‘goat’). This early alignment results in a steep fall, which significantly enhances the prominence of the accent. The second FOCAL accent is also a closing one - a fully-fledged rise-fall ending at the end of the proper name kama:l. Contrary to the first accent, this last accent is also followed by a break. This analysis is supported by the occurrence of phrasal high-vowel syncope in bitaːʕit (possessive marker) yielding il-miʃza btaʕit kama:l ‘Kamal’s goat’. This fact goes to show that while it is the terminal of the accent that contains the intonational meaning, this terminal is not brought about by a boundary. The case of the subject NP in (12) is especially revealing as it involves two content words with a quite long grammatical element between them. Thus all the assumptions made in the present model are born out in one single example: (i) the second L of default accent is aligned with the beginning of the following content word with tonal integration of the functional element into the fall of the accent as expected, (ii) the deviation from default in the closing accent type with a full rise-fall covering the lexical item the accent is associated
with, and the early alignment of the closing L, in which case the grammatical material is assigned a linking tone, (iii) the difference in intonation between TOPICAL and FOCAL constituents, and (iv) the downtoning of the presupposed constituents after narrow FOCUS in (Fig. 13b), i.e. its realization in a highly compressed pitch range. This is a case made for illustrative purposes and it unsurprisingly occurred in the read speech data. Maybe the more frequent rendition would involve tonal integration of NP constituents, especially in the FOCUS NP, which in spontaneous speech would most probably involve a hat pattern.

(12)

a. il-mīṣza btaṣit kama:l

L H

b. il-mīṣza btaṣit kama:l

L H L

Fig. 13: The long subject il-mīṣza btaṣit kama:l (the-goat belonging.to Kamal) ‘Kamal’s goat’ as a topic expression (panel a) and as a focus expression (panel b) (stressed syllables are capitalized).

Needless to say that TOPICAL and FOCAL constituents, especially if they are emphatic, are frequently expressed within a separate phrase, but this is not inevitably so. Importantly, the phrase boundary does not contribute to the shape of the accent, it only may enhance it by making a rise higher and a fall lower. The example in (13a) (Fig. 14a) shows a sentence with a topicalized object il-ʔakla di ‘this dish’ and narrow FOCUS of the subject te:ta ‘granny’ with a break after the TOPIC constituent, but no break after the FOCUS constituent.⁸ The accent associated with te:ta is a closing one with the last item ʕamalitha ‘she prepared it’ produced in a compressed pitch range (cf. Section 4.3.1).

---

⁸ Of course, phrasing is gradient and there are perceptionally different degrees of boundaries. Whenever I claim that there is no boundary, this not only rests on impressionistic, but also on acoustic and segmental cues such as the missing lengthening of a final segment or spectral characteristics and deletion or insertion of segments, cf. the phonetic rendition of te:ta ʔilli as [teːtɛlːi] as realized here versus to [teːtɛʔɛlːi] with a break between the noun te:ta and the relative marker ʔilli.
Another manifestation of the leading and closing contours is illustrated by the sentence in (13b) (Fig. 14b). This utterance has a low falling accent on the final adjunct imba:riħ ‘yesterday’, which is the FOCUS, while the topical constituent ma:lik id-dukka:n illi gambina “the proprietor of the shop next to us” is marked by a high level trend line, i.e. by the suspension of downdrift. This strategy was also observed by Norlin (1989) to be used for questions in EA and is employed in other languages with a high density of pitch accents such as Danish (Grønnum 1998) and Swedish (Gårding 1998). It seems that level or rising trend lines in languages with high pitch accent distribution are the logical equivalent of level or rising contours on unaccented material in languages with lower pitch accent distribution like English or German. Finally, what makes the sentence in (13b) not sound like an incidence of argument FOCUS (‘the proprietor next door’) is the lack of a closing accent on the end of the object phrase gambina ‘next to us’. The closing accent is rather associated with the adjunct imba:riħ ‘yesterday’.

\[(13)\]

**(a) QUIS corpus, translation task**

\[il-ʔakla di / [TE:TA]FOC lli ʕamal-it-ja\]

DEF-dish-DEM.SG.F granny REL make-3SG.F-3SG.F

‘This dish was made by GRANNY’

**(b) read speech data**

\[zur-na ma:lik id-dukka:n illi gambi-na]TOP [IMBA:RH]FOC visit-1PL proprietor DEF-shop REL beside-1PL yesterday

‘We visited the proprietor of the shops next door YESTERDAY.’

Fig. 14: Panel (a) shows a left-dislocated topic followed by a clause with subject focus il-ʔakla di / TE:TA lli ʕamalitha “this dish, GRANNY is the one who made it”; panel (b) shows focus on the adjunct imba:riħ “yesterday” “we visited the proprietor of the shop next to us yesterday.” (focus indicated by capital letters)

While the rise and the fall respectively can be viewed as prototypical realizations of leading and closing tones, the examples in Figure (14b) and Figure (11) show that high level and low level may be used as a substitute for rise and fall (more examples will be given in Chapter 5).
4.4.1.1 Experimental evidence for the tonal encoding of pragmatic relations

To test the hypothesis that TOPIC and FOCUS expressions exhibit different tonal prosodies, I conducted a production experiment with pairs of morpho-lexically and syntactically identical sentences with the same target word in TOPIC and FOCUS position, respectively. The information structural function was controlled by a preceding question, an example is given in (15) below. The sentences were presented on a computer screen in Arabic script, following the usual standards for the transcription of spoken EA⁹ (e.g. the use of the letters for semivowels /w/ and /j/ in some positions where they would not normally appear in written Arabic) to elicit the correct pronunciation. It has to be noted that none of the speakers had any difficulty in reading the ‘dialect’. Not a single utterance had to be discarded because it contained a word pronounced as in the written Standard. The participants were asked to read aloud the answers to the questions in a contextually appropriate manner. The questions that had been pre-recorded by the instructor were presented auditorily. The participants listened to the question after clicking on an icon to open the sound file. They were told that they could listen to the question more than once, if they liked. In Section 4.3.1.1 I have presented the data that were used for the first experiment. As noted, these data served as filler utterances for the data of this experiment. All utterances were presented in a pseudo-randomized order.

The data contained two sets with segmentally controlled target words. In each set, the target words varied with respect to type and position of the target syllable following the design of Hellmuth & El Zarka (2007). Target stressed syllables were of four types: light CV, heavy CVC and CVV, and superheavy CVVC; their position within the target word was antepenultimate (CV), penultimate (CV, CVC, CVV) or final (CVVC), which corresponded to initial and final position, respectively. All target syllables contained the vowel /a/ in order to eliminate effects of intrinsic F0, with sonorant consonants (predominantly nasals and liquids) around the stressed vowels to facilitate location of F0 targets in the F0 trajectory. The two data sets thus yielded 10 target words (14) that were put in two TOPIC positions and two FOCUS positions.

(14) target words (stressed syllable capitalized)

<table>
<thead>
<tr>
<th>set M</th>
<th>set N</th>
</tr>
</thead>
<tbody>
<tr>
<td>MA:la:k</td>
<td>NA:da</td>
</tr>
<tr>
<td>MA:la:ka</td>
<td>NA:ga:fa</td>
</tr>
<tr>
<td>MA:ri:am</td>
<td>NA:gi</td>
</tr>
<tr>
<td>MA:ri</td>
<td>NA:gi</td>
</tr>
<tr>
<td>gaMA:ri</td>
<td>maNA:ri</td>
</tr>
</tbody>
</table>

The experimental data were recorded in two sessions. The first session was recorded with six speakers, but the data of one participant had to be excluded, as already mentioned above. For the statistical

---

⁹ I am grateful to Tarek Eltayeb for helping me with the Arabic transcripts.
analysis of this experiment I also excluded the data of speaker (F2) for reasons to be discussed below. To supplement the original data that had been used as filler sentences for the contrastive focus experiment, a second recording session was carried out to complement the data of the first experiment. This session was recorded with only four of the six speakers. The data of the two sessions together contained TOPICAL target words in two positions: sentence initially (TOP1) and in second position (following the verb) (TOP2), and likewise two FOCAL target words in sentence initial position (FOC1) and in second position (FOC2) (an example is given in 15). There is a certain imbalance between TOPICAL and FOCAL targets because the data from the first session did not include FOCAL targets and TOPICAL targets in first position.

(15) Two target sentences as answers to four different questions (target word in bold face):

a. A1: \textit{mana:l \textit{ra:hit malha l-minja (TOP1)}}
Manal went fun.fair Al-Minya
‘Manal has gone to the Minya fun-fair.’
A2: \textit{mana:l \textit{fi-malha l-minja (TOP1)}}
Manal in-fun.fair Al-Minya

Q: \textit{mana:l fe:n?}
Manal where
‘Where is Manal?’

b. A: \textit{\textit{fuftu mana:l fi-malha l-minja (TOP2)}}
(no) saw.1PL M. in-fun.fair Al-Minya
‘We saw Manal in the Al-Minya fun-fair.’

Q: \textit{fuftu mana:l fe:n?}
saw.2PL M. where
‘Where did you see Manal?’

c. A: \textit{mana:l \textit{ra:hit malha l-minja (FOC1)}}
Manal went fun.fair Al-Minya
‘Manal went/has gone to the Al-Minya fun-fair.’

Q: \textit{\textit{mi:n ra:h malha l-minja?}}
who went fun.fair Al-Minya
‘Who went/has gone to the Al-Minya fun-fair?’

d. A: \textit{\textit{fuftu mana:l fi-malha l-minja (FOC2)}}
saw.1PL Manal in-fun.fair Al-Minya

Q: \textit{fuftu \textit{mi:n fi-malha l-minja?}}
saw.2PL who in-fun.fair Al-Minya
‘Who did you see in the Al-Minya fun-fair?’

Session 1 contained 2 repetitions and session 2 three repetitions. Only for speaker F0 the second session was divided in two sub-sessions with 2 repetitions each and a pause between them. In sum, the complete data set contained eleven repetitions of TOPICAL target words and five repetitions of FOCAL target words per syllable condition for three speakers and 12 repetitions of TOPICAL targets plus six repetitions of FOCAL targets for one speaker. This added up to $11 \times 5 \times 2 + 5 \times 5 \times 2 = 160$.
potential utterances for each of the three speakers M0, M1, F1, and \(12 \times 5 \times 2 + 6 \times 5 \times 2 = 180\) potential utterances for speaker F0, in sum 660 potential utterances for the analysis. 44 utterances had to be discarded because of lacking pitch values or disfluencies. Utterances were also excluded if they showed a pattern that clearly exhibited a strategy of marking another type of information structure. We must not forget that the task included different FOCUS structures in a pseudo-randomized order and required a high level of concentration. We may thus expect a speaker to make mistakes, when his/her concentration decreases. Unclear cases, however, were included in the analysis. Utterances containing a break were not excluded but marked accordingly. It has been shown in the literature that a break may change the behaviour of tones in various ways, both affecting alignment and scaling. The participants were asked to read the sentence in one go and their renderings were also influenced by the transcription of the answers in some cases. It is well known that EA has phonological rules of epenthesis and vowel deletion operating on phrase level (Kenstowicz 1980; Watson 2002). Thus the sentences in (15b, 15d) above, if articulated in one phrase, should be spoken with the vowel of the preposition fi ‘in’ elided if the preceding word ended in a vowel. This was also noted in the transcription by omitting the letter i and only writing the letter f as shown in (16). The decision whether there was a phrase break involved or not was taken in collaboration with a musicologist who has no knowledge of Egyptian Arabic. To reach a decision was not especially problematic, but in the undecided cases the example was labelled as involving a break. The statistics were run twice, once with and once without breaks to see whether a break would change the result significantly.

\[(16)\]
\[
\text{fu}f\text{na n}a\text{gwa f}malga? \text{il-musinni:n saw.1PL in-asylum DEF-old.people}
\]
\[‘\text{We saw Nagwa in the old people’s home.’}\]

The whole data of both sessions, including the omitted data of the two additional speakers were subjected to qualitative analysis and labelled by hand and finally quantitative analysis was carried out for 616 utterances.

4.4.1.1.1 Hypotheses

The aim of this experiment was to test whether the tonal contour differences observed in the corpus data would also differ significantly for information FOCUS elicited by a simple information question. Contrary to experiment 1, there were no explicit alternatives in the context. Nevertheless, the type of FOCUS tested was a narrow FOCUS occurring early in the utterance. The main hypothesis for the experiment was that TOPICAL and FOCAL constituents exhibit two different tonal contours. The assumption was that FOCAL constituents would exhibit a rising-falling contour while TOPICAL constituents were supposed to have a largely rising or perhaps flat contour (a leading or a linking accent), or a rise-fall with different alignment and scaling values than in FOCUS condition (i.e. a
default accent). At the same time, it was not assumed that post-FOCAL material would be deaccented and probably not even downtoned, as opposed to the contrastive focus cases of correction investigated in experiment 1 (4.3.1.1). Phonetic studies on the prosodic means of focus marking have proposed different correlates that are supposed to enhance the prominence of the focus accent. Although many of the studies deal with the distinction between contrastive focus and information focus, or narrow focus and broad focus, the present experiment was designed, following these different proposals and altering the experiment design according to the assumptions made here. Suggested acoustic correlates of FOCUS are pitch accent type (Selkirk 2002; Avesani & Vayra 2003; Baumann et al. 2006), peak height (Ladd & Morton 1997; Xu & Xu 2005; Baumann et al. 2007; Féry & Kügler 2008; Hanssen et al. 2008; Féry & Ishihara 2010), excursion size (Xu & Xu 2005; Hanssen et al. 2008), alignment of tones (Peters 2002; Baumann et al. 2007; Hanssen et al. 2008) and duration (Xu & Xu 2005; Féry & Ishihara 2010; Baumann et al. 2007; Hanssen et al. 2008).

In the present experiment, the difference in rising and falling contour was supposed to be phonetically encoded primarily by the slope of the fall to the next L-tone after the high tone has been reached. As we have seen above, rises and falls may span the domain of one to several accents. A short target of one word was chosen to elicit the respective contour within one single accent that could be schematically represented as in Figure 15. Note that

![Diagram of leading (dashed line) and linking (dotted line) TOPIC accent and a closing (solid line) FOCUS accent.](https://example.com/diagram.png)

**Fig. 15:** schematic representation of leading (dashed line) and a linking (dotted line) TOPIC accent and a closing (solid line) FOCUS accent.

As outlined in 2.2.4.1, the tones in a default accent are supposed to have a certain alignment, based on the findings in Hellmuth (2006b) (cf. 2.2.3). It was supposed that the FOCUS accent should exhibit earlier alignment of the second flanking L tone (L2) (Rastegar-El Zarka 1997; Hanssen et al. 2008). It was further assumed that the peak should be aligned earlier, i.e. within the stressed syllable in all syllable types - either as an effect of the earlier alignment of L2 and/or as an effect of a more precise articulation.

Based on the findings of Norlin (1989) and Hellmuth (2006b, 2010b), the scaling of the peak could be significantly higher in FOCUS than in TOPIC condition and the post-focal material could be realized
within a compressed pitch range. It was also supposed that FOCUS constituents should exhibit a higher duration than TOPIC constituents. It has been shown that duration is affected in various ways, effects have been shown in the stressed vowel, the onset and coda consonants, the stressed syllable or the stress foot as a whole (Baumann et al. 2007; Hanssen et al. 2008). A third parameter correlating with the other parameters mentioned so far is velocity and amount of rise and fall, respectively (Xu & Xu 2005; Hanssen et al. 2008).

4.4.1.2 Methods and data

Based on the findings in the literature and the hypotheses outlined above, the following parameters were tested:

a. **Relative scaling of H tones**: The difference in peak height between a FOCAL peak and a TOPICAL peak and the accent following them was calculated. A difference in peak height between successive accents would point to the existence of prominence differences between them. That could mean that either the FOCAL accent was made more prominent by a higher pitch excursion and/or by compression of the following accents, or both.

b. **Absolute scaling of tones**: To determine whether it was the expansion of the FOCAL accent or the compression of the accent following it (or both) that was relevant, absolute values were looked at as well. Additionally, the scaling of the flanking L-tones was measured.

c. **Comparison of rises and falls**: In line with the hypothesis that the second part of the accent is the communicatively important one, the falls of the FOCAL accents were compared to the falls in TOPICAL accents in terms of their amount and the velocity of pitch change (following Xu & Xu 2005). I assumed that FOCAL falls should be steeper and quicker than TOPICAL falls. In addition, the rises of the different accent types were also compared with each other.

d. **Duration values**: Based on the assumption that FOCAL accents are characterized by higher duration, different duration values were calculated, following Hanssen et al. (2008) and compared between accent types.
   - duration of the word
   - duration of the onset of the stressed syllable
   - duration of the vowel of the stressed syllable

e. **Alignment**: To test whether the alignment of the individual tones differed between accent types, the alignment of both L tones and the H tone was calculated with respect to the stressed syllable and to the right word boundary. In addition, the relative position of the H tone inside the word was calculated. Based on the investigation of Egyptian MSA intonation (Rastegar-El

---

10 Note, however, that according to Hellmuth (2010a) these features only apply to contrastive focus.
Zarka 1997), I assumed that FOCAL accents should exhibit an earlier L2 than TOPIC accents and probably that some TOPIC accents would show a later alignment of the H tone than in the default case. In addition, it was assumed that TOPIC accent would be linked to the following accent domain and thus do not show a L2.

f. **Comparison of values at wordend:** The most important measure to test the main hypothesis was the amount of fall from H to the end of the FOCUS and TOPIC constituents. Assuming that what perceptually counts is whether the tonal contour falls, rises or stays the same within the morpho-lexical boundaries of a constituent and not within a phonological constituent such as the syllable or the foot, all pitch fall values were also calculated with respect to the end of the lexical item expressing the TOPICAL or FOCAL referent. I.e. not only the whole amount of the fall from an H tone to the following L tone was looked at, but also the amount that is realized by the end of the word itself. This is due to the fact that interpolation between the individual tones is not always even. This method was chosen to take plateau intonations into account that seem to be quite frequent in TOPICS in EA before the F0 starts going down rapidly until it reaches the next low turning point. Plateaux cannot be captured by Standard pitch accent notation, unless a phrase tone is assumed.

Segmental and tonal labels (Tab.5) were assigned using the speech analysis system PRAAT. Labels included the boundaries of the target words, the segments, and the L and H tones associated with pitch accents (cf. Fig. 16).

<table>
<thead>
<tr>
<th>Labels</th>
<th>TOP</th>
<th>FOC</th>
</tr>
</thead>
<tbody>
<tr>
<td>L1</td>
<td>elbow before the pitch peak</td>
<td>L3</td>
</tr>
<tr>
<td>H1</td>
<td>maximum F0 of the pitch accent</td>
<td>H2</td>
</tr>
<tr>
<td>L2</td>
<td>elbow after the pitch peak</td>
<td>L4</td>
</tr>
<tr>
<td>C01</td>
<td>beginning of onset consonant of stressed syllable</td>
<td>C02</td>
</tr>
<tr>
<td>V01</td>
<td>beginning of vowel of stressed syllable</td>
<td>V02</td>
</tr>
<tr>
<td>C11</td>
<td>beginning of first consonant after stressed vowel (coda or onset)</td>
<td>C12</td>
</tr>
<tr>
<td>C12</td>
<td>beginning of second consonant after stressed vowel (=onset of next syllable)</td>
<td>C22</td>
</tr>
<tr>
<td>V12</td>
<td>beginning of vowel of poststress syllable</td>
<td>V22</td>
</tr>
<tr>
<td>H2</td>
<td>peak of next accent</td>
<td>h</td>
</tr>
<tr>
<td>l</td>
<td>lowest point at or near the end of an utterance</td>
<td></td>
</tr>
</tbody>
</table>

Tab.5: Labels used in the experiment for TOPICS in the left column and for FOCI in the right column.
Fig. 16: Example screen shots of two labelled utterances: a TOPICAL target word in the upper panel and a FOCAL target word in the lower panel.

Pitch values were measured in semitones re 100 Hz. The values were extracted using a PRAAT script written for the purpose at hand. The calculated relative values (amounts of pitch change and duration values) were subjected to statistical analysis using SPSS.

4.4.1.1.3 Results

A qualitative analysis of each speaker separately revealed different prosodic strategies for the individual speakers (Tab. 6); typical examples of contours for the four speakers are shown in Figure 16. The fifth speaker's (F2) utterances were excluded from statistical analysis because her strategy to mark the FOCUS was completely opposite to the strategy pursued by the other speakers. Her FOCI were characterized by a much wider rise and a significantly longer duration than her TOPIC constituents and usually were followed by pause. Three speakers (M0, F0 and F1) were quite consistent in the strategy employed, while speaker M1 did not show a clear pattern. Perceptually, it is frequently not possible to distinguish between his FOCI and his TOPICS. It is well known that subjects differ as to their ability to fulfil a reading task of some complexity. There is some evidence that speaker M1’s performance is simply poor; his data will be discussed below.

Speaker F0's speech is vivid and her utterances typically exhibit strong accents on every content word, even in a 'neutral' reading contour, which is falling from the beginning to the end. I.e. the topics are

---

11 Table 6 only includes the observation for the three speakers F0, F1 and M0 because they all exhibited clear, albeit partly heterogeneous, FOCUS marking strategies.
high, with strong intensity and the FOCUS domain is falling to low. The accents on the target words differ, however, according to expectations, although speaker F0 tends to truncate falls in closing accents. Still, her TOPIC accents are significantly different from her FOCUS accents, frequently continuous (LH-) and occasionally rising altogether (LH'). She makes less pervasive use of downtoning of the post-focal material than speakers F1 and M0, but she still marks the difference. Her downtoning also never comes close to deaccenting as is the case with the other two speakers. The comparison of the contours in Figure (17a) and (17b) reveal this difference, not only in the target words themselves, but also in the following part of the curve which is realized in a compressed pitch range after the FOCAL accent (17b). Her breaks are evenly distributed between TOPICS and FOCI.

The interesting aspect of speaker F0's renditions is the fact that even without clear downtoning the FOCAL accent can be clearly distinguished from the TOPICAL accent. I suspect that upon hearing the accent on the target word alone without the rest of the utterance it should be possible to decide whether the target is a TOPIC or a FOCUS. So far, this is just a speculation and it will not be possible to support the hypothesis or falsify it here. A perception experiment should shed some light on this issue. If I am on the right track, this means that downtoning - and deaccenting as an extreme case - only helps identify the FOCUS instead of being the decisive factor as commonly assumed. This would at least be true for a language like EA that quite consistently marks word accent by pitch.

Speaker F1 is a less vivid speaker. Her speech rate is very high. Many of her utterances do not show the bumps of the wave pattern but are smoothly falling throughout (17c). Accordingly her topics are frequently flat and non-prominent, but there is also a fair amount of continuous and rising topics. Speaker F1 nevertheless marks FOCUS very strongly with a clear rise-fall and downtoning of the following material. To enhance prominence of the closing accent, her final Ls are sometimes aligned very early and she also frequently uses a break after FOCUS.

Speaker M0 behaves as expected. He almost always uses closing accents in which L2 is reached by the end of the FOCUS domain with quite consistent downtoning afterwards (17f). His topics frequently rise or are continuous, and his default accents only fall slightly. He prefers a hat pattern between topic and the first one of the FOCUS accents as displayed in Figure (17c); but also default falling contours are frequent. His breaks are evenly distributed between TOPICS and FOCI.

As already noted, speaker M1 is the odd one out. Auditory analysis does not reveal a clear perceptual difference between TOPIC and FOCUS. Nevertheless, speaker M1 seems to try marking FOCUS. His FOCUS constituents mainly exhibit longer duration and a frequent break after the target word. In general, speaker M1 is a not a gifted reader (and not a gifted narrator either). Some evidence for speaker M1's lack of concentration is his rather monotonous reading and the frequent use of the redundancy contour (17h) which can be interpreted as indicating boredom and lack of concentration. Another interpretation would be that the redundancy contour is associated with presupposed material.
Recall that the redundancy contour consists of a fall-rise-slump stretching out over two accents. When used after the FOCAL target, the first accent is always an inverted one, mostly aligning the low with the stressed syllable (L*) and rising afterwards. This is an accent typically associated with given elements that selects a variable from the background (Gussenhoven 1983: 384). Table 6 illustrates the different TOPIC and FOCUS marking strategies employed by the three speaker F0, F1, and M1.

![Diagram of F0 and F1 speech patterns](image)

Tab. 6: Different strategies of TOPIC and FOCUS marking from three speakers.
Fig. 17: Typical examples for four speakers; TOPICAL targets in the left column and FOCAL targets in the right column.

Turning now to the quantitative analysis of the data: Normality was tested using a Shapiro-Wilcoxon text and normally distributed data were subjected to a t-test with two independent samples. For not normally distributed data the non-parametric Mann-Whitney-U-test was applied. Only the most important figures will be given in Tables 7 and 8.  

The first columns indicate significance level, n.s (not significant), * (significant on 5% level), ** (significant on 1% level), *** (p<.0005) values. The second columns contain mean values with standard deviations in parenthesis. Some values for speaker M1 are not included in the table, but will be discussed later in the text. For most of the data, all TOPIC data were contrasted with all FOCUS data. For some of the data, however, there were clear effects of position, which was either noted in the text, or the values were given for first and second position separately. This was especially the case when calculating a mean value did not make sense, notably in the case of absolute pitch values (Table 8, lines 1-4). As already mentioned, statistical analysis was carried out twice, once for all the data including utterances containing a break after the TOPIC or the FOCUS and once for the utterances without a break. The breaks only had a significant effect for some of the parameters investigated, most notably duration values. In the following, the values for the non-break data are given. Especially relevant results are highlighted in bold print.

---

12 A detailed description of the experiment will be published separately.
Regarding duration, the results show that duration, especially word duration, is an important acoustic correlate of narrow FOCUS in EA. All speakers show highly significant duration differences on word level. While vowel duration also differs significantly for all speakers, onset duration does not. However, all mean duration values on segment and word level are larger in FOCUS position than in TOPIC position for all speakers. Note that word duration was only measured from the beginning of the first consonant to the end of the final vowel. The duration of the final consonant was not included in the measurement to make extraction of pitch values possible. This decision, however, has the effect that the major correlate of a break, the lengthening of the final segment was factored out which resulted in a less significant effect of a break on duration values than expected.

Concerning alignment, the hypothesis of the earlier alignment of L2 is not wholly supported by the data, at least not on the level of fall duration (the distance between L2 and the preceding peak). As shown in Table 7 (line 7), mean fall duration was even remarkably stable between TOP and FOC conditions across the three speakers F1, M0 and M1. Only speaker F0 exhibited a significant difference between TOPIC and FOCUS, caused by a longer fall in TOPICS. On the other hand, L2 alignment with respect to wordend (Table 7, line 8) showed the expected difference, which is highly significant for the three speakers F0, F1 and M0. For all three speakers the next low typically lies after the word boundary (mean values of 120 ms for F0, 115 ms for F1, and 115 ms for M0) in TOP condition, whereas it is situated at the end of the target word in FOC condition (mean values of 20 ms for F0, and and 35 ms for F1 and M0 after the end of the vowel). The alignment of L in FOC position is quite stable at wordend (the median values for all three speakers being 0,00), while the TOP values show a wider variation. Figure 18 shows the respective box plot diagrams for speakers F0, F1, and M0; note the strong concentration of L2 alignment in FOC condition around zero, especially in the

**Duration and alignment**

<table>
<thead>
<tr>
<th></th>
<th>F0</th>
<th>TOP</th>
<th>FOC</th>
<th>F1</th>
<th>TOP</th>
<th>FOC</th>
<th>M0</th>
<th>TOP</th>
<th>FOC</th>
<th>M1</th>
<th>TOP</th>
<th>FOC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Onset dur</td>
<td>*</td>
<td>59</td>
<td>63</td>
<td>n.s.</td>
<td>57</td>
<td>64</td>
<td>n.s.</td>
<td>54</td>
<td>65</td>
<td>n.s.</td>
<td>50</td>
<td>53</td>
</tr>
<tr>
<td>Vowel dur</td>
<td>***</td>
<td>69</td>
<td>94</td>
<td>(18)</td>
<td>***</td>
<td>65</td>
<td>98</td>
<td>(25)</td>
<td>*</td>
<td>97</td>
<td>(27)</td>
<td>*</td>
</tr>
<tr>
<td>Word dur</td>
<td>***</td>
<td>278</td>
<td>335</td>
<td>(51)</td>
<td>***</td>
<td>261</td>
<td>334</td>
<td>(58)</td>
<td>***</td>
<td>316</td>
<td>(51)</td>
<td>***</td>
</tr>
<tr>
<td>Peak delay</td>
<td>***</td>
<td>65</td>
<td>49</td>
<td>(24)</td>
<td>***</td>
<td>72</td>
<td>51</td>
<td>(37)</td>
<td>***</td>
<td>70</td>
<td>(19)</td>
<td>n.s.</td>
</tr>
<tr>
<td>H to wordend</td>
<td>***</td>
<td>90</td>
<td>160</td>
<td>(70)</td>
<td>***</td>
<td>90</td>
<td>160</td>
<td>(71)</td>
<td>***</td>
<td>87</td>
<td>(66)</td>
<td>nv</td>
</tr>
<tr>
<td>H to Vend</td>
<td>***</td>
<td>40</td>
<td>33</td>
<td>(66)</td>
<td>***</td>
<td>33</td>
<td>62</td>
<td>(45)</td>
<td>***</td>
<td>61</td>
<td>(54)</td>
<td>n.s.</td>
</tr>
<tr>
<td>H to L2</td>
<td>***</td>
<td>220</td>
<td>180</td>
<td>(62)</td>
<td>n.s.</td>
<td>190</td>
<td>190</td>
<td>(160)</td>
<td>n.s.</td>
<td>210</td>
<td>(70)</td>
<td>n.s.</td>
</tr>
<tr>
<td>L2 to wordend</td>
<td>***</td>
<td>120</td>
<td>20</td>
<td>(65)</td>
<td>***</td>
<td>115</td>
<td>35</td>
<td>(186)</td>
<td>***</td>
<td>127</td>
<td>(115)</td>
<td>n.s.</td>
</tr>
</tbody>
</table>

Tab.7: Duration and alignment values are given in ms, except the values for relative H alignment in the fourth row, which are given as a percentage of the distance between the beginning of the accented syllable and wordend.

Tab. 7: Duration and alignment values are given in ms, except the values for relative H alignment in the fourth row, which are given as a percentage of the distance between the beginning of the accented syllable and wordend.

The alignment of L in FOC position is quite stable at wordend (the median values for all three speakers being 0,00), while the TOP values show a wider variation. Figure 18 shows the respective box plot diagrams for speakers F0, F1, and M0; note the strong concentration of L2 alignment in FOC condition around zero, especially in the
third diagram. In sum, we may note that in narrow FOCUS cases there is a tendency of the closing L tone to seek the boundary of the semantic constituent it is associated with.

Fig. 18: Three box-plot diagrams illustrating L2 position with respect to wordend for speakers F0 (left), F1 (middle), and M0 (right). The horizontal red line indicates word end position.

The second hypothesis was that the peak would be aligned differently in TOP and FOC conditions. It has been suggested that H tones may be aligned earlier under narrow FOCUS than in ‘neutral’ cases (Smiljanić & Hualde 2000 for the Zagreb dialect of Serbo-Croatian). Gussenhoven (2004: 60-61) suggests that this could be interpreted as a more precise location of the peak inside the stressed syllable in cases of narrow FOCUS. The EA data indeed show a significant difference of peak alignment. The H position under the two experimental conditions was calculated in three ways, once as the time distance to the end of the target word (H to wordend), once as a percentage of the duration from the beginning of the accent (L1) to the end of the target word (peak delay) and another time in relation to the stressed syllable (H to Vend). As shown in lines 4 to 6 of Table 7, these three calculations exhibit highly significant differences. The results clearly show that in the case of TOPICS the H tone was considerably closer to the end of the target word (a peak delay of 70% in TOP condition and between 36% and 51% in FOC condition). Furthermore, as can be seen from the figures in line 6, the peak was indeed aligned after the stressed vowel in TOP condition (cf. the typical alignment facts for default accents as established by Hellmuth 2006b, see Chapter 2.2.3) and within the stressed vowel in FOC condition (as suggested by the negative values in the Table). The assumption of meaningful variation would have to be rejected, if earlier peak alignment could be shown to be phonologically conditioned, e.g. by an upcoming boundary and/or boundary tone (Silverman & Pierrehumbert 1990; Prieto et al. 1995) causing tonal repulsion. In the light of the present results, the first of these assumptions is easily refuted, as the distance between the peak and the boundary of the target word is significantly smaller in TOP than in FOC condition, also in those utterances that did not contain a break (Table 7, line 5). The results for fall time, however, suggest that peak alignment and L2 alignment correlate. Note, however, that fall time is not really constant for all speakers, specifically speaker F0 shows a significant difference between TOP and FOC conditions. On the other hand the peak was aligned within the stressed vowel in the majority of FOC tokens for the
three speakers F0 (78%), F1 (67%) and M0 (88%), which suggests that peak alignment within the stressed syllable might in fact be due to a more precise articulation as suggested by Gussenhoven.

The other acoustic correlate investigated by the experiment is scaling and direction of pitch change. Table 8 gives three types of pitch values: (i) absolute values (line 1-7), (ii) relative values, i.e. scaling differences between H tones (line 8) and the amount of pitch change within one accent domain (line 9-11) and finally velocity values for rises and falls (line 12-13).

The results for absolute pitch values are interesting, suggesting that higher peak scaling (line 1-2) and even greater excursion size is no necessary correlate of a FOCUS accent. The absolute values are given for first and second position targets separately. Peak scaling shows a significant difference between first and second position, i.e. scaling only differs between TOP and FOC conditions for first position targets. The only speaker that shows a highly significant difference is speaker M1 who does not mark the difference between FOCUS and TOPIC realization as strongly as the other speakers, specifically his FOCUS utterances do not exhibit any downtoning, i.e. no compression of pitch range after the FOCUS accent. But also the other speakers only show higher H values in first position targets. This higher scaling, however, correlates with a higher beginning of L1 so that the excursion size of the rise does not differ very much between TOP and FOC conditions for most speakers. This can be seen in the values for rise amount in line 11. For speaker F0, the amount of rise under TOP and FOC conditions does not differ significantly, speaker F1’s FOCUS accent excursions are considerably greater than her TOPIC accent excursions, whereas speaker M0’s mean value of FOCUS rise amount is even slightly smaller than that of his TOPIC rise amount.

<table>
<thead>
<tr>
<th>Absolute pitch values</th>
<th>F0</th>
<th>TOP</th>
<th>FOC</th>
<th>F0</th>
<th>TOP</th>
<th>FOC</th>
<th>M0</th>
<th>TOP</th>
<th>FOC</th>
<th>M1</th>
<th>TOP</th>
<th>FOC</th>
</tr>
</thead>
<tbody>
<tr>
<td>H POS1</td>
<td>**</td>
<td></td>
<td></td>
<td>**</td>
<td></td>
<td></td>
<td>6.82</td>
<td></td>
<td></td>
<td>7.67</td>
<td></td>
<td></td>
</tr>
<tr>
<td>H POS2</td>
<td>n.s.</td>
<td></td>
<td></td>
<td>n.s.</td>
<td></td>
<td></td>
<td>6.19</td>
<td></td>
<td></td>
<td>5.91</td>
<td></td>
<td></td>
</tr>
<tr>
<td>L1 POS1</td>
<td>***</td>
<td></td>
<td></td>
<td>15.15</td>
<td></td>
<td></td>
<td>1.64</td>
<td></td>
<td></td>
<td>2.45</td>
<td></td>
<td></td>
</tr>
<tr>
<td>L1 POS2</td>
<td>n.s.</td>
<td></td>
<td></td>
<td>17.17</td>
<td></td>
<td></td>
<td>0.54</td>
<td></td>
<td></td>
<td>2.42</td>
<td></td>
<td></td>
</tr>
<tr>
<td>L2 POS1</td>
<td>n.s.</td>
<td></td>
<td></td>
<td>17.06</td>
<td></td>
<td></td>
<td>0.54</td>
<td></td>
<td></td>
<td>2.42</td>
<td></td>
<td></td>
</tr>
<tr>
<td>L2 POS2</td>
<td>n.s.</td>
<td></td>
<td></td>
<td>17.06</td>
<td></td>
<td></td>
<td>0.54</td>
<td></td>
<td></td>
<td>2.42</td>
<td></td>
<td></td>
</tr>
<tr>
<td>l values</td>
<td>n.s.</td>
<td></td>
<td></td>
<td>8.38</td>
<td></td>
<td></td>
<td>-5.26</td>
<td></td>
<td></td>
<td>-5.31</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| H POS1               | 24.19 | 25.47 |     | 19.13 |     |     | 7.67  |     |     | 12.04 |     |     |
| H POS2               | 23.51 | 24.06 |     | 16.83 |     |     | 6.19  |     |     | 10.87 |     |     |
| L1 POS1              | 15.15 | 16.48 |     | 14.28 |     |     | 1.64  |     |     | 8.25  |     |     |
| L1 POS2              | 17.17 | 17.23 |     | 13.80 |     |     | 0.54  |     |     | 8.33  |     |     |
| L2 POS1              | 17.06 | 16.74 |     | 10.77 |     |     | 0.36  |     |     | 7.66  |     |     |
| L2 POS2              | 17.06 | 16.74 |     | 9.33  |     |     | 0.36  |     |     | 7.66  |     |     |
| l values             | 8.38  | 8.66  |     | 8.02  |     |     | -5.26 |     |     | 2.52  |     |     |

<p>| H POS1               | 25.47 | (2.10) |     | 19.13 | (2.15) |     | 7.67  | (1.05) |     | 12.04 | (2.11) |     |
| H POS2               | 24.06 | (0.94) |     | 16.83 | (1.83) |     | 6.19  | (1.53) |     | 10.87 | (2.10) |     |
| L1 POS1              | 16.48 | (1.83) |     | 14.28 | (1.08) |     | 1.64  | (0.82) |     | 8.25  | (2.54) |     |
| L1 POS2              | 17.23 | (1.63) |     | 13.80 | (0.06) |     | 0.54  | (1.61) |     | 8.33  | (2.50) |     |
| L2 POS1              | 16.74 | (1.90) |     | 10.77 | (2.00) |     | 0.36  | (1.28) |     | 7.66  | (2.49) |     |
| L2 POS2              | 16.74 | (1.90) |     | 9.33  | (1.83) |     | 0.36  | (1.28) |     | 7.66  | (2.49) |     |
| l values             | 8.66  | (1.40) |     | 8.02  | (2.14) |     | -5.26 | (0.67) |     | 2.52  | (2.44) |     |</p>
<table>
<thead>
<tr>
<th>Relative pitch values</th>
<th>F0</th>
<th>TOP</th>
<th>FOC</th>
<th>F1</th>
<th>TOP</th>
<th>FOC</th>
<th>M0</th>
<th>TOP</th>
<th>FOC</th>
<th>M1</th>
<th>TOP</th>
<th>FOC</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>HtoH diff.</strong></td>
<td>***</td>
<td>4.28</td>
<td>(2.18)</td>
<td>6.98</td>
<td>(2.27)</td>
<td>***</td>
<td>1.45</td>
<td>(2.16)</td>
<td>7.15</td>
<td>(1.66)</td>
<td>***</td>
<td>1.56</td>
</tr>
<tr>
<td>Fall</td>
<td>n.s.</td>
<td>7.57</td>
<td>(2.73)</td>
<td>8.38</td>
<td>(2.73)</td>
<td>***</td>
<td>3.18</td>
<td>(1.62)</td>
<td>8.04</td>
<td>(3.01)</td>
<td>***</td>
<td>3.83</td>
</tr>
<tr>
<td><strong>Fall to wordend</strong></td>
<td>***</td>
<td>2.98</td>
<td>(2.27)</td>
<td>7.35</td>
<td>(3.32)</td>
<td>***</td>
<td>1.76</td>
<td>(1.25)</td>
<td>7.16</td>
<td>(2.59)</td>
<td>***</td>
<td>1.59</td>
</tr>
<tr>
<td>Rise</td>
<td>n.s.</td>
<td>7.51</td>
<td>(2.55)</td>
<td>8.33</td>
<td>(2.22)</td>
<td>***</td>
<td>2.31</td>
<td>(1.70)</td>
<td>4.52</td>
<td>(1.82)</td>
<td>*</td>
<td>5.34</td>
</tr>
<tr>
<td><strong>Velo of fall to wordend</strong></td>
<td>***</td>
<td>29.48</td>
<td>(16.52)</td>
<td>49.41</td>
<td>(15.46)</td>
<td>***</td>
<td>24.38</td>
<td>(22.74)</td>
<td>48.05</td>
<td>(17.01)</td>
<td>***</td>
<td>19.83</td>
</tr>
<tr>
<td>Velo of rise</td>
<td>**</td>
<td>51.33</td>
<td>(18.2)</td>
<td>61.06</td>
<td>(19.17)</td>
<td>***</td>
<td>18.94</td>
<td>(12.15)</td>
<td>30.04</td>
<td>(11.10)</td>
<td>***</td>
<td>29.80</td>
</tr>
</tbody>
</table>

Tab. 8: Absolute pitch values of H, L1 and L2 low given separately for sentence positions (lines1-7); relative pitch values between successive H tones (line 8) between H and next L (amount of fall) (line 9), H and wordend (amount of fall to wordend) (line 10) and L1 and H (amount of rise) (line 11); velocity of fall (line 12) and velocity of rise (line 13).

In Figure 19, prototypical TOPIC and FOCUS accents for speakers F0, F1 and M0 are illustrated schematically using the mean values of all their utterances. The utterances have, however, been normalized for duration so as to show the difference in tonal alignment between the two conditions. Figure 20 gives two schematic illustrations for speaker M1’s accent realizations. His utterances had to be treated separately for each target position, because he obviously pursues different strategies in the two positions. He shows a slightly earlier L2 alignment in both FOCUS accents. Contrary to the comparable results of the other speakers, this difference is however not significant. The more striking strategy is the higher scaling of the peak in position 1 and the more pronounced fall in position 2. This second strategy is the most salient feature of a FOCUS accent in the speech of the other speakers.

Contrary to the amount of rise that does not differ significantly between experimental conditions (note, however, that the velocity does differ between conditions because of the earlier alignment of H), the amount of fall shows a clear difference, brought about by the lower scaling of the closing L tone and not necessarily by higher peak scaling. This is especially salient in speaker F1’s and M0’s speech, and also in the second position targets of speaker M1. Speaker F0, however, shows no clear difference in the amount of fall between conditions. But her accents also exhibit a significant difference in the amount of fall by the time the end of the FOCUS constituent is reached (line 10). Conspicuously, the magnitude of the fall by wordend was considerably larger under FOC condition than under TOP condition for all speakers (2.5:1 to 4.5:1). The quite consistent alignment of the closing L in the vicinity of the end of the target word for F0, F1 and M0, as illustrated in Figure 18 and 19, results in compression of the pitch fall (Ladd 1996; Grabe et al. 2000) and lends more prominence to the accented word.
Furthermore, the meaningful difference between fast fall or slow fall in an accent domain, especially to wordend (Table 8, line 12), in EA supports the assumption that it is the second (falling) part of the accent that bears the functional load. The quite stably aligned rise with the stressed syllable in EA suggests its primary function to be a correlate of word accent, whereas the intonational meaning is mostly expressed by the continuation of the contour after the peak. In Standard AM Theory, this function is attributed to a phrase tone. However, this analysis is rather circular, as it assumes that the edge of a phrase is marked by an obligatory phrase tone, attributing every occurrence of a tonal event after the accented syllable to the occurrence of a phrase boundary, irrespective of whether there is a perceptual boundary present or not. Furthermore, the EA data show that the closing L tone may be present as a target distinct from the L target at the beginning of the next accent domain, even in TOPIC cases. This fact obviates the analysis of the closing L as a focus tone. 39% of the TOPIC utterances without a break for speaker F0, 37% for speaker F1, 70% for speaker M0 and 27% for speaker M1 exhibited a distinct L2 before the L1 of the following accent domain (as illustrated in Figure 13a above), in the other cases the two tones merged as an instance of tone linking. The difference between the speakers was clearly influenced by speech tempo and accuracy of speech. In the present model, the two L tones are interpreted as delimiting the accent domain with the higher functional load attributed to the second L whose alignment and scaling properties are varied to fulfill different functions.

13 This assumption is in line with the descriptive tradition of the British School of Intonation and the ToDI framework (Gussenhoven (2005)), but contra Standard AM Theory.
Finally, qualitative observation suggested that compression of pitch range after a narrow FOCUS was not only due to lower scaling of the following H tone, but also to higher scaling of the final L of the utterance (labelled l; see Table 8, line 7), meaning that pitch range was compressed from both sides, from above and below. The explanation for this fact would be that FOCAL falls are due to a purposeful lowering of the pitch range at the end of the utterance, indicating assertiveness, as illustrated in the one-accent FOCAL falls in Figure (19) and (20b) above (frequency code), whereas the subsequent lowering would only be due to the physiological mechanism of decreasing subglottal pressure (production code). The results of the quantitative analysis, however, did not verify this hypothesis. At least, the difference between TOP and FOC conditions were only significant for speaker F1 and almost significant for speaker F0. Three of the speakers, show a difference in the ‘right’ direction, i.e. the mean value of their utterance final L-tones was higher in FOC condition than in TOP condition. Speaker M0’s utterances exhibit exactly the opposite. However, the qualitative observation was quite robust, and it remains to be seen which factors influence this parameter. I hypothesize that a shorter post-FOCAL might have shown a significant difference.

Finally, the results also corroborate the hypothesis that intonational phrasing and tonal contour are a priori independent prosodic phenomena. The double calculations revealed a significant difference only for some of the results, most notably duration values. The values given in the Tables pertain to utterances without an intonation break and give ample evidence that the tonal behavior is not due to any kind of boundary tones.

4.4.1.1.4 Discussion and conclusions

Generally, the results of the production study corroborate the hypotheses put forward under 4.4.1.1.1. Specifically, it is shown that narrow FOCUS tends to be articulated with a closing accent (Bolinger’s A profile) and a TOPIC with a default accent or a leading accent, both subtypes of Bolinger’s B profile. Qualitative analysis of the data revealed that TOPIC articulation was subject to higher variability, varying between leading, linking and default contours. This resulted in a higher variability of peak alignment in TOPIC condition. It could be specifically shown that direction of pitch change is more stable than prominence manifested in greater excursion size and consequently differing scaling properties of the individual tones. This is especially true for the peak H that is not significantly higher in FOCI than in TOPICS for all speakers. While speaker F0’s FOCI were perceptually not much more prominent than her TOPICS, the difference in prominence was striking for the other two speakers. I assume that this prominence difference is brought about by the higher amount of downtoning after FOCUS applied by these speakers as well as by the steeper and lower fall of the accent. That means that excursion size is relevant for the perception of prominence, but in the present experiment the accents exhibited a downward rather than an upward expansion of pitch range.
Another stable correlate of FOCUS accent was longer duration. As already mentioned, intensity values could not be extracted. However, to get a complete picture of acoustic prominence correlates in information structural categories, intensity would have to be investigated as well. Concerning the question whether 'contrastive focus' and 'information focus' have different realizations in EA as suggested by Hellmuth (2006b, 2010b), we may interpret the results in line with the hypothesis that FOCUS as a pragmatic relation usually has the prosodic correlate of falling intonation, while focus as a pragmatic property, i.e. as highlighting important information is realized by a bundle of articulatory and acoustic features, such as more precise articulation (i.e. alignment of tones), expansion of pitch range, longer duration and higher intensity. Although the experiment was designed to elicit information FOCUS, it seems that every type of narrow FOCUS is necessarily accompanied by some measure of additional prominence. However, as the different results for speaker F0 on the one hand and speakers F1 and M1 suggest, this measure of prominence may vary considerably. While F0's utterances are more on the neutral information side of the continuum, F1's and M0's stronger accents give rise to an exhaustive interpretation and perhaps even a contrastive one as a conversational implicature. This final remark is of course only a hypothesis that would have to be further tested in a perception study.

To summarize, the results have shown that early narrow FOCUS in EA is articulated by a focus accent that is mostly of the assertive closing type. Speaker F2, whose utterances were excluded from the quantitative study, employed another strategy, using marked leading accents on the FOCUS constituent. She thus opted for highlighting the FOCUS constituent using a focus accent without assertion (this topic will be further discussed in Section 4.5). Instead, she obviously preferred indicating continuation. The results also suggest that precise tonal alignment as well as the low scaling of the L tones, especially L2, contribute to the increased prominence of the accent, as shown by Hanssen et al. (2008) for English. The results have also shown that intonation boundaries do not significantly alter the most important correlates of TOPIC and FOCUS, but rather enhance the phonetic features of tonal alignment and scaling as well as duration.

The different strategies used by the individual speakers indicate that there is no one-to-one correspondence between information structural categories and prosodic features. The results of the experiment, however, also show that speakers of EA somehow mark the difference in information structure in their prosodic articulation. We must not forget that reading sentences does not necessarily tell us what speakers actually do in spontaneous discourse. The sentences used in this experiment had an early FOCUS with given material following it, something that is not supposed to occur frequently in a natural situation. The whole corpus of spontaneous data does not include such an utterance. It is in fact highly unlikely that a speaker should repeat material that is not in FOCUS and not necessary to produce a comprehensible utterance and a coherent text. Thus, speakers normally do not utter two lexically specified objects of a di-transitive verb if one of them is given. Actually, the spontaneous
speech corpus contains only one instance of two lexical object NPs at all. In all other cases one of the objects is necessarily coded as a pronominal suffix due to its givenness. Experiments can, however, tell us what kind of strategies speakers use when forced to fulfil an unrealistic task. And there we may also discern language specific differences. What clearly distinguishes EA from the familiar West-Germanic languages is that deaccenting is not a preferred option and that strong downtoning is also not always opted for. Looking at the spontaneous data reveals a strong tendency to put a major prominence in three different positions, either the right edge of an intonation phrase or the left edge or the prefinal accent. Note that this tendency is paralleled by word accent patterns across the world's languages which are predominantly paroxytonic, oxytonic or exhibit the main stress on the initial syllable of the word (Hyman 1977; van der Hulst 1999) - in line with a preference for falling feet (Hurch 1996, 2000). I strongly suspect that an experiment design with only one potential accent after the FOCUS would have resulted in a higher frequency of deaccented or at least downtoned elements with all speakers.

To summarize the results of the corpus study and the experiments reported in Sections 4.3 and 4.4, we may note that the EA data give evidence for a model that makes a distinction between the prosody of pragmatic properties, which is related to the effort code, and the prosody of pragmatic relations that is related to the frequency code. Semantic contents, the speaker for some reason wishes to put to the foreground, is given more prominence. Foregrounding or focus clearly pertains to contrastive semantic items. The results of the second experiment suggest that a focus accent also applies in narrow information FOCUS. It remains to be seen whether focus accents may also appear in wide FOCUS constructions, in cases where they cannot be attributed to contrast. The pragmatic relations of TOPIC and FOCUS have been identified with leading and closing contours, respectively.

It has to be noted, however, that this is only the prototypical way of encoding a bipartite topic-comment sentence, for instance. The examples above display such prototypical intonation contours. Research in the tonal encoding of information structure has shown that there are no invariant correlates of information structural categories. For example, Hedberg & Sosa (2008) investigated the claim that topics in English are marked by an LH* accent, using a sample of spontaneous dialogue data. They conclude that such a hypothesis is too strong and that there are no dedicated topic and focus accents. Such a position is also held by Féry (2007: 161) who argues that

[...] the phonology, syntax or morphology are unable to define information structure. It is a common mistake that information structural categories are expressed by invariant grammatical correlates, be they syntactic, morphological or phonological. It is rather the case that grammatical cues help speaker and hearer to sort out which element carries which information structural role, and only in this sense are the grammatical correlates of information structure important. Languages display variation as to the role of grammar in enhancing categories of information structure, and this variation reflects the variation found in the ‘normal’ syntax and phonology of languages.
4.5 Interactions

4.5.1 The interaction of pragmatic relations and discourse functions: FOCUS, TOPIC, continuation and other issues

Féry (2007: 172-173) specifically argues that it is misled to identify topics and foci with specific tonal contours because of the instability of this relation. While I agree with the contention that information structure is no unambiguously coded by the grammar, I will assume that the correlation between specific prosodic contours and pragmatic relations is the default case in EA. I.e. everything else being equal, TOPICS and FOCI are encoded in the predicted way. But the crux of intonational categories is their multi-functionality. In accordance with its iconic motivation a rising contour may not only be a cue to the aboutness function of a constituent it is associated with, thus being a cue to its topic function, but it may also indicate continuation or questioning - which are in fact all related. I assume that the conflict between different interpretations of a tonal configuration is individually solved in favour of one function that so-to-say overwrites the other functions. Take the example adduced by Féry (2007: 173) to show that the correlation of focus (i.e. FOCUS) with a certain type of accent is false.

(17)
{Wer hat wen gesehen?} (German)
(Die PRÄSIDENTIN)_{p} (hat den DEKAN gesehen)_{p}.
the president has the dean seen
‘The president has seen the dean.’

Féry argues that while the second focus accent is falling, the first one is indeed rising. She concludes that the direction of the pitch accent is only related to the position of the accent in the sentence and not to its information structural function. Although it is certainly true that initial position promotes the occurrence of a rising accent and final position the occurrence of a falling accent as a result of what Bolinger ascribed to the dichotomy of tension and relaxation, it is equally true that falling accents may occur in initial position and rising ones in final position. For English, it has been noted by a number of scholars (Jackendoff 1972; Brazil 1975, 1997; Büring 1997) that if the first accent is associated with the FOCAL part and the second one with the TOPICAL part of the sentence, the typical rise-fall intonation of the hat pattern is reversed as shown in the following example from Brazil (1997: 83), where the TOPIC-FOCUS relation is reversed in the sentences in (18b):

(18)

<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>a.</td>
<td>MARy BROWN</td>
<td>is a TEACHer</td>
<td>and</td>
<td></td>
</tr>
<tr>
<td></td>
<td>MARy BROWN</td>
<td>is a TEACHer</td>
<td></td>
<td></td>
</tr>
<tr>
<td>b.</td>
<td>MARy BROWN</td>
<td>is a TEACHer</td>
<td>and</td>
<td></td>
</tr>
<tr>
<td></td>
<td>MARy BROWN</td>
<td>is a TEACHer</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
The sentence in (18) could arguably also be uttered with two rise-falls. To my intuition this invokes a slightly different interpretation in terms of information structure. While clearly both articulations involve two accents, the rendition suggested in (18) also involves a topic-comment articulation. Such an analysis of course is only possible if we assume that topic and focus - according to terminology used here: TOPIC and FOCUS - are not mutually exclusive. Note that the sentence has two presuppositional structures as it also involves two predications, (i) one being about the president with the presupposition die Präsidentin hat X gesehen (the president has seen X) and (ii) one that says that it is the president who has seen someone with the presupposition X hat den Dekan gesehen (X has seen the president). I therefore assume that X is a topic and a FOCUS at the same time. I further assume that the speaker is the one who chooses which of the relations to put into the foreground, whether to encode the topicality relation between the referent and the proposition or to encode the rhematic function of the expression die Präsidentin (the president). We will be looking at topicality in more detail in Chapter 5. Furthermore, it is possible, although mere speculation at that point, that there is a difference between an ordinary topic-comment structure without the subject being in FOCUS and the structure in 15. As a native speaker of German I would interpret the three answers in (19) differently.

(19)

a. Q: Wen hat die Präsidentin gesehen?
   (who has the president seen)
   A: Die Präsidentin hat den DEKAN\ seen.
   (the president has the DEAN seen)

b. Q': Der General hat den Oberst gesehen und die Kanzlerin hat den
     Staatssekretär gesehen. Aber wen hat die Präsidentin gesehen?
     (the general has the colonel seen and the chancellor has the state
     secretary seen, but whom has the president seen?)
   A': Die PRÄSIDENTIN hat den DEKAN\ gesehen.

c. Q'': Wer hat wen gesehen?
   (who has whom seen)
   A'': Die PRÄSIDENTIN hat den DEKAN\ gesehen.

While die Präsidentin in (19a) is likely to be deaccented being a ratified topic (noted by the underscore sign as a symbol for a linking topic), it is a contrastive topic in (19b) and therefore most likely to be spoken with a rising accent as suggested by Büring (1997, 2003). At the same time there supposedly is a difference in prominence between the two accents, the FOCAL accent being the stronger one and the topical accent being less prominent as indicated by the small caps. (19c) is a repetition of (17), quoted again here for comparison. Note that the suggested contours in A’ and A” are almost identical, only differing in the prominence relations between the two accents. These suggestions for German in fact only rely on native speaker intuitions, but let us now turn to EA.

The EA corpus shows numerous cases of rising FOCI many of which seem to be conditioned by the lack of completion of the utterance. This is especially the case in the QUIS corpus, where most tasks
are basically descriptions of pictures or short scenes. In these descriptions we mostly encounter a listing intonation as in (20) and (21). But interestingly, the rises found are more pronounced than expected for ordinary continuation rises. The examples, however suggest that a high rise, like the Glasgow and Belfast varieties of British English (Cruttenden 1986, Hirst 2013), may also be used without the implication of continuation (20). We have already alluded to that intonational type in Chapter 2.1. High rising focus accents have also been attested in Spanish narrow FOCUS (Face 2002). There is some indication that with younger speakers the rising terminals are supplanting a former low fall as used by older speakers. In the corpus a clear difference can be observed between male speakers of advanced age who hold a certain social position and young students. While the former not only use significantly more low terminals, they also use significantly more closing accents all throughout their utterances, which adds a decidedly authoritative notion to their speech. This is a case of interaction between the paralinguistic and the linguistic use of the same intonational feature, namely high or low tone in accordance with the iconic meanings of the frequency code - a fact that goes to show how difficult it is to disentangle the linguistic and the paralinguistic in intonation.

(20) 28ARZ-A01GNT-00F130

\[ajwa jaif-a:-ha\]

\[H \quad L^H \quad H\%

yes see.PTCP-SG.F-3SG.F

'Yes, I see it.'

Example (20) is from a map task. One participant asks the other one whether she sees a certain mark on the map and the other one answers 'Yes, I see it.' There is nothing supposed to follow - at least in her own speech - and the first speaker quite naturally takes the turn, and it is not meant as a question either. Furthermore, a positive answer is the prototype of an assertive utterance and a falling contour would definitely be appropriate. But probably, we are faced with another case of overwriting intonation here. The rise could be interpreted as a signal to offer the turn to the other interlocutor implying something like: 'Yes, I see it. What am I to do next? Tell me!'

In the next example (21) from the QUIS corpus the speaker has to describe a certain geometrical form.

(21) 55ARZ-A01GNT-05F1301 (QUIS corpus)

\[huwwa takwi:n \quad /jakl-u\quad jaifi\quad zajja\quad :\quad sˤali:b /\]

3SG.M formation / form-3SG.M FILL like ah cross

\[L^H\quad L^H\]

\[14\] This observation was suggested to me by Mushira Eid (pers. comm.).
\[15\] A striking example is an interview conducted by the famous actor Hussein Fahmy (born in 1935) with the young Egyptian journalist Marwa Rakha (born 1974) or the corpus of television broadcast \[ʔumsiija \ ʔaqa\afija\] recorded in the early 1990ies presented by Faruq Shusha, analysed in Rastegar-El Zarka (1997).
and-last piece in DEF-cross / the left(one)
L^H L^H

nazl-a l-taht
go.down.PTCP-SG,F DEF-down
!H L*%
'It is something shaped like a cross and the last piece on the left side is going down.'

The utterance involves three FOCI takwi:n 'formation', zajjə s'ali:b 'like a cross' and nazlə l-taht 'going down' of which only the last one has a falling contour. The example illustrates the same situation as in the German example (17) above for the constituent takwi:n which is the FOCUS of a sentence and at the same time functions as a left-dislocated topic for the next proposition fakl-u zajjə s'ali:b 'its form is like a cross', which also ends in a rising contour to indicate continuation.

Phonetically speaking, the FOCAL accents are highly prominent, the prominence being brought about by a high rise followed by an IP boundary. Consequently this high rising accent may be analysed as involving a boundary tone. While the rises in (21) are on the final stressed syllable where a boundary tone may only be assumed to exist underlingly, the rises in (22) clearly show the occurrence of a final H resulting in a rise on the final (unstressed) syllable (Figure 21).

(22) 35ARZ-C09GNT-04M1200
ra:gil / wa:xid is-siga:ra / wi-bi-j-wallaʃ-ha / man take.PTCP DEF-cigarette and-IND-3SG.M-light-3SG.F
bi-l-wallaʃ-a / wi-tˤ-t'afa:j a / ma-fi:-haʃ
with-DEF-lighter and-DEF-ashtray NEG-in-3SG.F-NEG
wala ha:ga
NEG thing
'A man... has taken the cigarette... and lights it with the lighter ... and the ashtray has nothing in it.'

We can make some interesting observation in this example: First, every proposition is phrased in a separate intonation phrase (cf. 4.5.3); second, the first three phrases clearly have one accent that is more prominent than the others which is the last one in the phrase (cf. 4.5.2); third, the final most prominent accent in these is always on the referent that is newly introduced into the utterance (cf. 4.5.2). The first phrase introduces the topic of the next two clauses that remains implicit there. The sentence topic is changed in the final clause, where the referent of t'afa:j a 'ashtray' is articulated as a left-dislocated topic that re-appears as a resumptive pronoun within the clause. This topic is not FOCAL as it has been introduced in the preceding picture. The fourth important observation thus is that the topical rise – at least in this case – differs from the FOCAL rise being much shallower, not involving a boundary tone (the H is reached within the stressed syllable and that level is maintained to the end of the constituent. Also note that there is no IP boundary between the left-dislocated topic and the core clause.
It seems, however, that a TOPIC accent may be rising even higher than a preceding FOCUS accent, if it is associated with the same term as the preceding FOCUS. It is a common feature in EA that a FOCUS constituent is literally repeated as the topic of the following sentence. The examples illustrating this fact permit the observation that the two high rising contours may differ perceptually. What distinguishes the topical rise from the FOCAL one is the scaling of the H tone associated with the stressed syllable, which is much lower in the FOCUS accent than it is in the TOPIC accent. The same is true if the FOCUS is repeated for the purpose of continuation, a so-called second occurrence focus as in (23). The example is displayed in Figure 22. It shows that the second rise is somewhat higher than the first one, but it crucially differs from the latter in the way the fact that the rise is covering the stressed syllable.

(23) (fi: hitta) min il-tuffa:ha - qitʕa min it-tuffa:ha ...
    (EXIST piece) from DEF-apple - piece from DEF-apple
    '(there is a piece) of an the apple - a piece of the apple...'

Such facts suggest that it might not only be relevant that a tone is higher than another one within the same reference frame, but also how much it is higher, which is equivalent to including register differences into the phonological representation of accents as suggested earlier by Ladd (1994).
To sum up the results of this section, we have seen that the default articulation of TOPIC and FOCUS, namely a leading or a linking contour and a closing contour, may be overwritten by other functions of intonation, be they linguistic or paralinguistic. I have further concluded that there is a difference between closing and leading accents in initial position. In the leading topics, the aspect of topicality is encoded by a rising contour, whereas in the falling topics, which may be FOCAL at the same time as in example (17) from German rhematicity and assertion is overwriting the topic function. Another issue is the observation that rising FOCI seem to differ phonetically from rising TOPICS by being more prominent, the rise being significantly higher involving a high boundary tone. This takes us to the issue of focus discussed in Section 4.3.1. The data suggests that the rising accent in the FOCUS involves a focus of interest that is probably not present in all types of rising TOPICS. This proposal is, however, not unproblematic and the problem of defining the relationship of focus, FOCUS and TOPIC will be dealt with in the next section.

4.5.2 The interaction of pragmatic properties and pragmatic relations: focus, FOCUS and TOPIC

In Section 4.3.1 we have seen that there is indeed an EA equivalent to the contrastive narrow FOCUS cases of English, German and Dutch that have been discussed in the literature for decades - with the remarkable difference that what is deaccented in the latter languages is normally not deaccented in EA. Nevertheless we have seen that there are huge prominence differences between accents in EA that have exactly the function of accenting vs. deaccenting in English, namely to make the narrowly focused item more prominent and to downtone presuppositional material that is taken for granted. Thus, we may claim that downtoning in EA is equivalent to deaccenting in West Germanic languages, under the assumption that EA word accents almost always carry a pitch accent, while English word stress is not necessarily associated with a tonal correlate. The main difference between a language such as English and a language such as EA is thus not whether it allows deaccenting or not, but rather what it is that may be downtoned. The data presented in 4.3.1 have shown that givenness is not a sufficient condition to trigger downtoning in EA. Prior mention alone does not lead to the downtoning of information, if it is in the FOCUS domain of an utterance. In Hellmuth's experiment reported above one picture shows a blue triangle and another picture a blue crescent. Note that there are two different referents involved that accidently have the same colour. But as it seems, in EA this fact is not sufficient to put important information to the background in order to stress the difference in shape taking the sameness of the colour for granted. In EA, it seems, the whole event is rather viewed as newsworthy information - which it in fact is. Nothing in the triangle's being blue is a sufficient condition to presuppose that the colour of the crescent should also be blue. On the other hand, if already mentioned information is in the presupposition, it may be downtoned in the described way. As we have seen, there is one case where information in the FOCUS domain may be downtoned: in the

212
experiment reported in 4.3.1.1 above correction clearly evoked a special emphasis on the corrected item. Let us assume that *blue crescent* is a correction of a false assumption *blue triangle* in a context such as 'Do you have a blue triangle on your card?' and the answer is 'No, I have a blue CRESCENT.' In this case the two expressions have the same referent and the fact that it is blue may be taken for granted, as it is not corrected by the second speaker. This is a case where EA would also resort to downtoning the already mentioned information and emphasize the new one. Such cases have been analysed above as involving a *focus* within the larger *FOCUS* domain. (24) shows a corrective focus of exactly this kind. The assumption was that there is a piece of chocolate and a piece of cheese involved and the answer is 'No, TWO pieces of chocolate'.

\[(24)\] 26ARZ-D03FCT-03M1100

\[
\text{laʔ [ITNE:N] foc } |fukula: t'a| foc
\]

no TWO chocolate

'No, TWO pieces of chocolate'

Fig. 23: Utterance with a focus within a FOCUS domain: \text{laʔ ITNE:N fukula: t'a 'no, TWO chocolates'}

In this example the given referent 'chocolate' is clearly downtoned by the use of a downstepped accent. The question that poses itself in this respect is the following: If what we dubbed a 'focus of interest' only occurs adjacent to a downtoned constituent, isn't it enough to represent the downtoning of that constituent? Such attempts have been made in the literature (esp. Schwarzschild 1999) to get rid of the problematic issue of focus marking. If, on the other hand, such highlighting of referents or concepts also occurs without surrounding downtoning, we will have to resort to focus marking, for example when a *TOPIC* constituent involves one or more foci as for instance in the case of contrastive topics as suggested in Chapter 1 and Section 4.5.1 above. But not only clear cases of contrastive topics may involve special emphasis but also cases of topic changes as shown in the following examples from a frog story (23).
The three topics in (25) all involve switch reference. They are, however, associated with three different configurations: a high rising contour on the demonstrative in (25a) and a closing contour in the second clause of the example\(^\text{16}\), while (25b) is a case of double accentuation - a high rise with a following downstepped accent on an otherwise stressless syllable. Double accentuation is a frequent device of focus marking as already shown by Rastegar-El Zarka (1997) for Egyptian MSA. The first topics are clearly deviant from default as they involve a (strong) accent on a function word while in the third case the double accent is associated with a proper noun that would by default carry an accent anyhow. Thus we may say that every one of these types of topic marking clearly is a deviation from default. The three topics are perceptually highly prominent, the most prominent being the third one - not only because it involves double accentuation, but also because of the following break.

The examples cited so far may be thought of in terms of structure, such as the marking of topic change or narrow focus or contrast. The next type that clearly involves added prominence does not seem to have any structural motivation. Its motivation seems to be nothing than simple foregrounding of the denotation signified by the word carrying the focus accent, i.e. in some sense simple emphasis. Emphasis, however, can hardly be viewed as a linguistic concept, and it is precisely for this reason that Bolinger regarded the focus accent to be unpredictable. In experimental data, such cases are not particularly frequent as experimental data mostly exhibit detached and so-called 'neutral' intonation. But as soon as one investigates spontaneous speech such cases abound. No structural reasons can be adduced to explain the special prominence given to the 'words' in the following examples - only that they are related to concepts the speaker wishes to emphasize. The first type of focus accent is found

\(^{16}\) It has to be noted that hijja 'she' could be associated with a rise or a rise-fall but the demonstrative in (25a) is rather too short to easily accommodate a rise-fall.
within frames, both adverbial phrases and clauses. The expression of initiality or anteriority ʔawwil 'first' and ʔawwil-ма or ʔawwil lamma 'before (conj.)' in lively speech almost inevitably carries (one of) the strongest accent(s) in an IP (26).

(26)

a. Mar_Shab_FG3_20

\[\text{fa-ʔawwil} \quad \text{jo:m} \quad \text{li-jjaj} \quad \text{fi-g-gam\dagger}\a\]

so-first day for-1SG in-DEF-university

LHL LHL LH LL*L^H%

'so, the first day for me at the university...'

b. Mido_02_F0_08

\[\text{ʔawwil-ma} \quad \text{ra}:h \quad \text{na:t'it\i'} \quad \text{in-nat'it\'a} \quad \text{kibi:r-a}\]

first-SUB COV jump.PTCP.SG.M DEF-jump.MAS-IN DEF-big-SG.F

LH*L LH L^HL L^H*L L^H-

'as soon as he set about jumping (the big jump)'

\[\begin{array}{c}
\text{Time (s)} \\
0 & 1.327 \\
\end{array}\]

\[\begin{array}{c}
\text{Time (s)} \\
0 & 2.344 \\
\end{array}\]

Fig. 24: Adverbial phrase (panel a) and adverbial clause (panel b) as frames with initial focus accents.

The initial accent in panel (24b) is especially emphatic due to the very low and rapid fall, the material afterwards is typically downtoned being an exact repetition of the FOCUS in the previous clause. As an aside, it may be noted that the continuation rise is not only indicated by the right boundary, but stretches out over the last to accents that are gradually upstepped with respect to the preceding ones. Thus (24b) could also be analysed as a case of narrow focus with downtoning of the immediately adjacent accents due to givenness, we cannot adduce the same explanation for (24a). We could, of course, argue that only the second case involves a focus accent, while the first case is only the result of emphasis. Phonetically speaking, the difference is a gradient one in any case.

The following example (27) shows a focus accent on the nominal predicate mabsˤu:tˤa 'happy' followed by a less prominent complement clause. The information in the complement clause is new, at least in the sense of not having been mentioned before. But we could argue that the information is presented as presupposed by the speaker as it refers to Mido and his dog who have been pushed into the water by a deer who is now happy because she defeated them.
If we have a look at the focus accents in (26) and (27) we find that they all are closing accents as far as their tonal shape is concerned. A closing accent that involves a full rise-fall associated with little segmental material sounds more prominent by the mere fact that it involves two contours (a rising and a falling one). In some of these accents the prominence is even enhanced by early alignment and low scaling of the closing L and/or by 'punctual' (i.e. within the stressed vowel) alignment of the H in addition to its high scaling. All this makes a focus accent differ from default. But there are also other strategies to enhance the prominence of an item, for instance double accentuation or accentuation of a functional element (28) or prolong duration (29).

(28) 

a. Mido_02_F0_09

\( w\)-\textit{mabsut\textlil{a}n} \quad \textit{u-farhanil{a}n} \\
and-happy.PL \quad \quad \text{and-joyful.PL} \\
\text{^LHLH} \quad \quad \text{^LH^L!HL} \\
'and they are happy and joyful'

b. Mido_01_F0_05

\( w\)-\textit{\textlil{a}mmil{a}mil} \quad \textit{ji-hamil} \\
and-make.PTCP.INT.SG.F \quad \text{3SG.M-try} \\
\text{L}^\text{L} \quad \text{H}^\text{L} \quad \text{H} \quad \text{H}^\text{L} \\
'and he keeps trying...'

c. Mar\_HF\_FG3\_01

\( a\)\text{na} \quad \text{mi}f/ \quad \text{raf}d\text{-a} / \quad \text{wug}d/d \quad \text{ra}gil \\
1SG \quad \text{NEG} \quad \text{refuse.PTCP-SG.F} \quad \text{presence} \quad \text{man} \\
\text{L-} \quad \text{H}^* \quad \text{H}^* \quad \text{LH} \quad \text{H}^\text{L} \\
'I don’t REFUSE taking a husband' (lit. the presence of a man)

d. Mar\_Siq\_FG3\_06

\( w\)-\textit{\textlil{a}tittajjar\textlil{a}} \quad \textit{ma\textlil{a}} \quad \textit{ba\textlil{a}d\textlil{a}} \quad \textit{f-nafs} \quad \text{il-ittija:h} \\
and-IND-change-3PL \quad \text{with} \quad \text{part} \quad \text{presence} \quad \text{man} \\
\text{LH} \quad \text{^LºH} \quad \text{^LºH} \quad \text{LºH} \quad \text{H}^\text{L} \\
'and you change together in the same direction'

Fig. 25: Focus accent in the FOCUS domain, indicated by a closing accent with high pitch excursion.
(29) Mido_01_F0_05

\[ \text{ʃa:f} \quad \text{xalijja} \quad \text{maljaːːna} \quad \text{nahl} \]

see3.SG.M honeycomb full bee.COLL

\[ \text{^LH} \quad \text{LH} \quad \text{L^H-} \quad \text{!H'L%} \]

‘he saw a honeycomb FULL of bees’

Fig. 26: Emphatic accent in the FOCUS domain, indicated by high scaling and long duration.

Fig. 27: Emphatic accent in the FOCUS domain, indicated by double accentuation.

These facts raise the rather problematic issue whether it is possible at all to distinguish between focus as a quasi-linguistic concept and emphasis as a clearly paralinguistic one. We will return to that issue in section 4.5.4 below. In sum, there are three important questions that arise regarding the notion of focus:

i) Is focus structure-based?
ii) Are focus and contrast the same? and
iii) Can we distinguish focus from ordinary emphasis at all?

Bolinger’s and Halliday's notions of focus is obviously related to what is called focus here, while our TOPIC and FOCUS are clearly related to Halliday’s Theme and Rheme. As already mentioned, in many recent theories the two dichotomies are collapsed into one dimension (Lambrecht 1994; Vallduví 1992; Vallduví & Engdahl 1996). The present work opts for a distinction between the two dichotomous basic notions. We have already sorted out TOPIC-FOCUS structure and come to a conclusion as regards their prosodic coding. The concept of focus, however, presents itself as notoriously difficult to capture. A view into the literature on the topic confirms this impression, researchers have dealt with focus rather uneasily. While older theories simply equate focus and accent (Halliday 1967; Selkirk 1984, 1995; Rooth 1985, 1992) - all dealing with English - this equation has more recently been subject to criticism (Schwarzschild 1999; Büring 2003). Before we turn to

217
questions (ii) and (iii) dealing with the relationship between focus, contrast and emphasis in section 4.5.4, we will look at the 'interaction' that has been among the main issues in intonational phonology: focus and phrasing.

4.5.3 Focus, FOCUS and phrasing and the myth of the nucleus

In the British tradition of intonational phonology it has been assumed that every intonation phrase has a nucleus, which is supposed to be the strongest and at the same time the final accent in an IP. On the other hand, in the American tradition there is no such assumption and consequently Bolinger questions the concept of the nucleus. He notes that a primary accent is not necessarily always the most prominent one in its domain (Bolinger 1986: 58ff.). On the other hand, there is a tendency to put the most important information - and thus the strongest accent - in utterance final position because this position is psychologically advantageous in perception (Bolinger 1986: 48f.).

In one line of AM theory, albeit the early conception of it (Pierrehumbert 1980), that followed Bolinger in that respect, the nucleus was simply the final accent in a phrase with no special phonological importance. In another line (Ladd 1996), the nucleus was introduced into AM theory as a phonological concept, dividing an intonation contour with more than one accent into prenuclear and nuclear accents. I believe I am not mistaken to say that most models that deal with the prosody of focus (FOCUS) today implicitly or explicitly make this assumption.

It is also commonly assumed that the nucleus is the one accent that indicates FOCUS. In Halliday's (1967) framework what is referred to as nucleus here is called the tonic (Halliday 1967: 203). Halliday views the tonic, i.e. the focus accent, to be reflecting "the speaker's decision as to where the main burden of the message lies” (p. 204). He also allows for the domain of the focus to "extend over the whole of the information unit" (p. 243). In early generative models (Chomsky & Halle 1968) the nuclear stress rule (NSR) determined the position of the nucleus by rule as the final element in a syntactic domain which may project focus to the whole domain. As a result, it has commonly been assumed that there is marked and unmarked focus (FOCUS) structure. While in the unmarked case the nucleus is placed at the end of the domain and may project focus (FOCUS) to a higher domain, an early nucleus immediately means that the focus (FOCUS) is narrow. To that effect, Ladd (1980: 75) states that "while most of the possible accent placements in a sentence signal narrow focus, one leaves the focus broad or unspecified”.

Gussenhoven (1983), in his seminal paper Focus, mode and the nucleus, lies the foundation of what is still viewed as the Standard view of focus-to-accent structure today, arguing for the rule-governed placement of the nucleus on the assumption that one of the chief functions of its location is "to signal
the focus distribution of the sentence” (p. 377). The Standard view of focus projection is expressed in Selkirk (1995: 555), given in (30).

(30)

**BASIC F-RULE:** An accented word is F-marked.

**F-PROJECTION RULES:**

a. F-marking of the head of a phrase licenses the F-marking of the phrase.

b. F-marking of an internal argument of a head licenses the F-marking of the head.

Selkirk’s rules do not essentially differ from Gussenhoven’s rules, only that her algorithm works in a bottom-up manner, deriving focus projection by rule, while Gussenhoven's SAAR (Gussenhoven 1983; cf. Chapter 3) could rather be described as a top-down rule that defines the focus domain first and places the nucleus accordingly. On such a view, focus ‘projection’ is rather a matter of interpretation. Be that as it may, both models rely on two main assumptions: (i) that the nucleus or focus accent may signal focus (FOCUS) on a larger domain and (ii) that the nucleus falls on an argument and not on the predicate in the unmarked case. In recent publications, these focus projection rules have come under criticism (Schwarzschild 1999; Büring 2006; Féry & Samek-Lodovici 2006), but the main assumptions are not questioned.

In the present work, I take a different approach, at least to some extent following Bolinger’s radical focus-to-accent approach. In the present approach, intonation phrases are not supposed to be defined in structural terms, i.e. they are not necessarily co-extensive with syntactic phrases. They are identified independently on solely phonetic criteria (see 2.2.6). On the other hand, focus is neither identified with respect to a syntactic phrase, nor with respect to an intonation phrase. It is assumed to be directly influenced by the semantic and pragmatic weight of a concept, i.e. mostly the denotation of a lexical item. Let us now look at the data for both assumptions in turn.

In section 4.3.1 above, we have already noted that deaccenting is not allowed or at least dispreferred in some languages - which is in fact equivalent to the claim that the nucleus will have to be located phrase-finally. According to Ladd (1996: 177), Italian is such a language. Under the assumption that prosodic marking essentially is marking via accentuation, Vallduví (1990) proposed his typological distinction between plastic accent and non-plastic accent languages, referring to languages such as English and German as plastic accent languages that exemplify a considerable freedom of nuclear accent placement, whereas many Romance languages fall into the category of non-plastic accent languages which are characterized by edge-marking nuclear accents. As Vallduví (1990) notes, non-plastic accent languages rely on syntax instead for marking information structure. The similarity of Italian and EA in terms of syntax and prosody has been pointed out by Hellmuth (2010a). On the other

---

17 For more details on the different models the reader is referred to the literature overview in chapter 3.
hand, Face & D’Imperio (2005), based on data from Italian and Spanish, also show that the syntactic vs. intonational focus marking categories are really a typological continuum, showing that the traditional assumption that Italian and Spanish mark focus only syntactically cannot be maintained in the light of the more detailed investigation about the focus-intonation interface in these languages. As EA word order is also comparatively flexible, it could be assumed that EA is a non-plastic accent language - a position implicit in Hellmuth’s work.\(^\text{18}\) The question is, then, if EA belongs to the languages of type A or to the languages of type B in Ladd’s suggested dichotomy (Ladd 1996: 167)

\[
\begin{array}{ll}
\text{Language A} & \text{Language B} \\
\text{This is a book RED} & \text{This is a BOOK red} \\
\text{I bought car NEW} & \text{I bought CAR new} \\
\text{He has nose BIG} & \text{He has NOSE big} \\
\end{array}
\]

We have already seen above in section 4.3.1 that EA definitely is not a language of type A in the sense that this would be the only occurring accentuation pattern. In fact, in EA both types of accentuation may occur. In this section, I will adduce more examples to show that a focus accent - but note that the concept of focus accent is different from that of the nucleus as defined above - may occur in various positions within a phrase. In addition, I will shed more light on the relation between focus and phrasing as suggested by the corpus data.

\[
\begin{array}{llllll}
\text{Mido}_02\_\text{F0}_03 & \text{hazin-a} & \text{nnø} & \text{mi:du} & \text{ha-ju-xrug} & \text{wi-j-sib-ha} \\
\text{sad-SG.F} & \text{COMP} & \text{M} & \text{FUT-3SG.M-go.out} & \text{and-3SG.M-leave-3SG.F} \\
\end{array}
\]

‘sad that Mido was going out and leaving her’ [being said about a turtle]

(32) is an example of a single focus accent in the initial position of an IP (Fig. 28). The focus accent is made prominent in two ways: firstly, it has by far the largest pitch excursion within the IP and secondly, it departs from a default accent concerning the alignment of the second L tone which is at the end of the focused word and the beginning of the complementizer instead of being situated at the beginning of the following content word. In the default case the functional item would be integrated in the fall, but in the present case the functional element carries a linking tone, which means that the two accents are not linked in the sense defined in chapter 2, but separated. Furthermore, the constituent is the one with the strongest intensity within the utterance.

\[\text{\textsuperscript{18} However, Hellmuth ((2010a)) argues that there is no functional complementarity involved in marking thetic sentences in EA.}\]
In terms of information structure, the focus accent highlights the core of the predicate phrase with subsequent downtoning of the complement clause. It is important to note that both clauses contain accessible information (we will return to that aspect in Section 4.5.5 below). The whole utterance is appositive to a foregoing one that states that the turtle was sad (this time zašla:na) and pulling her head in. So in fact the focused word could even said to be given (as the two expressions are synonyms denoting the same concept) while the contents of the complement clause is predictable from the context. This example, as much as many others shown so far, is evidence against a range of assumptions that have been made: It shows that expansion of pitch range and compression of pitch range are not only found in cases of contrast. It also shows that focus is not necessarily associated with what is informationally new and downtoning with what is informationally given. We may thus conclude that focus is a category in its own right independent from TOPIC or FOCUS and also independent from information status or the given-new distinction.

The examples illustrated in the previous section give evidence that focus in EA may occur in any position in an IP: the default position clearly is the phrase-final one (e.g. 28c); another frequent position is phrase-initially (25, 26, 32), and finally focus may also occur in the middle of a phrase (27, 28d). It could be argued that (27) and (32) are in fact examples for a phrase-final position if phrases are defined in syntactic terms, i.e. any kind of smaller IP or prosodic phrase, such as the minor phrase (Selkirk 2000) or Truckenbrodt’s p-phrase (1995, 1999), the position would indeed be phrase-finally. Well, firstly, this phrasing definition suffers from circularity because it assumes a boundary of a prosodic domain to be co-extensive with the end of a functional domain to argue that the functional domain is delimited by a prosodic boundary. And secondly, there is plenty of evidence against it, if one is not to throw any phonetic evidence for phrasing overboard. Consider examples such as the following:

(33)

a. Mido_02_F0_03
ra:ħ na:tˤitˤ / ?ig-ge:b bita:ʕ mi:du
go.3SG.M jump.PTCP.SG.M inside / DEF-pocket POSS M
L- LHL H- LHL- L°H°L
‘and then he jumped into Mido’s pocket’
In (33) the phrases containing a focus are split by a prosodic boundary. This is in fact not a production error or caused by a hesitation pause, but a frequently employed highlighting strategy. In (33a, c), the split is between the prepositional head and the noun, whereas in (33b), the split is within an annexation between the two nouns. In both cases the phrasal heads have their own accents. As we have already seen before, an additional accent on the preposition of a PP adds to the prominence of the whole PP - and a boundary further enhances that prominence.

This is not to say that syntactic phrases and prosodic phrases do not normally coincide. Their coincidence, in fact, is a matter of common sense, if information is to be structured in a reasonable way. The point is, however, that they belong to different parts of grammar that interact without one of them being completely determined by the other. Coincidence thus is default, flouting the default is marked and thus enhances prominence, thereby offering a good strategy for focus marking. Phrasing thus is a means of foregrounding important information. Furthermore, phrasing is intimately related to speech style and tempo, which in turn depend on discoursal factors such as situation or genre - a heated discussion is supposed to differ greatly from a news report broadcast on the radio. Speech style and tempo also differ from one person to another.

The EA corpus contains incidences of all different sizes of phrases. It can, however, be observed that intonation phrases are frequently quite short, containing from one or two accents, of which one - usually the final one - is normally stronger than the other. This type of phrasing is in accordance with what Chafe (1994) called the *one-idea-constraint* of informational chunking. A good example of this type of structure is example (21) above, in which almost every argument and every predicate phrase are phrased in an independent IP. In other words, the clauses are divided into a topic (first argument) and a comment, the FOCUS, each of them consisting of one or two accents. Another example is the following stretch from a frog story that shall be given here to illustrate a slightly longer stretch of discourse. Accented syllables are marked in capital letters for the discussion below.
Now, today, we are going to tell a very beautiful story...

A new adventure of Mido...

and his friend, the frog...

Mido one day...

and spent the beautiful day with his friends...

the dog, the frog and the turtle...

being very very happy’

In the first part of the quoted text, almost every concept, whether it is an event or an entity, is coded in a separate phrase. At the same time, most IPs consist of two accents, exhibiting a very strict rhythm typical of Egyptian Arabic (Rastegar-El Zarka 1997, El Zarka 2005), especially of certain conventionalized genres, such as narratives, speeches and the like. The text thus is a good example illustrating the interaction of informational and rhythmic preferences, i.e. of grammar and prosody. In rare cases we may also find a series of IPs consisting of one single accent domain resulting in a rather marked prosody to the effect that the single prominences are further enhanced (35).

A more natural rendition of the paragraph would entail the phrasing in (36a), while (36b) illustrates an example of infelicitous phrasing.
Although the phrasing in (36b) is perfect in terms of rhythm, it is odd in terms of information structure. Note that there is nothing wrong in phrasing the adverb together with only part of the topic as shown in (37), if subsequent phrasing is informationally reasonable. The examples illustrate that it is always a balanced interaction between prosodic (rhythmic) considerations and functional (information structural) considerations that produces real utterances.

So far we have not said a word about focus. In the text, I only marked the accents without specifying any focus accents. The next question now is how many foci may occur in one IP. Halliday (1967) allows for one focus (the tonic) per information unit, occasionally followed by a secondary one. Halliday's information unit is supposed to usually have the extension of a clause (Halliday 1967: 201). Gussenhoven (1983) is rather explicit about the extension of what he calls a focus domain, which he defines in semantic terms. A focus domain is characterized by exactly one accent and may have the size of an "argument", a "predicate" or a "modifier" (an attributive adjective or an adjunct), with the provision that a predicate will form one focus domain with an adjacent argument which in this case carries the accent - or the focus. Neither Halliday nor Gussenhoven were concerned with intonation phrases and their boundaries at that time. Rather, the units as they define them are tonal units, which is why Ladd (1983) suggested that Gussenhoven's focus domain should rather be called accent domain. An IP may contain one or more accent domains whose last accent is usually supposed to be the nucleus or focus exponent, while the pre-nuclear accents are supposed to be subsidiary secondary accents. We are thus left with two different definitions of focus. Why should that be so? I think, the answer lies in the nature of English prosody, which was and - albeit to a lesser extent - still is the main basis of all theorizing. We have seen above that even in EA, a pitch accent-rich language, utterances are frequently rather short and may contain only one prominent accent (cf. Fig. 21, for example), this fact is even more applicable to English which has pervasive deaccentuation - a fact that also explains Halliday's claim that an information unit is usually equivalent to a clause while at the same time mostly containing only one accent. EA, however, is clearly different in that respect. This raises the
question of whether there is a nuclear accent in the first place and whether focus is projected at all. And could it not be the case that there is more than one focus in a domain, i.e. in an IP? The issue is important enough to devote to it a separate subsection.

4.5.3.1 Focus, nucleus and focus projection

The definition of focus is notoriously difficult, both formally and functionally. We have seen that it is not possible to define focus structurally, i.e. neither information status is a sufficient condition to predict it, nor is it the case that focus is always based on explicit contrast. The question now is what makes a focus accent different from an ordinary accent if we do not want to say that every accent is a focus. Considering the examples cited so far it is clear that this cannot be the solution. But we will defer this question to the next section. For the time being let us assume we already know what a focus accent looks like.

It has become sufficiently clear by now that a focus need not always be the last accent in the phrase. It is also not the case that every accent after a focus is produced in a compressed pitch range. Consider example (38), illustrated in Figure (29): The utterance consists of two IPs with a number of focus accents. But can an IP have more than one focus? Evidence from EA suggests it can. In (38) the FOCUS domain (the second IP) contains a series of three equally strong accents - at least perceptually - and this is what counts (but note, for instance, that all three FOCAL foci align the high tone right in the middle of the stressed syllable).

(38) Mar sıq FG3.16
\[\text{in-niskafiha:}-ja \quad \text{ba\textipa{\textperiodcentered}a} / \quad \text{tafa:\textipa{\textperiodcentered}al} \quad \text{ma\textipa{s}} \quad \text{al-mawqif} \]
\text{DEF-nescafé-IN} \quad \text{SEQ} \quad \text{interact.3SG.M} \quad \text{with} \quad \text{DEF-situation}

‘Now, the Nescafé... reacted to the situation’

Fig. 29: Pitch track of a TOPIC-FOCUS utterance with several foci.

A possible argument against the analysis could be that every such accent projects an own FOCUS domain. But what would that mean in terms of FOCUS? It can hardly be assumed that the preposition alone is any kind of FOCUS - except perhaps in narrow FOCUS cases, and even there this could be
doubted. What the three focus accents indicate is a special emphasis of the FOCUS being the predicate phrase 'reacted to the situation'. Which of the focus accents would be the one to project focus? Is it the final one on the nominal part of the prepositional complement (as suggested by the rules set up for English) or is it the accent on the verb? To answer this question let us first look at some cases where only one focus accent can clearly be identified (39a and b, illustrated in Figure 30a,b).

(39)

a. Mido_01_F0_05
rax / ʔarsˤ-u min-manaxi:r-u
go.3SG.M bite.PTCP.SG.M-3SG.M from-nose-3SG.M
'So he bit him in his nose'

b. Mar_Shab_FG3_12
mihta:g-a maghu:d kiti:r ʔawi
need.PTCP.SG.F effort much very
'It needs a lot of effort'

c. Mar_shab_FG3_03
ʔiddi:n ʔaswaʔ is-sinarju:ha:t
give.IMP.SG.1SG worst DEF-scenario-PL
'Give me the worst-case scenario'

d. Mido_02_F0_11
mif ʔajz-i:n-ak ti-ʃmil-li-na fadˤajːih
NEG want.PTCP-PL-2SG.M 2SG.M-make-for-IPL shame.PL
'we do not want you to bring shame on us'

e. Recip_F3
baʃd kida [baʃk:id] ha-n-gi:b bafar-a (corr. basˤal-a)
after so FUT-1PL-bring [error] (corr. onion-IN)
'and then we take an onion' (says 'skin', but means 'onion')

![Fig. 30: Four pitch tracks of early focus accents (panel a, b) and final focus accents (panel c,d).](image-url)
If we assume that FOCUS in the above cases is not narrow, but broad, we would have to assume that it is projected by all different types of constituents, such as the verb in clause initial position (39a), the nominal argument in prefinal position (39b), the adjectival nominal head of an annexation (39c), the nominal object in clause-final position (39d, e) etc. If the Standard projection rules for English were to hold for EA, the second two cases would be fine, but the first three would not be. Furthermore, the cited cases only tell part of the story. There is variation in almost every case. For instance, the clause cited in (32) could equally well be uttered as in (40a, b) without the information structure being changed, while the utterance in (40c) necessarily implies narrow FOCUS of the subject.

(40)

a. *haz:ni-a nno mi:du ha-ju-xrug wi-ji-sib-ha*
   sad-SG.F COMP M FUT-3SG.M-go.out and-3SG.M-leave-3SG.F
   '(she was) sad that Mido was going out and leaving her'

b. *haz:ni-a nno mi:du ha-ju-xrug wi-ji-sib-ha*

c. *haz:ni-a nno mi:du ha-ju-xrug wi-ji-sib-ha*

Similarly, we find variation in the accentuation of modifiers which may either receive a focus accent as in (41) or be downtoned and integrated into the tonal domain of a focus accent on a preceding predicate or argument as in (42a).

(41)

a. *Mar_si_q FG3_08*
   *wi-da bi-j-gi:-l-i kiti:r ?AWI*
   and-DEM.SG.M IND-3SG.M-come-to-1SG much very
   LH-^H-^HL% /LH
   'Such things come in very often' (i.e. are posted in an internet blog)

(42)

a. *Mar_si_q FG3_08*
   *wu-huwa da:xil ʔala mawqif /*
   and-3SG.N enter.PTCP.SG.M PREP situation LH HL
   'When he is faced with a situation...

HA:JIB minn-u ?awi
fear.PTCP.SG.M of/from-SG.M very L^HL-%
he is very much afraid of'

b. *Mar_Shab FG3_12*
   *wu-miht:a:g-a MAGHU:D kiti:r ?awi*
   and-need.PTCP.SG.F effort much very LH L^HL-% ^H-L
   'And it needs a lot of effort'

In the face of this evidence we may ask whether the notions of nucleus and focus projection are at all useful. I suggest they are not. Another piece of evidence against a 'nucleus' projecting FOCUS is the
fact that a 'nucleus' may occur on all kinds of lexical items as it seems. If we were to analyse the final accent in the first IP of (33c) as a nucleus, it would be a nucleus on a preposition that is separated from its complement by an intonation break. Where then should the focus project to? And what about nuclei of IPs that contain frames or other TOPIC phrases as in the first IP of example (38)? The final accent is on the modal particle *baʔa* 'now, so, then etc.' If a nuclear prominence were to percolate up the syntactic tree, it would have to project TOPIC phrases as well as FOCUS phrases (cf. also Büring 2006 for the same argument). It thus seems that the notion of focus projection is highly problematic. But fortunately we do not seem to need focus projection - at least not for EA. Here we come back to what has been the subject of earlier sections of this chapter, where a distinction between pragmatic relations and pragmatic properties was drawn. A focus as conceived of on this view is the opposite of downtoning, namely highlighting what the speaker wishes to put to the foreground. FOCUS, on the other hand, is signalled by contour shape instead and does not need to be projected by a single focus accent.

Let me once more quote Bolinger (1989: 363) here who admittedly based many of his claims on introspection only. But it may be noticed that Bolinger's introspection is much more reliable than the statistically significant results of many of the quantitative analyses that are so popular today. Bolinger (ibid.) claims that any of the five possible focus (i.e. main accent or nucleus) positions given in (43) (Bolinger's examples 30 and 31) is suited as an answer to the given question and thus is apt to signal broad focus. Comparable examples from EA have been given in (41) and (42).

(43)
A: What is wrong with this society?  
B: The average American expects too much of people.  
a) The average American expects too much of people.  
b) The average American expects tóo much of people.  
c) The average American expects too múch of people.  
d) The average American expects too much óf people.  
e) The average American expects too much óf péop.  

Recently, some researchers working within the generative framework have also expressed doubt about the validity of the Standard focus projection rules even for West-Germanic languages (Schwarzschild 1999; Büring 2006; Féry & Samek-Lodovici 2006). From a functional point of view, Lambrech (1994: 304) suggests that any accent within the VP could project (unmarked) "predicate focus" while an accent on the subject is the marked case and may either be interpreted as "subject focus" or "sentence focus". My view - regardless of the different theoretical underpinnings - concurs with Büring (2006) who also suggests that focus projection rules are unnecessary. He interprets Selkirk's focus projection rules (given in 30 above) as "vertical" (2a) and "horizontal" (2b) focus projection. The "horizontal" projection rules are in fact similar to what Fuchs (1976, 1980) and Jacobs (1993) called integration. I will return to the important notion of integration in 4.5.5. Büring (2006) believes the vertical rules to be unnecessary, also for English. The conclusion he draws from this analysis is that
the accents are placed by more general rules of prominence assignment. On this view, focus and accent are not the same thing.

The one-to-one relation of focus and pitch accent was already called into question by Jacobs (1988: 116) who noted that an accent on a topical constituent could be viewed as a feature of sentence phonology to make the sentence rhythmically more balanced. He also suggested that an early accent in exclamative sentences could not be explained by the presence of alternatives and that a left-dislocated TOPIC or theme is also necessarily marked by a thematic accent (p. 115). Similarly, Féry & Samek-Lodovici (2006: 131) refute the idea of

pitch accents and focus as directly related to each other via an intermediate distribution of abstract F-marks which link to pitch accents on one end and to the discourse status of syntactic constituents on the other. In this discussion note we argue for the opposite view, showing that pitch accents are unrelated to F-marks and that instead their distribution follows ENTIRELY from the interaction between the constraints that govern the prosodic organization of the clause and the constraints STRESS-FOCUS and DE-STRESS-GIVEN that govern the prosodic expression of discourse status.

In other words, more recent work on the prosody of focus views it as an interaction of information structural and prosodic, i.e. rhythmic, principles, which is in fact the position I argued for in chapter 2. The notional difference between Féry & Samek-Lodovici’s proposal and my own is that I do not believe that focus marks “ENTIRELY” follow from the interaction of these principles, but rather are influenced by additional considerations, such as the semantic value of individual lexical items and the speaker’s wish to emphasize a certain notion, even if it counts as given.

To summarize, the view taken here is that accent is a phonological phenomenon intimately tied to rhythm and that not every accent is a focus accent. In addition, I view focus as partly independent of information structure. It may occur within TOPIC as well as FOCUS domains to point out a major point of information. A focus accent is taken to mark a constituent with an especially high information value, at least in the eyes of a speaker at a specific point in discourse. It comes as no surprise, however, that foci are especially common in FOCUS domains. In the EA corpus, roughly 80% of all foci occurred within the FOCUS domain. An ALL-FOCUS utterance may in fact consist of a number of equally strong foci, if all constituents have an equally high information value as (44).

(44) Translation task (F0) (SFB 632, D2 project)
    fi: kalbə bi-ḥud’də ra:gil
    EXIST dog IND-3SG,M-bite man
    ‘There is a dog biting a man’
Fig. 31: ALL FOCUS utterance with three foci.

We can see that all three accents are closing accents with the typical early alignment of L2 and the linking tone between the accentual excursions. Note that the utterance is unmistakably produced within one IP, indicated by Schwa-insertion after kalb 'dog' and bi-ʃudˤdˤ 'he is biting'. The accents all sound very prominent and may thus all be analysed as foci. The problem, however, is how a focus accent should be defined phonetically and how it differs from an 'ordinary' accent? This is the topic of the next section.

Cases such (44) are in clear opposition to the notion of an obligatory nucleus or sentence accent. A sentence accent - if the term is justified at all - is clearly only one option and here again its relatedness to a whole 'sentence' is only a matter of interpretation or "focus construal" in the sense of Lambrecht (1994). The decision where to put it is determined by other considerations, a point that was also made clear by Fuchs:

'Sentence accent', then, is 'sentence' accent in the sense that the informational focussing effected by it - which may be bestowed on any of its constituents, and on several at a time - is fully interpretable only against the background of the relations the constituents stand in within the sentence. The term must not mislead us into thinking that what an accent functionally relates to is simply the sentence as a unit. If it did (in a 'delimitative' function), the rules for its placement would in all probability be simpler than they are (for how should the accent help delimit the unit, otherwise?), and an unalloyed 'syntactic' approach might have been sufficient to capture them. (1984: 162)

4.5.4 Focus - contrast - emphasis

Let us now come back to the question what a focus accent stands for. Many approaches, in the past and at present, have somehow tried to differentiate between information focus or new information and contrastive focus. At present, there are broadly two different views on that issue as far as their prosodic encoding is concerned: (i) There are approaches that view contrastive focus as non-distinct phonologically from new focus in the languages they investigated (Bolinger 1961, 1986, 1989; Jackendoff 1972; Ladd 1980; Gussenhoven 1983, 1992; Selkirk 1984; Lambrecht 1994; Büring 1997; Schwarzschild 1999) and (ii) there are those who view it as a separate category coded differently (Chafe 1974, 1976; É. Kiss 1998; Vallduví & Vilkuna 1998; Frota 2000; Selkirk 2002, 2008; Féry &
Some approaches belonging to the second group explicitly view contrast as a notion orthogonal to focus (Vallduví & Vilkuna 1998; Molnár 2006; Brunetti et al. 2010). The upshot of the most recent discussion on the prosody of focus (Samek & Lodovici 2006; Kratzer & Selkirk 2007; Selkirk 2008; Katz & Selkirk 2011), however, is that only contrastive focus is ‘(F-) marked’ at all and that ‘discourse new’ is not marked, but accented by default reflecting prosodic constituency. In these studies a distinction between ‘contrastive’ in the sense of identificational focus (É. Kiss 1998) or alternative focus (Jackendoff 1972; Jacobs 1988; Krifka 1992; Rooth 1992) and simply ‘new’ information is drawn. According to Kratzer & Selkirk’s proposal only F-marking accents contribute to the meaning of an utterance:

There appear to be two distinct sources of pitch accents in intonational languages – default pitch accents, which are nonmeaning-bearing tones epenthesized in the phonology, and pitch accents that are tonal morphemes […] In English, Dutch and German a variety of pitch-accent types have been isolated […] some of which are claimed to make a distinctive meaning contribution to the sentence […] (Kratzer & Selkirk 2007: 128; emphasis added)

In principle, this proposal is similar to the one made here, as it disentangles the notions of ‘new information’ and specially highlighted concepts. This is tantamount to differentiating between a marked and an unmarked intonation. The main difference between these proposals and the present one is twofold: (i) While the generative studies deal with the different coding of individual expressions and their denotations (e.g. Selkirk 2008) based on their information status in their distinction between ‘contrastive’ and simply ‘new’, the present model makes a principled distinction between FOCUS as coding the new information in a pragmatically structured proposition¹⁹. A new concept itself is not yet to be considered as new ‘information’ which is necessarily propositional – its informational status is only achieved through its relation to other concepts of the proposition. Thus, the present model differentiates between rhematic FOCUS as new information and focus as a highlighted concept. The commonality with the generative proposals is the observation that what is ‘new’ is not necessarily highlighted by special accents. (ii) The second difference pertains to the notion of contrast. I have argued above that a focus need not be related to contrastiveness. We have, however, seen that a focus normally evokes alternatives²⁰. The concept of alternative used here is actually broader than the concept of contrast, or more explicitly paradigmatic contrast. I assume that there is at least a conceptual distinction between the FOCUS types as answers to either question ‘What would you like to drink, tea or coffee?’ or ‘What would you like to drink?’ Drawing a conceptual distinction, however, does not mean that there has to be a prosodic correlate of that distinction (in EA). I only speak of

¹⁹ Note that this term is adopted from Lambrecht (1994) and is not to be confused with the terms structured proposition or structured meaning in Generative Semantics (Stechow (1991); Krifka (1991/92)).

²⁰ Again, the use of the term ‘alternative’ is to be taken in its everyday meaning and does not imply the adoption of semantic definitions put forward in Alternative Semantics (Rooth 1985, 1992). However, I believe that there is an essentially common idea behind ‘alternative’ as used here and the term in its strict semantic definition.
contrast if the alternative set is closed or ‘restricted’ as argued by Molnár who views contrastive focus as the “pragmatically most restricted type of explicit mentioning (or exclusion) of alternatives” (Molnár 2006: 210; emphasis added). Likewise, Bolinger (1961: 87), one of the proponents of the ‘gradient’ view, suggested that “in a broad sense, every semantic peak is contrastive” and Lambrecht (1994: 290), following him in this contention, states that there are “clear and […] less clear instances of contrastiveness, and it accounts for our intuition that the clearest instances are those in which a focus designatum explicitly contradicts a stated or predicted alternative.” How is all this related to the question of narrow FOCUS vs. broad FOCUS? On the view adopted here, narrow FOCUS as an answer to an information question is necessarily identificational. But identification is not coterminous with exhaustiveness or contrast as suggested for instance by É. Kiss (1998). Consider the following hypothetical dialogue: B’s answer does not necessarily imply that no one else went to the Zamalik club.

(45) A:  
\textit{mi:n ra:h na:di z-zama:lik?}  
‘Who went to the Zamalik Club?’

B:  
\textit{nagafa ra:hit na:di z-zama:lik.}  
‘Nagafa went to the Zamalik club.’

A:  
\textit{wi-mi:n kama:n?}  
‘And who else?’

On this assumption, there are two questions to be asked: (i) Is narrow FOCUS necessarily different from broad FOCUS? and (ii) Is contrastive and/or exhaustive FOCUS a special category within narrow FOCUS, exhibiting categorically different prosodic realization. Unfortunately, both issues cannot be finally settled at the present state of knowledge. The results of the two of experiments reported in Sections 4.3.1 and 4.4.1 do not give conclusive evidence for one solution. We have seen that correction as the strongest type of contrast typically involves a focus accent on the correcting expression, but the spontaneous data contained corrections that were not especially necessarily marked prosodically. However, there may be a difference between self-correction that need not be marked and contradiction in an interaction that is more likely to be explicitly coded. On the other hand, experiment II, which was originally designed to elicit narrow information FOCUS, yielded ambiguous results. Two speakers marked the FOCUS with a focus accent that significantly differed from default with accompanying downtoning of the post-FOCAL (presupposed) material. However, others used strategies that are not as unambiguous. Specifically, one speaker’s utterances – although differing significantly from topic-comment articulations could not necessarily be interpreted as involving a focus accent. This may be interpreted in the sense that the individual speakers produced different types of narrow FOCUS. While one speaker presented the requested information in a neutral manner, the other speakers gave it an exhaustive and more contrastive interpretation and expressed it more emphatically. As the type of FOCUS was in principle controlled for, the results lead to the conclusion that exhaustive or contrastive interpretation may arise as a conversational implicature. Judging on the qualitative corpus study, I suggest, for the time being, that contrastive FOCUS and exhaustive FOCUS
are no grammatical categories in EA, but that these cognitive notions are usually have the formal correlate of a focus accent. The challenge now is to see whether we can distinguish focus from ordinary emphasis formally. Defining the category of focus as a pragmatic property that is prosodically encoded by a focus accent calls for the phonological identification of that accent. I have tried to do so in the preceding sections by making a distinction between default accent and marked accent types.

In doing so, we unavoidably step on slippery ground, trying to introduce a category in hopelessly gradient matter. The question to ask is whether there is any difference between focus and emphasis at all. Recall that Halliday (1967: 204) defined his notion of focus as "one kind of emphasis" that "reflects the speaker's decision as to where the main burden of the message lies". A similar view is also held by Bolinger (1958, 1972). Hirst & Di Cristo (1998: 31; emphasis in the orig.) also define focus as "a more general notion [...] which attempts to account by means of a single mechanism both for the given/new aspect of an utterance and at the same time for what has traditionally been treated under a separate heading as emphasis or contrast". The linguistic problems in defining the semantics of focus are paralleled (and related to) the phonological problems in defining a special accent type for focus. On the other hand, we have seen that recently attempts have been made to distinguish between information focus and contrastive focus on a phonological basis.21

Let us first look at some of the proposals made in the literature on other languages. It has been claimed for many languages that narrow and broad focus or contrastive and information focus are articulated by categorically different accent types. Selkirk (2002) reports an experiment that investigated the prosodic correlates of "contrastive FOCUS" vs. "presentational focus" in English. Selkirk (2002: 644) identifies a medial “FOCUS” accent as most commonly L+H* (82%) followed by a low edge tone (89%) of the cases. She also notes “the presence of a level L plateau extending inwards from the right edge of the verb [which was the focus in the experiment, DE]” in cases where the accented syllable is followed by two more syllables. If we translate her notation and observations to our terminology, it becomes clear that also in English the emphatic rendition of contrastive focus most frequently is a closing accent with a punctual peak (I take it that in Selkirk's description H* means association and also phonetic alignment of the peak with the stressed syllable) and an early aligned closing L tone. Similarly, Toepel & Alter (2004), based on perception experiments, suggest that in German information focus categorically differs from contrastive focus, the former being coded with a bitonal accent L+H* with a late peak, while the latter is more marked exhibiting a closing low tone L+H*L. Frota (2000) suggested that in European Portuguese focal falls have a pointed hat accent (in

21 In the following discussion I will simply equate the use of ‘contrast’ in the literature I refer to with my notion of focus.
22 Note that Selkirk's use of upper case and lower case letters is exactly the opposite of the way it is used here.
our terminology) in narrow FOCUS and a downstepped accent (also in our terminology) in wide FOCUS cases. For Dutch, Hanssen et al. (2008) investigated the differences between broad information focus, narrow information focus (i.e. ‘identificational’ focus in É. Kiss’s terminology) and (narrow) ‘corrective’ focus as a case of contrast. They found that the closing L2 is lowered in all types of narrow focus (i.e. information and contrast), but contrary to what is commonly assumed, the peak was not higher in narrow focus cases, but even lower. The falling part of the accent (basically a rise-fall) in narrow focus cases was steeper than in broad focus while the rise was not affected. Finally, the peak and the second elbow (L2) were aligned earlier. The authors conclude from these results that there is no categorical difference between focus conditions, but only a phonetic one in the sense of hyperarticulation (Lindblom 1990) of the falling part of the accent, which they also take as evidence for the higher communicative value of the second part of the accent. In general, the phonetic facts found in all studies are strikingly similar to the facts found for EA (see Figure 32), cf. specifically the results of experiment II, reported in Section 4.4.1.1. It is only the analyses that differ between the approaches.

What all of the phonetic features reported in the literature – whether they are interpreted phonologically or phonetically – have in common, is that they enhance the prominence of an accent. The notion of focus accent in the present model is ultimately based on the assumption that prominence differences may be auditorily detected. A focus may be discerned quite easily when all material surrounding the accent is downtoned. The problematic cases are those foci that are not surrounded by low prominences. I have suggested above that a focus accent is an accent that departs from default, meaning that its tones are aligned and/or scaled differently. Scaling has always been a much debated issue and the issue whether there is a categorical distinction between an 'emphatic' and a 'normal' accent is not yet settled in the current debate. Ladd & Morton (1997) tested listeners' perception and interpretation of single-accent utterances covering a variety of pitch ranges and they found that they actually interpreted some of the accents as 'normal' and others as 'emphatic'. In an earlier paper, Ladd (1994) had made the attempt to make such a distinction in terms of pitch levels. This attempt was not accepted due to the gradient nature of pitch range manipulation. This is why most prosodists distinguish between gradient and therefore phonetic (pitch range manipulation) and categorical and therefore phonological features (accent type). In fact, the alignment of tones on which the distinction between accent types is based in intonational phonology is an equally gradient matter, as I hope to have shown so far. That means that a distinction between what is supposed to be 'normal' or default and what is 'emphatic' or focal - if it can be made at all - will have to rely on continuously varying cues. However, the human ear is able to identify different intervals in music, for instance, so why not in speech. Furthermore, language offers additional cues that are of equal importance in interpreting prosody and these cues being the linguistic contents as well as the linguistic and non-linguistic
context. Both types of factors, the various acoustic and the linguistic and discoursal cues, are also implied in the following quote from Ladd & Morton (1997):

*A variety of acoustic and pragmatic parameters play a role in this decision* [whether an accent is normal or emphatic, DE], *including pitch range, voice quality, lexical content, discourse background, relationship between the speaker and listener, and so on*. Any or all of these parameters may be continuously variable, and the continuous variability may be directly perceptible as such, and there is thus no true categorical perception. Yet the interpretation computed on the basis of all the input parameters nevertheless normally falls unambiguously into one category or the other. (Ladd & Morton 1997: 339, *emphasis added*)

If Ladd & Morton's experimental results are to be interpreted in favour of a distinction between two categories, the decision I have taken here is on the right track. Nevertheless, the decision, what counts as a *focus*, is not an easy task as there always remains a fuzzy zone between the ends of the continuum that give the clear cases. According to the tentative definition of "departing from default" we may identify two major types of focus accents. These focus accents may be further subdivided according to the position they occur in within the IP:

(46)

a. *closing* accents  
   i. within an IP  
   ii. IP final  

b. *leading* accents  
   i. within an IP  
   ii. IP final  

Non-final foci are mostly closing. If they are leading, their pitch range will be noticeably higher than that of the following accents or they will involve double accentuation. Final leading accents may be indicating a focus or simply a continuation rise. As a result, leading accents pose a severe problem for the phonological identification of a category of focus accent, and I will not have much to say about them for the time being. It may suffice here to remind the reader of the discussion in Section 4.5.1, where I have shown some rising focus accents with high boundary tones. In the remainder of this section, I will concentrate on closing accents. Their deviation from default may be in two different ways which may all be employed together (I do not claim this to be an exhaustive characterization of the closing accent, however): (i) higher scaling of the H (‘H) and lower scaling of L2 (‘L) (and occasional lower scaling of L1) and (ii) medial alignment of the H (H*) and earlier alignment of L2 (L‘) (and occasional later alignment of L1 (L‘)). This description is schematically represented in Figure 32.
A special case are phrase-final accents. We have already mentioned the rising accents that involve a high boundary tone. Due to final lengthening, any type of final accent sounds prominent to some extent. In fact, boundary phenomena and final lengthening were the phonetic basis for the assumption that the final accent has a special status as the nucleus of the phrase. There are especially three prosodic features that make a final falling accent especially prominent, two of them being related to pitch. A final accent in a declarative utterance that reaches a pitch level somewhere in the vicinity of the bottom of a speaker’s range sounds especially prominent. As final downstepped accent is the default in declaratives, an accent that is not downstepped, but has a pointed hat instead sounds more prominent than its downstepped counterpart. Downstepped accents in turn may be made more prominent by early alignment of L2. A very important feature, however, is intensity. At the end of an IP, intensity is supposed to be rather low and fade out. If it does not and is perhaps even higher than the intensity of the preceding accent, the perceived prominence is significantly stronger. As noted before, intensity could not be systematically investigated in the corpus because of the special problems involved in intensity measurements, but visual inspection of the intensity tracks in many of the figures serve to illustrate that point. In sum, we may say that deviation from default in the above mentioned way serves to enhance prominence, either by increasing the excursion size (cf. also Féry & Kügler 2008 for German; Féry & Ishihara 2010 for German and Japanese; Hanssen et al. 2008 for Dutch) and/or by narrowing the width of an accent (Hanssen et al. 2008); other factors such as intensity and duration (Féry & Kügler 2008; Hanssen et al. 2008) also play an important role.

Finally, let us return to the question of whether focus is any different from emphasis. I cannot give a decisive answer to this question for the time being. But there is a chance that it might be possible to identify features that are especially associated with the evocation of alternatives and others that only relate to emphasis. A good candidate for indicating emphasis is overall register. Richter & Mehlhorn (2006) found in a production experiment on Russian that the prosodic correlates of emphasis, contrary to those of corrective focus, were not bound to a specific syllable and no constant pitch accent choice could be determined, rather high F0 correlates with positive emotions and low F0 with negative
emotions. One feature that presumably only signals emphasis without indicating a *focus of interest* in EA is drawling pronunciation with extensive vowel lengthening as illustrated in (47) or (29) above (presented in Figure 33 and 26, respectively).

(47) Mido_01_M1_03

\[
\textit{mi:du} \, tˤi\textit{li}\textit{s} \, \textit{šala} \, f-jagar-a \, la?a \n\]

Mido get-up.3SG.F on DEF-tree-IN find.3.SG.M

Mido got up the tree and found ...

\[
\textit{fi:-ha} \, \textit{fatha} \, \textit{gbi::ra} \, ?awi \n\]

in-3SG.F opening big very

LH- L-

a **big** hole in it'

![Pitch track and spectrogram](image)

Fig. 33: Panel (a) shows a pitch track of an utterance including the emphatic rendition of *kibi:ra* 'big' by lengthening the stressed vowel, panel (b) shows the pitch track on the adjective on a broadband spectrogram (the sharp drop in F0 is due to glottalization within the lengthened vowel).

Another possible candidate to be signalling emphasis is the double accentuation in (28b) (Fig. 26). Functionally both cases involve emphasis on the semantic contents of the word. In the case of *kibi:ra* 'big' the size of the hole is iconically stressed, but it is perhaps not suggested that it could have been anything else than big. Likewise, we find in the corpus instances of a drawl on other properties, like *gami:š* 'beautiful' and *tˤix:i*n 'fat', in the latter case in addition to a wide movement in a lower pitch range and lower centralized vowels. An emphatic articulation of *rufajja:* 'slim, thin' would share the drawl, but be pronounced in a very high pitch register with higher and more peripheral vowels. In the case of the doubly accented *šamma:l* 'continuing to' (28b, Fig. 27), the double accent helps emphasizing the long duration immanent in the meaning of the word itself. All these are clearly paralinguistic features. On the other hand, double accentuation is also involved in example (28c) and many similar examples in which the assumption of an *alternative* makes perfect sense, especially as *rafḍ'a* 'refusing. SG.F' in the utterance *ana mif raisḍ'a / wugu:d ra:gil* 'I am not refusing having a man on my side' is in the scope of the negation and contradicts the previous utterance that says something like 'as you are refusing to take a husband'. As already noted, the issue has yet to be thoroughly investigated and the above remarks must be taken as a preliminary hypothesis suggested by the data in the corpus whose interpretation is entirely based on introspection. I think, however, that the hypothesis could gain some evidence (or be falsified) from experimental data, especially acceptability experiments using different prosodic cues and explicit contrast. In their study on educated EA Gary &
Gamal Eldin (1982: 49) suggest two types of "sentence emphasis" they term "non-contradictory" and "contradictory" emphasis. They note that in the first type "the entire sentence contour is raised, with the S [sentence, DE] rising and sentence final syllable prolonged and followed by a slightly rising pitch" (ibid.). They cite an example that in fact always comes with that characteristic emphatic tune that suggests a strong involvement of the speaker. I replicate the example here in (48), adding prosodic annotation as used in this study.

(48) Gary & Gamal Eldin (ex. 357; p. 49)

\[
\text{sˤaħb-ak} \quad da \quad \text{rụγa:j bi-faːkl}
\]

friend-your this talkative in-appearance

LH- °L^H

For “contradictory emphasis” they cite examples with a negated presupposition first and an assertive correction following it. They note that the "greater stress and higher intonation" is on the asserting constituent. This description given by Gary & Gamal Eldin notes two important features also observed in my EA data for emphatic utterances: register shift and prolonged pronunciation (cf. the pitch track in Figure 22). On the other hand the description of “contradictory emphasis” is reminiscent of a focus accent.

To summarize the discussion on focus, contrast and emphasis, the EA data clearly shows that differences in prominence can be employed to mark the speaker's special interest in certain semantic items. Hellmuth's (2010b) claim that only contrastive focus is marked by an expanded pitch range, while information focus is not marked at all, cannot be verified by the evidence from the present investigation. Her results, however, are supported if her ‘new information status’ is not conflated with the notion of information focus, as Hellmuth actually does.

The results of the present study also suggest that narrow FOCUS, whether contrastive or completive, is marked by a focus accent. And the difference between a 'contrastive' focus accent and a focus accent in a narrow information FOCUS case is most probably a gradient matter. From a phonetic perspective, the results of this study are in line with the findings of Norlin (1989) and Hellmuth (2010b) for EA, and Chahal (2001) for Lebanese Arabic, as there is optional pitch range expansion of a FOCUSED constituent and pre- and post-focal pitch range compression. It could, however, be shown that it is relative prominence in the first place that indicates focus accents, higher scaling only being an optional correlate of this prominence. It was further shown that the closing accent (Bolinger's A accent) is a frequent option for a focus accent due to its inherent prominence. Its assertive character makes it especially suited for the indication of FOCUS. I follow Bolinger (1958: 145) in the assumption that "[a]ccent A is assertive. It is used with items that are separately important, contrastive, and/or new to the discourse.” The main claim I am making here is that a focus accent is not necessarily associated
with the FOCUS of a proposition, but is rather a matter of expressing a speaker's interest in the semantic content coded by the expression associated with the focus accent. Foci in TOPIC constituents will be dealt with in Chapter 5.

Finally, it was suggested that there may be a difference between simple emphasis and focus as evoking alternatives, but this distinction is rather tentative and is in need for more evidence from experimental studies. Generally, the basic distinction between a focus accent and an ordinary accent is not without problems. It may be the case that prominence relations between accents are in fact fully gradient. More research will have to shed light on these issues.

4.5.4.1 Parallel focus

In the preceding section we have discussed the issue of contrast as a paradigmatic notion. There is another type of contrast mentioned in the literature, called "parallel focus" by Dik (1981; Dik & Hengeveld 1997). This type of contrast is syntagmatic. In EA, it is quite regularly realized by a special intonation involving two opposite contours, in the case of a declarative this tune is a leading, followed by a closing contour.

(49) HRuh_02_F0_02
\[
\begin{array}{l}
\text{wa-la-ʕadd-ʕa} & \text{na:fi} & \text{daha:n} / \text{wa-la} & \text{ʔasman} \\
\text{and-NEG-come.back.3SG-NEG} & \text{useful} & \text{paint} \quad \text{and-NEG} & \text{mortar} \\
\text{L-LH} & 4^\text{HL} & \text{L}^{\text{H}} & \text{L}^{4\text{HL}} \\
\end{array}
\]

'And neither paint, nor mortar could help anymore'

(50) HRuh_09_F0_01
\[
\begin{array}{l}
\text{haddo} & j-\text{bi:-ʕ} & \text{il-ʕa:li} & / & \text{bi-ɾxizʕ} \\
\text{someone} & \text{3SG.M-sell} & \text{DEF-expensive} & \text{with-cheap} \\
\text{LH} & \text{L}^8 & \text{L}^{4\text{HL}} & \text{L}^\text{H} \\
\end{array}
\]

'Would anyone trade the valuable for the worthless?' (lit.: [is there] anyone who sells the expensive for the cheap)

(49) is a typical example for a parallel foci in a declarative tune (Fig. 34). In (50), the order of the contours is reversed due to the question intonation used at the end of the whole utterance. In doing so, the contrast between the two items can be maintained. Another instance of overwriting the parallel contrast intonation by question intonation is shown in (51) - but this time the leading-closing contour is employed, only that the final closing accent is not fully realized to indicate non-finality.

Fig. 34: Pitch track of a parallel contrast with the leading contour on the first part and the closing contour on the second part of the contrasting pair.
Parallel contrast seems to be confined to FOCUS phrases, most frequently parallel verb phrases - exhibiting the typical parallel structures of Arabic discourse (Johnstone 1991; Dickins et al. 2002). The pattern was especially frequent in the read speech data, but can also be found in spoken discourse. Frequently the parallel phrases are introduced by adversative connectives, such as 'either-or' and 'neither-nor' as in the final example (51), displayed in Figure 35. The examples examined here all contained a clear intonation break between the two parts. But whether this is an obligatory feature of the pattern would have to be verified with a larger amount of data.

Contrasting topics, however, are not encoded in this way, as shown in the next example of a contrastive topics and identical FOCUS phrases (53). Here, parallelism is coded using the same contour type with a level tone accompanied by an additional drawl in the stressed vowel of the parallel verbs, bestowing a monotonous tone on the whole utterance. The topics, especially the second one, are coded by a closing accent indicating topic shift (cf. Chapter 5). This example is of special interest as it clearly shows the difference between EA and English or German, in which languages the topic contrast would most probably be marked by a focus accent and accompanying deaccenting of the verb, either of both instances or only of the second one. This is thus another example for the difference between the languages as shown in Section 4.3.1. Recall that it was suggested that the main difference is that EA - contrary to English and other languages of that type - does not downtone given information that is in FOCUS (see 4.5.2).
4.5.5 Integration and Separation

We may now turn to the important issue of integration called "horizontal focus projection" by Büring (2006). Contrary to "vertical" focus projection I believe integration to be a very important concept. It was first suggested by Fuchs (1976, 1980) and taken up by Jacobs (1993) who applied it to syntactic constructions. In these accounts dealing with English and German, the formal correlate of integration is deaccentuation of constituents and their integration into the scope of one accent. As EA does not employ deaccentuation very frequently, I have modified the idea of integration in chapter 2 based on Bolinger’s “dimension of ‘connectedness-separateness’ and all semantic inferences that can be drawn from this” (Bolinger 1986: 166). Bolinger (ibid.), following Pike (1945), continues this idea noting that “[w]hile A [i.e. a closing accent] singles things out, B [i.e. a leading accent] ties them in.” Thus the concept of integration used in the present study is to be seen as one of “connectedness” that serves to mark out chunks of information that, casually speaking, in some sense belong together. In chapter 2 the different types of integration assumed here have been laid out with a fair number of examples for illustration. In this section, I will concentrate on the informational aspect of integration. The concept of integration plays a major role in the expression of thetic utterances which will be discussed in chapter 6.

In Chapter 2, Section 2.2.4.4, I have referred to the notion of integration in conjunction with accent linking. In this section, I will try to categorize integration distinguishing between two major types:

(54) a. Integration I: integration of equal constituents
   b. Integration II: integration with subordination

In Section 2.2.5.2, we have already come across two patterns of type (54a), namely the hat pattern and the roof pattern. A number of syntactic uses have been given as examples for both patterns, and it has also been suggested that the pattern is used for topic-comment sentences. Based on the information structural framework that has been laid out in the preceding sections, we may now more explicitly suggest that integration in terms of a leading + closing contour (whether hat pattern or roof pattern) is typically used for the bipartite construction of TOPIC and FOCUS, following Bolinger’s observation that “the B+A [i.e. leading+closing tone or tune, DE] suggests that the first has already been introduced” (Bolinger 1986: 145) or may be presupposed via accommodation, as I would like to add. Examples have been given in Section 4.4.1 above. In chapter 5 we will look in more detail into the
question of what counts as a TOPIC. ‘B+A’ is an instance of integration of equal constituents, TOPIC and FOCUS, which are integrated in the sense that they belong together semantically, that there exists a pragmatic relation between them. Contrary to that is the notion of separation. Bolinger (ibid.) notes that two closing accents (A profiles in his terms) “may refer to two separate items of information”. This is why the closing accent is thus well suited for the marking of focus as a centre of information. Recall that in (44) we have seen a FOCUS domain containing three individual foci, which were claimed to give every constituent its due weight. Integration I (54a) may be understood as connecting two equal constituents as a TOPIC phrase may also contain individual foci. In this case there is separation on the lower accent level and integration on the higher phrase level.

Integration II (54b) is what is commonly referred to as ‘integration’ in the literature, where it refers to the formation of an accent domain with one accent having scope over more than one constituent (Fuchs 1976, 1980; Jacobs 1993). In the present model, integration is defined on a tonal basis and subordination is related to prominence. In Sections 4.3.1 and 4.5.2 many examples were presented. Phonologically, the pattern involves a (focus) accent on the element to be highlighted and downtoning of the element(s) that for some reason is/are to be deferred to the background. As we have seen above, downtoned elements usually encode given or at least accessible information that is taken for granted in the utterance at hand. We will however see two cases of subordination of information under the scope of a single accent that does not necessarily fulfil that precondition. One will be addressed in Chapter 6 where we will be dealing with thetic utterances and the other one is described here.

We have seen the different explanations given for the fact that in English, arguments frequently carry an accent while adjacent verbs are deaccented. Ever since Schmerling’s (1976) seminal study this fact has mostly been interpreted in terms of structure (cf. Gussenhoven 1983; Ladd 1996). In his answer to what he calls the “Ladd-Gussenhoven” rules, Bolinger (1989: 235) countered that the accent on nouns is based on their relevance and informativeness and the deaccenting of verbs by their relative redundancy in most cases, arguing that

[n]ouns far outnumber verbs. They are more explicit, and can more readily stand for the actions that are performed in connection with them than the actions can stand for the sense of the noun. […] The verb, for all its importance, is relatively redundant. […] But this is not a ‘rule’. If the Predicate is more interesting than the Argument, then it gets the accent.”

Conceding that "on the basis of the English cases alone, it is very clear that relative semantic weight is a key factor in the location of accent even in many broad focus cases” (Ladd 1996: 186), Ladd (pp. 190f.) notes that a comparison across languages reveals that there are other languages which treat arguments and predicates alike with respect to accentuation. He argues that in the following semantically and syntactically equivalent sentences in English and Italian the main accent occurs in different positions (p. 191).
a. I have a BOOK to read.
b. Ho un libro da LEGGERE.

Another case for structure that has been made as early as the beginnings of structure-based accent placement (Chomsky & Halle 1968; Liberman & Prince 1977) is the accentual difference between phrase accent rule (NSR) that puts the main accent at the end of a phrase and compound accent rule that puts the accent at the opposite (initial) end of the domain.

We have seen that in EA, every content word carries an accent and that verbs frequently also carry a closing accent being the ‘exponent’ of predication (4.4.1). In any case, EA would have to be considered a language that does not make a difference between word classes. But if we take a closer look at the corpus data, we find that verb accents in fact show enormous differences in prominence. Take, for instance, the verb series consisting of a coverb (Maas 1995, 2000; Ahmed & Selmy 2000) and a full content verb. In such syntagms, the coverb is frequently subordinated to the full verb, i.e. less prominent. This subordination in prominence may even result in tonal integration into the accent domain of the full verb (56, Fig. 36).

(56) Rap_F0_22
raḥ-it MINAZZIL-A-L-HA faʕr-aha
go-3SG.F let.down.PTCP-SG.F-to-3SG.F hair-3SG.F
'And then she let down her hair.'

Fig. 36: Pitch track of an utterance with a complex predicate with subordinated coverb and main accent on the content verb, an active participle.

The same is true if a verb is more or less predictable from a noun it collocates with as in the cases Bolinger adduces for his argumentation of informativeness as the basis for accent placement and accent strength. The different prominence values associated with the verb-noun syntagms in (57) below are convincing evidence for Bolinger’s view.

(57) ʔaːl / jimkin idʕ-i dʕidʕa btaʃ-ti
say.3SG.M maybe DEF-frog POSS-1SG
LHL L!H L^HL-
'Maybe my frog ...
Fig. 3.7: Pitch and intensity tracks of an utterance containing a verb plus a prepositional complement.

In (57), illustrated in Figure 3.7, Mido and his dog are looking for the frog in the woods and on their way they see a hole in the ground and suspect that the frog could be in it. The verb *daxal-it* 'she entered' is clearly lower in prominence than the following PP containing the noun *ħufra* 'hole' which is associated with high pitch values and high intensity, whereas the verb is low in prominence, being largely predictable from the noun phrase, i.e. if the frog really is inside the hole it must have gone into it in the first place. What is of interest here is where the frog might be and not the activity of going into a hole. We will see in Chapter 5 that such instances may be analysed as involving a TOPIC that includes the verb.

This takes us to the important issue of the factors governing integration by subordination. We have already seen that it is not always the case that what is informationally new is accented and what is given or accessible is downtoned. Another example is (56) above. Here the direct object is subordinated to the focus accent on the content part of the verb, but the downtoning of the object expression cannot be accounted for by its being 'given' and the strong accent on the verb neither by its placement by default, nor by its being 'new'. Instead, the whole utterance is a case of "iteration" as suggested by Fuchs (1984), which means that an utterance is repeated with a characteristically deviant accent pattern from a "first-instance" utterance. (56) is taken from the story of the famous fairy tale *Rapunzel* where the enchantress orders Rapunzel "to let down her hair" so that she can climb up to Rapunzel's chamber. I give the full text of this scene in (57) as told by the gifted narrator *Heidi Abdel Shahid*.

(58) Rap_F0_21-22

\[
\begin{align*}
ga-t & \quad \text{is-sahir-a} & \quad w-ʔaːl-it \\
\text{come-3SG.F} & \quad \text{DEF-enchantress} & \quad \text{and-say-3SG.F}
\end{align*}
\]

Then the enchantress came and said:
The last line of the text (quoted in 56) is completely given by prior mention in the preceding IP. As suggested by Fuchs (1984: 148), in such an utterance the deviant accent structure is used to indicate focus on the "factuality" of the predication, not on the propositional content itself. In the EA example, the early focus accent can be interpreted along these lines. A possible paraphrase would be "So she acted as ordered" or "So she did the letting-down-the-hair procedure". As the last paraphrase indicates, integration in the example at hand is indeed denoting the conceptual unity of the proposition.

While in this example the whole content of the integrated phrase was given, it is new in the following EA equivalents of compound phrases, to which we finally turn now. As already noted, EA has a means of complex noun formation called annexation, or id'a:fa in Arabic. An annexation exactly mirrors an English compound in terms of word order, having the order head-modifier as opposed to the English modifier-head. As an annexation involves two content nouns we predict a construction with two accents. These two are often integrated via Integration I as has been shown in 2.2.5.2 (example 18).

Such examples were called "intermediate compounds" by Bolinger (1989: 245). On the other hand, we may also find many instances of Integration II. If the assumption made in the literature (esp. by Hellmuth 2006b; inter alia) that EA (presumably like Italian) has the main accent always on the final element is true, we should only encounter cases of integration of the type (. x), but in fact the other type (x .) is probably even more common, if it occurs at the end of an IP. We have already discussed such cases in 4.3.1.1 above. Two examples from the contrastive focus experiment exhibiting the two types of Integration II are given in (59) and presented in Figure 38. Both were given as an answer to the same question and we may safely assume that they have the same information structure.

(59)

a. MADJCV3M01
Q: Where did you see Malaka?
A: šuf-na malaka [fi manzil lamlu:m]FOC
see-1PL Malaka in house Lamloum
'We saw Malaka [in Lamloum's house]FOC

b. MADJCV2F01
Q: Where did you see Malak?
A: šuf-na malak [fi gni:nit šadna:n]FOC
see-1PL Malak in garden Adnan
'We saw Malak [in Adnan's garden]FOC
If it is neither information status, nor information value that is responsible for the variation in accentuation here, what else is at stake? I suggest that the different main accent placements are simply a consequence of the choice of the intonational tune made. In Rastegar-El Zarka (1997) I observed that the main accent in Egyptian MSA utterances may either occur at the beginning or at the end of a domain. I also suggested that an additional high prominence may be placed in the initial position of a prosodic phrase or an intonation phrase in addition to the phrase final prominence. This additional accent could even be the more prominent one. Another observation made in the MSA data was the occurrence of an early peak signifying greater finality and 'matter-of-fact' (cf. 2.1.3.1, esp. Fig. 9). The two phenomena of early peak and initial main prominence thus seem to be tied together in the choice of the intonational tune of 'declaration'. If, now, the final two elements happen to be part of an annexation, the resulting accentuation pattern will be either (x .) in the case of the declaration tune or (. x) in the case of some other tune, e.g. the redundancy contour mentioned above, or simply a type of accentuation that Bolinger (1986, 1989) called accent of power or more specifically climax by which "an extra degree of force is achieved by moving the terminal accent rightward." (Bolinger 1986: 75).

The semantics of the declaration tune are also compatible with a focus on 'factuality' as in (56) above.

It may be noted that the effect of integration with initial prominence is that the initial accent is more easily interpreted as having scope over the whole integrated construction than a late prominence. The reason for this effect may lie in the fact that an intonational tail is necessarily interpreted as belonging to the main accent preceding it, while a linking accent preceding a major one may logically either be anacrustic to this accent or be a 'tail' of the one before it. We will return to that issue in Chapter 6.

What are the repercussions of these facts on the structure-based vs. radical focus-to-accent debate? The facts of EA accentuation described in this chapter suggest that the position of a main accent or accent strength in general is influenced by three main factors: (i) the information value and the speaker's wish to put a certain denotation into the foreground for whatever reason there may be, (ii) solely prosodic consideration such as rhythm, end weight or intonational tune and iii) by the need to use a pattern deviant from what may be called the 'neutral' or 'default' one to indicate special nuances of information. We may refer to these three types of factors as (i) the semantico-pragmatic factor, (ii)
the phonological factor and (iii) the structural factor. These factors may either be in conflict or in harmony. For instance, they are in harmony when a highly informative item is placed at the end of an IP, and they are in conflict when this is not the case. Actually occurring utterances are always the outcome of an interaction of the different cognitive and linguistic levels involved, among them information structure, semantic content and prosody, the outcome of which is not fully predictable.

Thus, the whole issue of accent placement and intonational shape in general is a multifactorial procedure, just as accenting and melodic shape are multifunctional by virtue of the small number of possibilities offered - in comparison to the wealth of combinatory formations that are possible on the basis of an inventory of arbitrary segmental phonemes, for example. This fact alone makes the grammaticization of prosodic features so difficult to achieve.

To summarize, integration, as understood here, is the opposite of separation. Its function it is to tie concepts together in one informational chunk or separate them and give each of them its own weight, respectively. Integration and separation are realized tonally by tonal linking or the lack of it. As a result, separation, enacted by a series of closing accents, is best used to indicate focus on a linguistic constituent. If the material adjacent to the focus accent is downtoned or subordinated to the focus accent, this latter may have scope over the whole information chunk. If it is not - as in the case of ALL-FOCUS utterances with multiple foci (44) - the accent will only have a narrow scope over the word it is associated with. Integration, on the other hand, indicates the conceptual unity of the semantic concepts involved, whereby an early accent placement is more apt to be interpreted as the accent having wide scope over the right-attached material. In sum, this discussion of integration II has provided further evidence for the fact that a focus accent on the one hand and downtoning on the other hand may be employed for various reasons, among them contrast, emphasis of the semantic content denoted by an expression as well as other aspects of the semantic content of propositions, such as modality, factuality and the like.

4.5.6  A note on broad and narrow FOCUS

The notions broad and narrow FOCUS relate to the scope of FOCUS, not to its type. At the same time these terms refer to the phonological equivalents of what in semantic terms are sentence FOCUS and predicate FOCUS on the one side and constituent FOCUS on the other side. The classification proposed here is not entirely equivalent to Lambrecht's (1994) classification. FOCUS may be of different kinds or types, some of which have been alluded to in the above sections, such as identificational FOCUS as an answer to a constituent question (who, when, where) or types involving contrast, such as corrective focus or parallel contrast, as described above. An in-depth investigation of FOCUS types requires the thorough investigation of the interactions between syntax and prosody as
well as the contribution of other grammatical and lexical devices which is beyond the scope of the present study. I will only make two observations here for the time being because their consideration is most important for the identification of narrow FOCUS. One of these pertains to the question whether EA has a special syntactic position that hosts narrow FOCUS (contrastive or identificational) as has been claimed for other languages, e.g. preverbal for Hungarian (É. Kiss 1981, 1998; see Gyuris 2012 for an overview) or Finnish (Vallduví & Vilkuna 1998). Clefting has been identified as one of the major devices for the expression of argument FOCUS many languages (cf. Winkler 2012 and Klein 2012 for an overview on English and French, respectively).

Moutaouakil (1989: 19ff.) argues for a distinction between “New Focus” and “Contrastive Focus” in what he calls Modern Standard Arabic (MSA), within a Functional Grammar framework following Dik’s taxonomy (Dik 1881). He claims that in MSA, the preverbal sentence position is the one for contrastive focus (p. 24). The examples he adduces are given in (60) (transcription and glossing changed).

(60)

a. 
\(`\text{fA}:\text{J-an } \text{farib-a } \text{xa:lid-un}\) 
\text{tea-ACC } \text{drink-3SG.M Khaled-NOM} 
‘It was tea that Khalid drank’

b. 
\(`\text{xa:lid-un } \text{farib-a } \text{fA}:\text{J-an}\) 
\text{tea-ACC } \text{drink-3SG.M Khaled-NOM} 
‘Khalid drank tea’

There is one severe problem with Moutaouakil’s study. Nowhere does he make explicit what served as the corpus for his analysis. Although he assumes his analysis to hold for MSA, the sentences look like as if they have been quoted out of a traditional grammar of Classical Arabic (CA). Moutaouakil’s justification of that fact that MSA and CA display the same fundamental structure and therefore the constructions cited are supposed to hold for both “états de langue” (p. 1) may be doubted. Anyhow, whatever the validity of the claims made in this book for CA and perhaps MSA may be, they cannot be transferred to the modern Neo-Arabic varieties that are spoken in the individual Arab countries.

Contrary to Moutaouakil’s study, Brustad (2000) offers an investigation of the syntactic features of Spoken Arabic, covering the Standard varieties of Egypt, Syria, Morocco and Kuwait. Based on much more reliable data, Brustad (2000: 343) also distinguishes “focus of contrast” from information focus, following Chafe’s definition of the terms. She claims that “new information” occupies “sentence-final position” whereas “focus of contrast does not have one definable sentence position”. She argues that VSO is the unmarked word order in Arabic and thus the unmarked position of the object is post-verbal. Contrastive objects, however, may occur preverbally in OVS sentences and contrastive subjects may occur clause finally in VOS sentences. Brustad’s cursory remarks may serve as a general
guideline, but are definitely too crude to be conceived of a as a thorough description of the interaction of information structure and syntax. Furthermore, there are clearly differences among the individual Arabic varieties in that respect. One such instance is OV order for contrastively FOCUSED objects. Brustad (2000: 350) cites an example from Moroccan Arabic (transcription and glossing changed)

(61)

<table>
<thead>
<tr>
<th>Arabic</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>qal-3SG.M.aha</td>
<td>say.3SG.M.</td>
</tr>
<tr>
<td>hadik</td>
<td>DEM.SG.F</td>
</tr>
<tr>
<td>byi-t</td>
<td>want.1SG</td>
</tr>
<tr>
<td>ana,</td>
<td>give.IMP.SG.3SG.F</td>
</tr>
<tr>
<td>?ara-aha</td>
<td>sothat.1SG</td>
</tr>
<tr>
<td>baf</td>
<td>slaughter.1SG-</td>
</tr>
<tr>
<td>n-dbah</td>
<td></td>
</tr>
</tbody>
</table>

‘He said to her, that one I want, give it here so I can slaughter’

That the preposing of direct objects in FOCUS function is a possible construction in Moroccan Arabic, is also attested by Ouhalla (1997). In EA, however, a literal translation of this example would be ungrammatical. However, in line with Brustad’s claim, the EA data gives evidence that contrastive subject pronouns can appear in the right periphery as the following example from the corpus suggests:

(62) Mar_HF_MM5_02

<table>
<thead>
<tr>
<th>Arabic</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>ma-smi-fi-ha:-S</td>
<td>NEG-hear.2SG.F-3SG.F-</td>
</tr>
<tr>
<td>inti</td>
<td>2SG.F DEM.SG.F</td>
</tr>
<tr>
<td>di</td>
<td></td>
</tr>
<tr>
<td>/</td>
<td></td>
</tr>
</tbody>
</table>
| ‘You haven’t heard it before, this one…’

<table>
<thead>
<tr>
<th>Arabic</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>?ul-ta-aha:-i-ik</td>
<td>say.1SG-3SG.F-to-2SG.F</td>
</tr>
<tr>
<td>ana</td>
<td>1SG</td>
</tr>
</tbody>
</table>

‘Now you have heard it from me’ / ‘Now I have told you’

Example (62) is an instance of contrastive reference to the two interlocutors. However, in this example the referents of the pronouns are topics. The example shows that finality is not immediately associated with ‘new information.’ On the other hand, the same position could also be filled by an independent pronoun denoting contrastive FOCUS. Thus, the EA data also corroborates Brustad’s claim quoted above that there is no special position for contrastive FOCUS in EA.

We may therefore state that there is no special FOCUS position in EA. Although the main information tends to be positioned at the end of an IP, the initial position is also well suited for that purpose. Furthermore, the corpus data also confirms the hypothesis put forward by descriptive work on EA (Heliel 1976; Mitchell 1993) that FOCUS may stay in situ and does not have to be 'moved' to a special position. As suggested above, narrow FOCUS usually contains a focus and that is the strongest prominence of the whole utterance.

Brustad does not mention clefting as a means of information structure, and Moutaouakil only mentions the pseudocleft construction (Moutaouakil 1989: 20) with the element that constitutes the narrow FOCUS in sentence final position (63) – which is in line with his predictions. Although pseudoclefts definitely occur in MSA and also in EA as evidenced by the famous Egyptian proverb in (64), the
equivalent of an *it*-cleft is probably more frequent. Such cleft constructions have been studied by Ouhalla (1997) in Moroccan Arabic (MA) (examples of EA and MA are given below).

(63)
\[ l-laði\quad raðaj-tu-hu:\quad zajd-un \]
REL.SG.M  see-1SG-3SG.M  Zayd-NOM

‘The one I saw was Zayd’

(64)
\[ illi\quad faːt\quad maːt \]
REL  go.by.3SG.M  die.3SG.M

‘Let bygones be bygones’ (lit. what has gone by is dead).

A cleft sentence divides a proposition into two parts, a presuppositional part and a FOCUS part. Additionally the FOCUS constituent is usually viewed as involving an exhaustiveness effect (cf. Zimmermann & Onea 2011 for a discussion of this issue). The following pair of sentences are answers to the question 'Who ate the beans?' Interestingly, the translator translated the question with a cleft sentences (65), but offered two (equivalent?) translations to the answer: a FOCUSED subject in sentence initial position (65a) and a cleft version that marks the subject as the FOCUS constituent of the sentence (65b). In the latter example the main focus accent may be associated with the FOCUS NP, or this NP may be prosodically coded as a TOPIC, in which case the pronoun *hijja* 'she' shows the typical FOCAL closing accent. The pitch track of a sentence of the type of (65a) is presented above in Figure 13b, and a pitch track of (65b) is presented in Figure 39 below. In this example, the lexical NP and the pronoun *hijja* both carry focus accents. It would, however, be equally possible to have a focus accent only on the lexical NP or the pronoun only, in which case it would be scaled higher than the preceding TOPICAL NP. The main thing, however, is that the narrow FOCUS is marked syntactically and intonationally, obviating an explanation in terms of functional complementarity between syntax and prosody as suggested by Vallduví (1992). The EA facts are also in line with Ouhalla’s description of clefted FOCUS constituents in Moroccan Arabic. (Ouhalla 1999: 354) notes that in Moroccan Arabic both pronunciations given in (66a) and (66b) (Ouhalla’s examples 14a, b - capital letters are used for primary stress) are indeed possible with "no apparent difference in interpretation". The same is equally true for EA.

(65) (translation task (SFB 632, D2 project)
\[ miːn\quad illi\quad ṭakal\quad il-fuːl \]
DEF-goat  POSS  eat.3SG.M  DEF-beans

Who ate the beans? (lit. Who is the one that ate the beans?)

a. 82ARZ-C38FIT-01F0100
\[ \text{FOC} [il-miʃza]\quad \text{bitaː} \text{Sit} [bta\text{Sit}]\quad \text{is-}\text{sitt}\text{FOC}\quad \text{ṭakal-it}\quad \text{il-fuːl} \]
DEF-goat  POSS  DEF-woman  eat-3SG.F  DEF-beans

The WOMAN'S GOAT ate the beans.

b. 82ARZ-C38FIT-01F0101
\[ \text{FOC} [il-miʃza]\quad \text{bitaː} \text{Sit} [bta\text{Sit}]\quad \text{is-}\text{sitt}\quad hijja\text{FOC}\quad \text{lli}  \text{ṭakal-it}  \text{il-fuːl} \]
DEF-goat  POSS  DEF-woman  3SG.F  REL  eat-3SG.F  DEF-beans

The WOMAN’S GOAT - she is the one who ate the beans.
Fig. 39: Pitch track of a narrow argument FOCUS of the subject phrase, syntactically coded as a cleft.  

(66)

a.  
*Nadia* HIYYA lli ʔllf-at l-ktab  
N. PRON.she RM.the wrote-she the-book  
'It's NADIA who wrote the book.'

b.  
*NADIA* hiyya lli ʔllf-at l-ktab  
N. PRON.she RM.the wrote-she the-book  
'It's NADIA who wrote the book.'

A final observation I would like to make in passing here is the use of a special lexico-syntactic construction for the marking of a FOCUS domain that is very frequently employed in EA. Especially when narrating or describing, EA speakers frequently state what they want to say in the form of a question, more often than not in combination with the particle *baʔ(a)*, a kind of sequential marker, and always followed by a prosodic break, only to answer it with the FOCUS phrase in the next utterance. Such questions may elicit whole predicate phrases (67a) or single constituents (67b) that are the FOCUS of the utterance.

(67)

a. Recip_F0_03  
*ba*saʔ b-a-ʕmil ʔe: baʔa /
but IND-1SG-make what PRT  
'But what do I do (then)?'

ana ʕam-nafs-i  
SG from-soul/self-1SG  
*foCL*{b-a-ḥibb a-ḥutˤtˤ / il-lahma...}FOC  
1SG IND-1SG-love 1SG-put DEF-meat  
'I personally LIKE TO PUT THE MEAT...'

b. Recip_F5_01  
*t-rucʰi* ʔhatˤtˤ-a baʔ ʕale:-ha ʔe: /
2-go-SG.F put.PTCP-SG.F PRT on-3SG.F what  
'Then you put on it (the dough) (what)??...'

*FOCL*fwajjɪt / malḥ / wi-fwajjɪt...}FOC  
a.little salt and-a.little  
SOME SALT and SOME...

---

23 I have also noticed this strategy to be pervasive in Iraqi Arabic, at least with a speaker from Kirkuk (aged 55).
Furthermore, it may be observed that the question word frequently occurs right in front of an NP. As we have seen above, some languages seem to treat arguments and verbs treated differently with respect to accentuation, i.e. that arguments receive the sentence accent while adjacent verbs receive lower stress or are deaccented (Chafe 1974; Schmerling 1976; Gussenhoven 1983, inter alia). We have also seen that in EA this is not the case, the type and strength of the accent on a verb is being dependent on its semantic value. In (67b) the verb is semantically 'light' and the most important information clearly is in the direct object - the constituent asked for. Lambrecht (1994: 267) suggests that “[t]he prosodic difference between nouns and verbs can be seen as a consequence of the inherent difference in the way in which discourse-referential vs. non-referential expressions are processed”, Bolinger (1989: 235) suggests that nouns inherently carry more information than verbs. Whatever explanation is the right one, it is perfectly conceivable that speakers of EA use another strategy to single out the informationally rich arguments, using a break before the argument and optionally emphasising it by the use of a question word.

The upshot of the above discussion is the fact that it is not syntax alone or prosody alone that mark information structure in EA, but rather the conspiracy of both levels that produces a construction which as a whole may be interpreted as signalling a special type of information structure.

4.6 Summary and conclusions

In this chapter, I have given a functional account of pitch contours in Egyptian Arabic. Among other functions, intonation is used to encode information structure. To understand how this works in EA, I have suggested drawing a distinction between pragmatic relations and pragmatic properties, which are expressed by the shape of tonal contours and the relative prominence of individual accents, respectively. Furthermore, I have proposed that the shape of the tonal configurations (leading and closing) and their relative prominences can be attributed to the effects of two biological codes, the frequency code and the effort code. While contour shape is related to the frequency code, prominence relations with their tonal correlates of pitch excursion and compression as well as tonal alignment, and their other prosodic correlates, such as intensity, duration and double accentuation can be explained by the effect of expending more or less effort on the articulation. While prominence differences serve to single out informative elements and put them to the foreground, contour types serve to encode their relation.

In this final section I will give a summarizing review of the major tenets proposed in the present model of the prosody - information structure interface in EA, thereby touching upon some of the issues that have been a matter of debate in the field for a long time and suggesting my views on them in a
comprehensive way. To do so, it seems appropriate to repeat here the main questions that have been discussed in this chapter:

i. Why assume a two-dimensional system of pragmatic properties and pragmatic relations?

ii. What is the relation between pragmatic properties and pragmatic relations?

iii. What is the role of information status in expressing focus and background?

iv. What is the relation between focus and accent?

v. May narrow FOCUS be expressed 'in situ' or does the expression of narrow FOCUS necessarily involve a certain (marked) word order of the constituents involved?

vi. What is the role of 'contrast' and how is it related to the notion of focus?

vii. Is focus structure-based or simply a procedure of 'semantic highlighting'?

In Section 4.3, I have argued that downtoning in EA is the equivalent of deaccenting in West-Germanic languages. Although EA does not normally resort to complete deaccentuation of content words, it has been shown that the differences in prominence between the various constituents of an utterance parallel what is commonly treated as accentuation vs. deaccentuation in English and familiar languages. This suggests that Bolinger’s notion of "gradient accentuation" (Bolinger 1989: 243, et passim) is indeed the better way to think about this matter. I have also shown that only one type of pragmatic properties as defined in Chapter 1, namely the dichotomy of focus and background, has an impact on the prosodic encoding of focus in EA. While information status is neither a sufficient nor a necessary condition for downtoning, focus and downtoning are rather taken to be reflecting the speaker’s decision which parts of the utterance to put into the foreground and which parts to background. I have proposed that one reason for downtoning is that the speaker takes the information given 'for granted', which perhaps can be viewed as another way to say that the information is presupposed. This is in line with what Bolinger and Schmerling also propose for English, namely that "it is not anaphoricity that is directly correlated with lack of stress, rather, lack of 'significance' in some sense" (Schmerling 1976: 73), and that there is "no direct relation between stress and grammar, but direct relation between stress and pragmatic factors" (p. 80). Based on these observations, Schmerling then formulates her Principle I that "[c]ertain items in an utterance are treated by the speaker as relatively "insignificant" and fail to be assigned stress" (p. 75). According to these assumptions, it could be shown that focus is ultimately a matter of the speaker's choice and may vary with the FOCUS structure remaining constant. I have also suggested that focus usually evokes alternatives. At the same time I have, rather tentatively, tried to differentiate between focus as evoking alternatives and ordinary emphasis.

Consequently, the idea of a focus exponent being the indication of a FOCUS domain, as perpetuated in the different models on focus prosody has no function in the present model, which means that focus
projection is not assumed to be existent at all. The relation of focus and FOCUS is rather seen as an indirect one. FOCUS as a relational component of a pragmatically structured proposition is viewed to be in fact a type of predication that may be understood as the missing information in an open presupposition or the answer-part in a question-answer pair. Its prosodic correlate belongs to intonation proper, being a closing tone or tune as opposed to a leading (or linking) tone or tune for TOPICS. The prototypical prosodic coding of pragmatic relations may, however, be overwritten by other functions of the intonational tunes or they may interact.

We have thus arrived at the most intricate issue, the question of what constitutes a focus accent and whether - and if, then how - it could be distinguished from an 'ordinary' accent or an accent that is a matter of prosody alone, i.e an accent that lacks any specific meaning. I have suggested that accent and focus are not necessarily co-extensive, but that focus is meant to be foregrounding or highlighting by means of an accent - in particular, a certain type of accent. It has to be conceded that the distinction to be drawn here stands on shaky ground. I have proposed that a perceptually especially salient accent should be viewed as a focus accent whose acoustic correlates suggest that it is in some way deviating from default. The phonetic correlates of this type of accent all contribute to enhancing its prominence. The problem with that definition, however, is the gradience of these phonetic correlates and the fact that there is a great deal of overlap between a default accent and a focus accent in certain positions. Nevertheless, to draw a functional distinction between what Ladd & Morton (1997) have called an 'emphatic' and a 'normal' accent seems to be a promising path.

If, however, focus and accent are not necessarily co-extensive and if focus is rather signalled by accent, how can the relation between accent and focus be described, then? A frequent case is a focus accent signalling focus on the linguistic expression it is associated with. This is perhaps the most frequent type of focus. There are, however, two other - opposing - strategies to mark focus by accentuation, the first of which is double accentuation, where two accents signal a focus on a certain constituent, and the second one is integration where one focus accent has scope over a constituent larger than the lexical item it is associated with. Let me, for the purpose of illustration, repeat two of the examples given in this chapter that show the two possibilities. (68) is a case of focusing the accentuated element within a larger FOCUS domain where the alternative evoked is restricted to the concept denoted by the lexical item carrying the accent while (69) shows an example of a focus accent that evokes an alternative for the denotation of the whole integrated constituent.

(68) 26ARZ-D03FCT-03M1100
laːʃʃukulaːtˤaʃ fukulaːtˤaʃ
no TWO chocolate
'No, TWO pieces of chocolate'
(69) **Rap_F0_22**

\[ \text{FOC}_{	ext{rah-ūt} \ fafr-ahu} \]

\[
\text{go-3SG.F} \quad \text{PTCP-let.down-3SG.F} \quad \text{hair-3SG.F}
\]

‘So she let down her hair for her’.

In (68) - a repetition of (22) - the alternative evoked is another quantification of ‘chocolate’ which itself is not at stake here, but taken for granted. Contrary to that, in (69), it is an aspect of the whole predication that is ascertained by the focus accent. The alternative evoked is ‘not letting it down. Thus, it is the factuality of the predication that is focused. There is nothing in the form of the utterance that signals the scope of the focus accent, but it is the context that we have to refer to. We may thus state with Lambrecht that alternative focus readings

result from implicatures drawn on the basis of conversational contexts, not from grammatical rules of focus construal. When alternative focus readings [...] are to be made formally explicit, prosodic focus marking has to be supplemented with, or replaced by, morphosyntactic marking [...]. (Lambrecht 1994: 296)

Although Lambrecht’s statement refers predicate FOCUS, we may re-interpret his remarks as being applicable to the lower focus domain. This is in fact tantamount to saying that there is "horizontal focus projection" in the domain off focus, but no "vertical focus projection" projecting focus on the higher domains of FOCUS and TOPIC.

Lambrecht’s remarks still point to another important issue concerning the question whether focus marking (or the formal encoding of other information structural categories) is a matter of grammar or has a more general cognitive basis. It has been suggested that there is no one-to-one correspondence, neither cross-linguistically, nor within one individual language between an assumed information structural notion of focus and grammatical devices that indicate it (Féry 2007; Zimmermann & Onea 2011). It may be the case that in some languages the expression of focus (esp. contrastive/identificational focus) is a matter of grammar, for example by the use of a special syntactic position for focus (Moutaouakil 1989; É. Kiss 1998; Vallduví & Vilkuna 1998), or by morphological focus markers as in Gürüntüm, a West-Chadic language (Hartmann & Zimmermann 2009). In prosody, however, grammaticization is rather doubtful. The problem with prosodic categories is their multifunctionality. Regarding the prosody - information structure interface, and perhaps the interface between information structure and formal linguistic categories in general, it seems appropriate to differentiate between 'construction' or marking of focus for example, and 'construal' or interpretation of such constructions. The EA data give evidence for a position as expressed by Fuchs (1984: 161) that

[1]The choice of an accent pattern is not determined by context conditions, since it has an autonomous semantic function (although certain accent patterns may show a high degree of affinity to certain context types). Its interpretation, on the contrary, is systematically dependent on context givens of all sorts.
The EA data suggest that focus marking via accentuation is not structure-based, but rather based on more general cognitive processes, such as foregrounding certain elements of the speech flow, the semantico-pragmatic effect of which, however, is the evocation of *alternatives* and the structuring of utterances into *background* and *foreground*, which in turn helps identifying the relevant parts of an utterance and thus assists the hearer in processing the information received. In EA, we may identify prominence relations as the formal prosodic correlates of *focus* and *background*, while the prosodic correlates of *FOCUS* and *TOPIC* are melodic. The function of the former dichotomous pair is to highlight or to downtone parts of the speech flow, whereas the second pair helps structuring it in terms of linking an utterance to the preceding discourse and thereby establishing coherence with the context. Although accent patterns are never fully determined by the context, some accent patterns are clearly infelicitous in certain situations, as illustrated in (70) with a focus on the subject *Malaka* and rhythmic accents in the remainder of the utterance. The intonation pattern chosen gives rise to the interpretation of *Malaka* being the *FOCUS* of the utterance, which is not compatible with the preceding question.

(70) Q: Where is Malaka?
#A: MALaka f-MANZil lamLU:M
M. in-house L.
'MALAKA is in Lamloum's house'

Thus, the evidence for formal correlates of information structural categories is rather negative evidence than positive evidence. Nevertheless, as the experimental evidence in 4.3.1.1 and 4.4.1.1 suggests, there are certain tendencies that help predict the use of individual patterns to some degree and clearly guide their interpretation.

Another much debated question in the literature is the status of *contrast*. Contrast has frequently been viewed as a subcategory of *focus* (i.e. *FOCUS*) (Dik et al. 1997; É. Kiss 1998; Kratzer & Selkirk 2007; Selkirk 2008), but also as an independent notion orthogonal to information structural categories (Vallduví & Vilkuna 1998; Molnár 2006). The results of the present study suggest that contrast is not an independent category having a uniform formal correlate. It could be shown that *contrastive focus* crucially differs from *parallel contrast* (Dik et al. 1997) as regards its prosodic encoding. The expression of corrections - if having any prosodic correlate at all - have been shown to involve a focus accent with adjacent downtoning, while parallel contrast is most commonly associated with a rising-falling contour called (*integration* I) on the parallel constituents. Although there may be a marked difference between an argument *FOCUS* on the subject as neutral or contrastive information as shown in Figure (40c) and (40d) below, this difference does not hold for all speakers and not even for all utterances of this individual speaker. In general, there was no evidence, to draw a categorical distinction between a focus signalling information or identification as the answer to a wh-question and a focus signalling contrast as in the case of correction, contradiction etc. The distinction that may rather be drawn on the basis of the EA data is a distinction between narrow *FOCUS* and broad *FOCUS*. An intonational distinction between these two types has also been claimed by Grice (1995) and
D'Imperio (2002) for Italian varieties, Frota (2000) for European Portuguese and Sosa (1999) for Latin-American Spanish. In EA, narrow FOCUS usually involves one or more focus accents with optional concomitant downtoning of the backgrounded parts (in addition to optional syntactic strategies) while the broad FOCUS may or may not involve foci. Zimmermann & Onea (2011) state on the basis of a wealth of data from different languages that there is no categorical distinction between focus types and that contrastive focus may not be regarded as semantically different from the "weaker" information focus, but that there is "only a tendency to realize pragmatically more marked contrastive foci by means of more marked non-canonical structures". They suggest the weak implication that "if a language has two grammatical ways of realizing focus, we would not expect the more marked structure to realize instances of information focus and the less marked canonical structure to express contrastive focus" (Zimmermann & Onea 2011: 1665). I concur with this characterization as I concur with the basic claim that a phenomenon such as contrastive focus is largely dependent on how a speaker assesses a hearer's knowledge and readiness to accommodate unexpected information, which may lead him to use a more marked realization. Only, that in the model proposed here, this assumption is made for focus in general, as focus is viewed as 'focus of interest' thus reflecting the speaker's choice what to put in the foreground and what to background. This decision is, of course not independent from structural considerations, but not totally determined by them. Thus, the position here is in line with Bolinger's basic convictions about the role of accent - or in terms of the present model - focus accent.

The optionality of syntactic strategies for the expression of narrow FOCUS implies that narrow FOCUS in EA may be expressed in situ, which in turn implies that it may be signalled by prosodic means alone. Figure 40 presents six typical intonation contours for five FOCUS types (two involving broad FOCUS and four narrow FOCUS) from a pilot experiment to elicit FOCUS utterances. Note that the patterns in (40a) and (40b) are indeed alike, they both exhibit the pattern typical for a topic-comment sentence. We will be taking up this issue in the Chapters 5 and 6 again.
This raises the question of the relationship between focus and emphasis. The conclusion reached from the investigation of the EA data is a very tentative one. There is some evidence that we may distinguish between mere emphasis as is involved, for instance, in an emphatic, emotionally loaded articulation of certain expressions, such as adjectives like gami:l ‘beautiful’ or tˤixi:n ‘fat’, articulated with a drawl and perhaps in a specific register and focus as evoking alternatives.

The final conclusion drawn from the investigation of the prosodic realization of information structural categories in EA is that there are quite systematic reflexes of these to be found in the prosody of the language. It is mistaken to assume that the specific phonological properties of EA, being a language which employs a pitch accent on almost every content word, impede the prosodic realization of focus and the marking of relevant or newsworthy information, as claimed by Hellmuth (2005, 2006b, 2010b). The results of the study rather confirm the claims made by Heliel (1976), Gary & Gamal Eldin (1982) and especially by Mitchell (1993) who states that

> although Arabic seems to have a greater tendency to accent all words in the sentence and to treat the last accent in a given case as nuclear, this is by no means always so, and [English & Arabic] share the possibility of locating the nucleus differently among an unchanged form of words. (Mitchell 1993: 230).

The results of this study confirm that this is indeed the case, and they furthermore show how this is accomplished. Another important conclusion to be drawn from the study of prosody and information structure in EA is that the long-held assumption that given information is deaccented (Cruttenden 2006) and new information accented is not to be generalized, being a specific characteristic of West-
Germanic. The axiomatic conceptual dichotomy of 'accented' vs. 'deaccented' and the deduction that accentuation is either all-or-none led to the assumption of an obligatory nucleus and the overestimation of its relation to FOCUS structure. The present study provides evidence for a different view that builds on the assumption that focus is simply 'highlighting out of interest' and that the phonetic properties of accent are all gradient. As a corollary, there is no obligatory nucleus in EA intonational phrases as there is no focus exponent that projects focus to a higher domain. Finally, the results of the study contradict the assumption of functional complementarity between prosody and other formal means of encoding information structure, especially syntax. It rather seems to be the case that prosody in general, and intonation in particular, is an omnipresent, non-arbitrary feature of information structure that is fairly resistant to being grammaticized. Although I do not deny that grammaticization of intonational features is possible, it is more likely to have certain morpho-lexical items acquire a grammatical function than the up-and-down of intonational contours. Evidence for the non-complementarity of prosody and syntax also comes from other 'flexible syntax' languages such as Italian and Spanish (Face & D'Imperio 2005; Brunetti et al. 2010).
5  Topicality

Although the definition of 'topic' used by most researchers today is fairly uniform based on the notion of *aboutness*, the problems of identifying aboutness topics obviously have not been solved yet as inter-rater reliability tests of topic annotation in corpora show (Cook & Bildhauer 2013). In Chapter 1, I already explained in some detail the basic theoretical assumptions for this investigation and in this chapter I will provide evidence from EA for the theoretical choices made. The main issue is the distinction that I draw between TOPIC and topic. In most models that use the notion of ‘aboutness’ for a definition of topic, a topic is explicitly or implicitly required to be an entity. This view is also reflected in Krifka’s (2007: 41) definition given in (1).

(1) The topic constituent identifies the *entity* or *set of entities* under which the information expressed in the comment constituent should be stored in the C[ommon] G[round] content. (emphasis added)

Some models assume that there is only one possible topic per sentence, which is either identified by taking discourse factors into account (Reinhart 1982), leading to unfortunate vagueness, such as the assumption that a sentence is obviously more about one potential topic than about another (cf. Lambrecht 1994: 149f.). Or it is identified structurally as being the first such element in the background (Vallduví 1992).\(^1\) If we do not want to go either way, there remains some uncertainty when the definition of topic is applied to naturally occurring data, as linguists working with naturalistic corpora are painfully aware of. We will have a brief look at the annotation guidelines developed by the SFB 632 which have also been partly applied for the annotation of the EA corpus (2):

(2) An NP X is the aboutness topic of a sentence S containing X if

a. S would be a natural continuation to the announcement
   Let me tell you something about X
b. S would be a good answer to the question
   What about X?
c. S could be naturally transformed into the sentence
   Concerning X, S’
   where S’ differs from S only insofar as X has been replaced by a suitable pronoun.

   (Götze et al. 2007: 165)

Consider the following example from Cook & Bildhauer (2013: 125) who report that annotators disagreed in the identification of the sentence topic in such cases. As the authors point out, the

---

\(^1\) Vallduví uses the term ‘link’ for topic.
(con)text (given here in (3)) suggests two possible aboutness-topics for the sentence in (3b): the ORN-Bus aus Nieder-Olm ‘ORN-Bus from Nieder-Olm’ and the Spur ‘lane’, which are plausible topic candidates. The ‘bus’ is a likely topic being the subject of the sentence – subjects have been suggested to be unmarked topics. In this case the sentence would entail a topic shift. Taking ‘topic continuity’ (Givón 1983) as the more important factor would suggest the ‘lane’ to be the sentence topic, the topic of some of the preceding sentences.

(3)

a. **Diese Busspur ermöglicht die neue Buslinie, die ab 1. Juni eingerichtet wird: (. . . )**
   *Damit erhalten zum Beispiel die Bretzenheimer einen flotteren Anschluß nach Hechtsheim (. . . ) Auch in die Altstadt geht’s schneller.*
   ‘This bus lane made possible the new bus route, which will operate as of June 1st: (. . . ) The residents of Bretzenheimer will thus have a better connection to Hechtsheim (. . . ) It will be even quicker to get into the old town-centre too.’

   ‘The ORN-Bus from Nieder-Olm will also benefit from the lane.’

One typical example for the difficulty in identifying the sentence topic can be seen in the following passage from the EA corpus, coming from a frog story. The main character of the story is a little boy, Mido, who lives with his parents and his sister. Mido also has a lot of pets, a dog, a frog and a turtle who are staying with him in his room. In the present scene, Mido, who left his pets behind to go out with his parents to a restaurant, is returning home.

(4) Mido_F6_02_05

\[
\text{ba}^{ʔ} \text{da-kida } \text{huwa} / \text{ra}^{h} \text{il-} \text{ʔo} \text{d’a } \text{bta}^{ʔ} \text{st-u} / \\
\text{after-so } \text{3SG.M} \text{ go.3SG.M DEF-room POSS.SG.F-3SG.M}
\]

‘Then he [Mido] went to his room…

\[
\text{[la}^{ʔ} \text{ʔa]}_{\text{pop}?} \text{[k-kalb wi-s-sulhiba]}_{\text{pop}?} \text{mistanni-ji:n-u} \\
\text{find:3SG.M DEF-dog and-DEF-turtle wait.PTCP-PL-3SG.M}
\]

‘he found the dog and the turtle waiting for him.’

In the second clause of (5) we can assume that Mido again is the continued topic integrated in the verb laʔa ‘he found’, but we could likewise assume that laʔa only has a presentative function to introduce the new topic ik-kalb wi-s-sulhiba ‘the dog and the turtle’ about whom the state of being waiting is predicated, which notionally is perhaps the more plausible option.

Another problem connected with aboutness topics also mentioned by Cook & Bildhauer (2013) is the question whether a certain word order (e.g. VS order) should be taken as a clue to sentence focus, i.e. an all-FOCUS sentence without a topic. VS order in a language that arguably has a basic SVO word order has frequently been taken to mark sentence focus or theticity (which are considered to be identical by some authors, e.g. by Lambrecht (1994)). We will briefly deal with this issue in Section 5.1. within a short excursus on word order in EA to clarify that issue.
5.1 An excursus on word order

Evidently, word order is intimately related to information structure in most languages. However, it is not possible to deal with the interaction of these two components of grammar in the present book. To gain genuine insights into the issue requires a large-scale study of different aspects, not only pertaining to syntactic and discourse structure, but also to morpho-lexical characteristics, semantic roles and other semantic and syntactic features. This cannot be accomplished within the present study and must be referred to future work.

Classical Arabic and Modern Standard Arabic are said to have a basic VSO word order, whereas EA, like many other modern varieties of Arabic, have been claimed to have shifted to an SVO word order (Gamal-Eldin 1967; Woidich 2006; Edwards 2010). In an early description of “Egyptian Colloquial Arabic”, Gamal-Eldin (1967: 58) considers SVO to be the “favourite order”, VSO sentences being “few in number and […] mostly classicisms. As Ingham (1994) and Holes (2010) note, the assumption of a basic SVO in neo-Arabic varieties is an oversimplification, at least for the Eastern and central peninsular varieties, such as Najdi Arabic (Ingham 1994) and Gulf Arabic (Holes 2010). Similarly, Brustad (2000) and Dahlgren (1998) who both base their analyses on corpora of naturally occurring data, take discourse factors to be decisive in the choice of word order. Especially in the generative literature, attempts have been made to derive one word order from the other one that is classified as the ‘basic’ order (cf. Aoun et al. 2010).

Even though functional approaches do not subscribe to the 'movement'-view, the question of which word order is 'basic' or 'unmarked' or at least more frequent is a legitimate question. Brustad (2000) comes to the conclusion that neo-Arabic varieties have in principle retained the basic VSO order from Classical Arabic. The problem with Brustad’s account, however, is that she counts sentences without an overt subject as VS which enormously increases the frequency of VS sentences as compared to SV sentences.

The semantico-pragmatic basis of VS/SV alternation Arabic, even in MSA, was stressed by Holes (1995: 204) who notes that the statistically most frequent word order in MSA is VSCOMP which exhibits no morphological restrictions or semantic limitations on the type of verb used and which is predominantly used for “event-oriented” messages; SVCOMP, on the other hand, is common in sentences which describe background information or “entity-oriented” messages (cf. Holes 1995: 205). This fact also seems to hold for the Arabic vernaculars. In a similar vein, Ingham (1994: 34) states that Najdi Arabic (and Arabic in general) have “two frequently occurring sentence types one in which the verb precedes and one in which a noun phrase, not necessarily the verbal subject, precedes”, in his terminology the “Uninodal” and the “Binodal” type - a distinction that, according to Ingham (ibid.), corresponds closely, but not exactly to the distinction between the verbal sentence and the
nominal sentence of the Arab Grammatical Tradition. As defined by Ingham, a uninodal sentence is all-new with given information optionally expressed in an intonational tail (ibid.). A binodal sentence has two parts where the first part or node presents given information and the second part consists of new information. Ingham (1994: 36) explicitly likens the first node to Li & Thompson’s (1976) Topic. Similarly, Brustad (2000: 342) notes that in Spoken Arabic virtually any basic constituent may occur in sentence initial position. She proposes that Spoken Arabic has two basic different types of sentence structure she calls the topic-prominent and the subject-prominent type, each of which has its own discourse function (Brustad 2000: 315). Ingham’s proposal is especially interesting for us on two grounds: firstly, it is based on information structure, and secondly, it takes intonation into account. Ingham’s account of Najdi Arabic sentence types was also employed by Holes (2010) for Gulf Arabic.

In his study on word order in Modern Arabic vernaculars, Dahlgren (1998), following Hopper (1979) relates word order to foreground (representing the main story line) and background information (Dahlgren 1998: 63). He is able to show that in narrative discourse, the colloquials of the Eastern Mediterranean and the Mesopotamian colloquials exhibit 70-80% VS order in foregrounded events concerning the main story line and 70-100% SV order in background information (cf. Dahlgren 168f.). Interestingly, his results for Egyptian Arabic are not conclusive. Three of the investigated EA narrative texts showed approx. 40% VS syntax in the foreground and 15% in background information (p. 169), while the other two texts follow the pattern of the Eastern colloquials. Dahlgren (p. 219) thus concludes that in most dialects VSO is the dominant (unmarked) word order for foreground sentences and SVO for background sentences. This conclusion, however, does not hold for Egyptian Arabic. Dahlgren (p. 189f.) tentatively suggests that in (parts) of Egypt, contact with other languages may have led to a shift to SV order. Furthermore, foreground and background have also been shown to correlate with tense, aspect and mood. While foreground is associated with past tense, punctual and perfective aspect and realis mood, background exhibits preferably present, future, habitual tense, durative and imperfective aspect, irrealis mood (Dahlgren 1998: 61, 72). In a corpus study of MSA texts Pashova (2003) reports similar results for MSA.

The EA text corpus also shows instances of both word orders, SV and VS. The fully-annotated part of the corpus (excluding the experimental parts) was divided in two: a part including predominantly narrative genres (CorpNarr) and another part including predominantly non-narrative genres (i.e. dialogues and expository monologues; CorpExp). The frequency of SV and VS clauses was determined, counting only those instances that have either a subject directly followed by a verb or a verb directly followed by its subject, i.e. SXV and VXS clauses are not included in the count. Furthermore, subjectless sentences were counted separately. Table 1 shows the figures of the count.
Tab.1: SV, VS and subjectless sentences in the EA text corpus.

<table>
<thead>
<tr>
<th></th>
<th>CorpNarr</th>
<th>CorpExp</th>
<th>TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>VS(X)</td>
<td>116</td>
<td>8%</td>
<td>86</td>
</tr>
<tr>
<td>SV(X)</td>
<td>434</td>
<td>30%</td>
<td>533</td>
</tr>
<tr>
<td>no overt S</td>
<td>896</td>
<td>62%</td>
<td>1171</td>
</tr>
<tr>
<td>all clauses</td>
<td>1446</td>
<td>45%</td>
<td>1790</td>
</tr>
</tbody>
</table>

Firstly, the narrative corpus is marginally smaller than the 'expository' corpus. It comes as no surprise that over half of all predications (64%) do not contain an overt subject. Interestingly, however, these are also evenly distributed between the two corpora. Also the percentage of the SV clauses exactly matches that of the distribution of clauses in general. As expected, the narrative corpus contains relatively more VS clauses (57%). In relation to the overall amount of clauses per corpus, this percentage of the VS clauses is even higher (8% of all clauses in the CorpNarr versus 4.8% in the CorpExp). Judging on these figures, text type alone is obviously not a decisive factor for determining the word order. A detailed study of the discursive functions of VS and SV sentences within the text will probably be more revealing. For the present purpose, however, let us look at two narratives (two frog stories) of the corpus as recounted by two different speakers, Text 1 in (5) and Text 2 in (6).

(5) Text 1

Mido_F0_01_03_01

'And then, Mido woke up in the morning...'

Mido_F0_01_03_07-12

'... and the dog also kept looking around helping Mido, but Mido did not find anything at all.'
Mido went down quickly...

His dog, standing beside him,

saw a honeycomb full of bees.'

And then Mido left the mountain mouse

and went away. In the meantime,

And Mido found another hole in a tree...

And the gazelle reached a high place...
Mido went [in his room] to sleep, together with his dog.

And then the frog escaped from the glass. When Mido woke up...

Mido kept looking around in every corner of the room...

The striking difference between the two texts concerns the occurrence of VS structures. While Text 1 (speaker F0) exhibits a number of VS sentences, the short passage from Text 2 (speaker M1) does not show any VS structures. In general, speaker M1’s stories only contain VS structures in clearly thetic utterances such as tˤiliʕ it-l-u bu:ma (came.out-for-him eagle-owl) ‘an eagle-owl appeared’. Interestingly, when I discussed the written down version of the text with two of my assistants, they unanimously judged the SV sentences as ‘wrong’. One reason they gave me was that it was not ‘correct’ to have a coverb precede the content verb, i.e. they considered the order S CV V to be illicit. However, from a purely grammatical point of view there is nothing wrong with these structures. All sentences used by speaker M1 are grammatically sound. What my assistants meant by judging them as ‘wrong’ was their appropriateness in the text genre at hand. We only have to look at the English translation to see that the second text is not fully coherent. Speaker M1 simply is a bad narrator. My main language assistant and the other speaker involved in the elicitation session were both professional narrators having worked for an institution organizing events for children (birthday parties and the like) and story-telling belonged to their daily routine.

Let us consider some of the differences in the texts in some detail. Text 1 describes the activity of searching for the frog as wi-ʔašad ik-kalb ji-dawwarʕ huwa kama:n (and-kept-the-dog searching he also), while text 2 describes it as mi:du ʔašad ji-dawwarʕ (Mido kept searching). What now is the reason for this difference as both versions are clearly grammatical? A possible reason could be topic-shift, for instance. In text 2, Mido is a continued topic and in text 1 the topic is shifted to the dog. As we will see in the remainder of this chapter, explicit mentioning of a new or shifted or contrastive topic using a full lexical NP or an independent pronoun in the preverbal position is a very frequent
construction in EA. Thus, the postverbal position of Mido in text 1 cannot be explained by topic-shift. In terms of grammar the two sentences are also identical, both containing a coverb and a lexical verb. Let us now look at other instances of VS order in Text 1, such as \textit{wi-gat il-γaza:la ʕandā maka:n murtafiś} (and came the gazelle at [=to] a place high) or \textit{nizil bi-surʕa mi:du} (went down quickly Mido). What they do, is expressing some suddenness and thereby adding some vividness to the story. A second observation that seems to relate to that fact is the type of verbs that trigger VS or - conversely - SV constructions. The verbs \textit{gi(h)} 'come' and \textit{nizil} 'go down' are motion verbs and \textit{sˤaḥa} denotes a change of state. At least the first one belongs to the group of verbs that denote appearance or disappearance that have frequently been found with thetic utterances. It is well-known that some languages, i.e. languages with flexible word order, tend to use VS order for thetic utterances. Text 1, however, also shows instances of SV order. The issue of theticity will be the topic of the next chapter, but let me point out here that I do not think that the utterances of concern in Text 2 should be classified as thetic as they clearly involve topics, sometimes even continued topics as in the case of the gazelle-sentence. What, then, is the difference between the sentence \textit{gat il-γaza:la ʕandā maka:n murtafiś} (came the deer to a high place) and \textit{il-γaza:la gat ʕandā maka:n murtafiś} (the deer came to a high place)? For one thing, the first sentence adds some vividness to the story and takes forward the story line, while the second one is simply matter-of-fact. This shows that it is not enough to operate with narrow notions of information structure of individual sentences here, but that we have to take into account many other factors such as narrative structures and other discoursal factors like the shift in discourse topic or the beginning of a new episode, foregrounding and backgrounding. In addition, there are grammatical restrictions on word order (e.g. in some dependent clauses) and others imposed by the TAM features of verbs - e.g. \textit{mif la:mi} (Mido not finding) involving a participle predicate is most probably the only acceptable word order, \textit{mif la:mi mi:du} (not finding Mido) is at least highly marked and could only be expected to occur under very special pragmatic circumstances. Moutaouakil (1989: 85) claims for Classical Arabic that nominal predicates never precede their subjects. Additionally, verb semantics and other semantic features such as animacy seem to play a role. And finally, we must not forget that there may be an influence of traditional story-telling that might still be guided by the norms of Classical Arabic, being taught at school, norms that may also have been preserved in the traditional genre of story-telling. As already noted above, the investigation of the interaction of the different factors involved in word order phenomena calls for an independent in-depth study which is not in the scope of the present book.

Now, the experimental QUIS corpus that was elicited using different types of stimuli, mostly simple picture cards, but also some short films and a map task, contains almost no VS sentences at all. The

\footnote{It has to be noted that the second example also involves a topic shift.}

\footnote{The phonological difference between the verbs in text 1 (\textit{sˤaḥa}) and text 2 (\textit{sˤiḥi}) is due to dialectal differences, speaker F0 being from Alexandria and speaker M1 from Cairo.}
only instances found are with verbs of appearance, especially *gih* 'come', which may also be used as an aspectualizing coverb (Maas 1995; Ahmed & Selmy 2000). Even instances that involve a (sudden) change of scene were described using an SV sentence (7).

(7) 35ARZ-C01GNT-04M1200

\[il\text{-}makan\]  \[it\text{-}ayajar\]  \[humma\]  \[?a\text{-}id\text{-}i:n...\]
DEF-place  change.INTR.3SG.M  3PL  sit.PTCP-PL

'The place changed, they are sitting...'

The results of this preliminary analysis of word order in the corpus and what we know from the literature suggest two things: in EA, (1) SV is clearly more frequent than VS and (2) VS is contextually more restricted than SV, i.e. it is mainly restricted to thetic sentences (see Chapter 6) and to narrative discourse where the foregrounded main story is partly expressed by VS sentences, predominantly by COV-S-V structures. The problems with identifying the basic or unmarked word order in a language only in terms of frequency is made exceedingly clear by Dryer (1995: 119):

> [...] frequency will actually vary from discourse type to discourse type, from text to text, and from subtext to subtext. Where word order is sensitive to aspect, a word order that is more frequent in narrative may be less frequent than other discourse types [sic]. Where word order is sensitive to contrasting participants in a text, a text with two primary participants may exhibit rather different frequencies from one with a single primary participant.

These facts show that a description of Classical Arabic and MSA as VSO and the modern colloquials as SVO does not do justice to the linguistic reality. They point to the importance of discourse factors and a thorough analysis of actual texts to identify the functions of the individual word orders. Nevertheless, it is obvious that the Arabic varieties clearly differ in how they alternate between VS and SV order and it is also not sufficient to relate the individual word order phenomena to specific functions, stating that all types of Arabic are SV/VS alternating languages. It rather seems that at least some of the varieties, e.g. EA, have a "basic word order" which I take to be synonymous with what Dryer (1995) calls "pragmatically unmarked" word order. Dryer defines pragmatically unmarked word order as the default "if there are concise ways to characterize the situations in which other word orders are used, with the pragmatically unmarked word order being most easily characterized as the order that is used elsewhere." (Dryer 1995: 105f.; emphasis in the orig.). We have seen that it is virtually always possible to use SV order, with the exception of a thetic sentence with an indefinite subject which is ungrammatical with a SV order (9), i.e. if the subject is not introduced by an existential quantifier or some other presentative element. Thus, SV order is by far the most frequent word order in EA.

(8)  

\[^{*}\text{bu:ma}\]  \[^{*}\text{t\text{-}ala\text{-}i\text{-}l\text{-}u}\]
eagle.owl  come.out-3SG.F-for-3SG.M

'An eagle owl appeared before him.'

[^4]: It should most probably be "... in other discourse types".
A further piece of evidence in favour of SVO as the *canonical sentence structure* comes from hanging TOPICS. In Section 5.3.1 we will see that all kinds of arguments may be expressed extra-clausally in the left periphery of the sentence. In the case of direct objects and prepositional objects, the referent of the topic expression is referred to within the clause by means of a pronoun. It is also possible to put two full lexical arguments before the verb, however with the restriction that the second one be the subject (9). This may be taken to indicate that the subject belongs to the clause and is not an extra-clausal constituent that is referred to within the clause by means of the inflectional ending on the verb.

(9)

a. Sa:mi ga:b il-kita:b li-Salma
S (S) brought the-book (DO) to-Salma (IO)
'Sami brought Salma the book.'

b. il-kita:b Sa:mi ga:b-u li-Salma
the-book Sami brought-it to-Salma

c. Salma Sa:mi gab-la-ha l-kita:b
Salma Sami brought-to.her the-book

d. ?Salma il-kita:b gab-hu-lha Sami
Salma the-book brought-it-to.her Sami

e. *il-kita:b Salma gab-hu-lha Sami
the-book Salma brought-it-to.her Sami

I will thus assume that in EA the basic word order or the "canonical sentence structure" (Lambrecht 1994) is SV(X). For the time being, I will draw the preliminary conclusion that preverbal topics in the case of SV structures are 'explicit' topics in the sense of being explicitly set as the "starting point" or the "hitching post" to use Chafe's (1976: 44) expression. Postverbal subjects, on the other hand, may function as 'implicit' topic expressions, which means that their referents are topical in the proposition at hand in the sense of being the entity about which a predication is made, but without the aboutness relation being stated explicitly. On the other hand, there are also instances of topic shift expressed in VS sentences and there is always the possibility of a certain tonal contour and prosodic break after the subject to highlight the topicality of the postverbal subject (cf. 5.3.2). 'Implicit' topics are not foregrounded, i.e. their topical relation to the whole proposition is not expressed formally. In other words, it is the predication that is foregrounded and not the entity about which something is predicated. These functions between the two sentence types are reflected in Holes' distinction between *entity-oriented* and *event-oriented* sentences. Ingham's classification of *uninodal* and *binodal* sentences is akin to Maas' use of the notions of *narrative* and *qualitative* predication (Maas 2011) which goes back to Bloomfield's distinction between *narrative* and *equative* predication (Bloomfield 1935 [1970]: 173f.). Both Ingham's and Maas' characterization of the binodal/qualitative type, however, differ from the original Bloomfieldian typology as they also include 'verbal' sentences, i.e. sentences containing a verb, and not only 'nominal' ones whose predicate is nominal (NP, AP or PP). Matras (1995) suggests that VS order in Romani has a "connective function" in narrative discourse, while SV order serves to detach the proposition from the preceding context. This is based on the 270
assumption that a constituent in preverbal position serves as a base for the predication to come. If this base is lacking, the hearer has to retrieve the thematic element from the preceding discourse and thus connecting the predication to the one before.

To summarize the assumptions made in this subchapter, I will assume that SV, or rather SP, is the canonical sentences structure in EA, being the pragmatically unmarked and, as an epiphenomenon, also the most frequent word order. I will further assume that the difference between the two sentence types SV and VS is captured by the classification of qualitative vs. narrative predication. I do not assume, however, that this distinction is equivalent to 'all-new' or 'sentence focus' vs. topic-comment as implicit in Ingham's definition of uninodal and binodal sentences. I rather assume that preverbal subjects - depending on their prosodic coding - are marked as topics, while postverbal subjects may also function as aboutness topics, but they are mostly not marked as such. In this function, they resemble implicit continued topics that are only coded in the inflectional ending of a verb. However, a detailed analysis of the syntactic facts is yet a topic for future research. Especially, the proposal made by Matras (1995) for Romani deserves the investigation of the EA data from this perspective.

5.2 Topics and TOPICS

We have now clarified the question whether VS order necessarily indicates theticity posed in the introductory section of this chapter and reached the conclusion that this is not the case. If this is true, it follows that topics can occur postverbally or preverbally. We will see that in EA, topics also may occur sentence-finally, or rather in the right periphery of a sentence. In spoken language we not only rely on morphosyntactic marking, but also on prosodic cues. Nevertheless the identification of an aboutness topic is not always straightforward, as already noted above.

Let us first have a look on Brustad's account of topics in Spoken Arabic. Following the older tradition of information structure research (Halliday 1967; Chafe 1976; Li & Thompson 1976), she makes a distinction between subjects and topics which presumes that a topic is always sentence-initial. Brustad (2000: 361) classifies Arabic varieties as topic-prominent and subject-prominent at the same time according to the typology suggested by Li & Thompson (1976) whose distinction between subject and topic is exemplified in (11).

\[(10) \quad (= \text{ex. 1-2, Li & Thompson 1976})
\]
\[a. \quad \text{John} \quad \text{hit Mary.} \quad \begin{array}{c} \text{Subject} \\ \text{Predicate} \end{array}
\]
\[b. \quad \text{As for education, John prefers Bertrand Russell's ideas.} \quad \begin{array}{c} \text{Topic} \\ \text{comment} \end{array}
\]
According to another view, following work by Gundel (1976) and Reinhart (1982), the referent of John in (10a) would be considered a sentence topic just as in (11b). The left dislocated constituent as for education is a topic as well. Unlike John, however, which is an argument of the verb and part of the clause itself, the left dislocated topic is outside the clause and not in the scope of the verb. I will call this type of topics free TOPICS. Functionally, there is also a difference between the two topics, which will be explained below.

An explication of the term aboutness relies on Strawson's "Principle of Relevance" (cf. Reinhart 1982: 6), which states that we do not randomly convey "isolated and unconnected pieces of information" but rather that we "add information about what is a matter of standing current interest or concern" (Strawson 1964: 97). Based on this view, Lambrecht (1994: 131) defines

[a] referent as the topic of a proposition if in a given situation the proposition is construed as being about this referent, i.e. as expressing information which is relevant to and which increases the addressee's knowledge of this referent.

With this in mind, we may now reconsider the examples in (11). It is uncontroversial that both sentences increase our knowledge about John, but does the comment in (11b) really increase our knowledge about education? Following Chafe (1976) and Li & Thompson (1976), we could make a distinction between topics and subjects and take the subject as "what we are talking about" (p. 43) in continuation of the Aristotelian concept (cf. Ch.1) and the extra-clausal constituent as the topic, but now consider the example in (11). Although the construction is very similar to that in (10b), we intuitively feel that the proposition is at least as much about Rosa as it is about John.

(11) (= ex. 4.24a, Lambrecht 1994)

As for Rosa, John didn't really love her.

Just as in (11b), the formal construction 'as for...' - which is in fact used as a test for topichood (Gundel 1974) - suggests that Rosa is an aboutness topic of the whole sentence. Contrary to the situation in (10b), however, the proposition John didn't really love her can indeed be construed as being about Rosa and to increase our knowledge about her, whereas it is at least doubtful that John's preferences increase our knowledge about education. The main difference of the syntactically similar constructions is the syntactic role, the referent of the extra-clausal constituent plays within the clause. The referent of the expression Rosa is indeed referred to again within the clause by the pronominal form her, thus being an argument of the main proposition. This goes to show that extra-clausal constituent may in fact have different syntactic roles in the clause, just as they may have different pragmatic roles. As a corollary of the discussion of examples (10) and (11), we may differentiate between topics that are arguments and topics that are not arguments. As already noted in Chapter 1, Chafe's distinction between subject and topic is similar, though not identical and Matić (2003) referred to it as direct and indirect topic. Although both types are subsumed under the cover term topic in most
approaches (Lambrecht 1994; Maslova & Bernini 2006), the discussion of aboutness topic is in practice frequently restricted to argument topics (Reinhart 1982; Lambrecht 1994; but cf. Büring 1997 for another view).

As we have seen so far, it is certainly difficult to identify which of the argument topics counts as the aboutness topic or as the primary topic (Lambrecht 1994: 147). At the same time, the coding of argument topics is very much a question of morphosyntax, which I will not be dealing with here in the first place. We have also seen that a topic does not necessarily occupy sentence-initial position, contrary to Halliday's (1967) or Dik's (1978) themes or Chafe's (1976) and Li & Thompson's (1976) topics. However, the special affinity topics have to the sentence-initial position is grounded in the presumption of cognitive linearity, which is also implicit in Halliday's definition of theme: "The theme is what is being talked about, the point of departure for the clause as a message; and the speaker has within certain limits the option of selecting any element in the clause as thematic" (Halliday 1967: 212).

Given the vagueness of the aboutness-definition and the problems in identifying the (primary) topic of a sentence, I will therefore proceed from strictly formal considerations and look at constructions at the left edge of a sentence. A classification of these constructions will - in line with the general theme of this book - be primarily based on their prosodic coding but also consider the semantic content of these left-most constituents and their syntactic roles. Based on these formal criteria and their contextual embeddedness, the different pragmatic functions of these constructions shall be identified. As already outlined in Chapter 1, I will thus distinguish between entity topics and other types of topics, as I will suggest a formal category of thematic TOPICS based on sentence position prosodic shape. Let us first look at entity topics. We have already seen that a topic is not necessarily located at left edge of the sentence. In the preceding examples we have encountered only entity topics that are arguments of the proposition. As shown in (12), such arguments very frequently are placed in the left periphery of a sentence in EA.

(12) constructed examples

a. *salma ma-bi-t-ħibb-aʃ goz-ħa*
   S  NEG-IND-3G.F-NEG husband-3SG.F
   'Salma does not love her husband.'

b. *salma goz-ħa ma-bi-j-ħibb-ahaː-f*
   S  husband-3SG.F NEG-IND-3G.M-3SG.M-NEG
   'Salma, her husband does not love her.'

c. *salma gib-t-əl-ha sandwitʃə fuːl*
   S  bring-1SG-to-3SG.F sandwich beans
   'Salma, I got her a Ful-Sandwich.'
In all three sentences in (12) Salma is clearly a/the topic of the sentence. The syntactic role the expression Salma has within the proposition is variously that of the subject (12a), the direct object (12b) and the indirect object (12c), which are all within the valency of the verb. However, consider an example such as (13), which also exhibits a very frequent construction in EA.

(13)

\[
\text{salma ʔuxtaha ma-ga-t-f}
\]

S sister-3SG.F NEG-come-3SG.F-NEG

'Salma's sister did not come.'

In (13), Salma is not an argument of the proposition. Nevertheless she is a referent that is the topic of the sentence just as in the sentences in (12) or as Rosa is in (11). In (13), the proper noun Salma is the possessor in a possessive construction and the possessed object is an argument. The function of this topic is to identify the sister whom the hearer presumably is not acquainted with and to relate it to the familiar Salma, thereby introducing the otherwise unidentifiable sister, who is the other (possible) topic of the sentence, into the common ground. Structures such as the one in (13), (12b) and (12c) are the reason for Brustad's classification of Spoken Arabic as topic-prominent. The preverbal listing of the individual referents, one after the other, facilitates the hearer's task of identifying the referents first and then resolving the predication about (the final) one of them. The linearizing procedure of having pre-clausal elements identifying the referents of a clause, which are referred to by pronominal elements within the clause is following what Lambrecht (1994: 184-191) calls the Principle of Separation of Role and Referent (PSRR). The more complex construction ʔuxt salma 'Salma's sister' is violating the cognitive principle of moving from the more familiar to the less familiar. It is well-known that spoken language tends to use topic constructions as the one above more frequently, whereas written discourse is not as dependent on cognitive linearity and may make use of more integrated and complex structures (Miller & Fernandez-Vest 2006). However, in addition to the anchoring function, the topic may also be interpreted as an aboutness topic. Imagine it is Salma's birthday and she has invited some guests. Salma had hoped that her sister would also come to the party, but she eventually did not show up. One of the guests who arrived late finds Salma in a rather bad mood. Upon asking one of the other guests what was wrong with her, this other guest might answer Salma ʔuxtaha magatf 'Salma's sister has not come'. Here clearly, the sentence is more about Salma than it is about the sister. In case we did not already know, we have learned that Salma has a sister and that this sister was supposed to come, but did not show up. But at the same time, we also know something about Salma's sister, namely that she did not come. And finally we know something about their relation, namely that Salma expected her sister who did not meet these expectations - with all the inferences that may be drawn from this circumstance. If the ones who witnessed this situation should ever meet later, and by any chance Salma's sister would be mentioned - it could easily happen that someone said something like Oh, you mean the one who did not show up at the birthday party. I will not delve into the problematic issue of the exact meaning of aboutness here, let alone the issue of
cognitive processing and storage here. For the time being, I will stop at this point and assume that both instances, Salma and her sister are topics.

We may summarize that the topicalization of Salma apparently may serve two functions, that of indicating an aboutness topic of the sentence and that of anchoring another aboutness topic of the sentence in the common ground. I.e. an entity topic may be what the sentence is about, but it may also only serve as a cognitive link to the universe of discourse shared by the interlocutors. The other conclusion we may draw is that an entity topic may, but need not be an argument of the sentence.

However, not every aboutness topic is an entity, and not even referential. The following example from an Egyptian TV series broadcast in Ramadan 2010 shows a verbal noun in topic position. The interlocutors in the story are having an argument about a piece of land that belongs to one of them. This person knows that his vis-à-vis would like him to sell the piece of land and he wants to make it clear from the beginning that this is not going to happen. The topical expression is a formally nominal constituent outside of the clause, but it is not referential. Nevertheless 'selling' is clearly the aboutness topic of the sentence.

(14) overheard example
\[\overline{be:}$ʃ \ overheard \ \overline{mif} \ h-a-bi:ʃ\]
\[\text{selling.VN NEG FUT-1SG-sell}\]
'Sell I will not.'

It has been claimed that indefinite and quantified phrases cannot be topics, but indefinite and non-referential phrases appear in typical topic-comment constructions (cf. Ward & Prince 1991; Birner & Ward 1998). We will see below in Section 5.8 that this is also true for EA. The interim summary of the facts identifies 'topic' as a very heterogeneous category, both functionally and formally. Neither need an entity topic be an aboutness topic, nor is the aboutness topic necessarily an entity. Even if an/the (entity) aboutness topic can be unambiguously identified, it may appear in various morpho-syntactic manifestations, either as an inflectional agreement morpheme, as an independent pronoun or as a full lexical NP that may appear inside or outside the clause, preverbally (SV) or postverbally (VS), left-dislocated or right-dislocated.

This is why some researchers have questioned the utility of the notion topic. However, intuitively the formally and functionally different constructions encountered in natural languages have something in common that has given rise to the notion of topic and as Jacobs (2001: 675) remarks "the notion of topic will help to connect current and future research [...]" and is certainly valid as a heuristic notion. Jacobs (2001), however, questions the validity of topic as a grammatical category and the possibility of a unitary explanation of the different phenomena commonly subsumed under the notion of topic. He offers a more flexible and multidimensional approach to topic-comment structures, suggesting four semantic attributes that are shared by different topic constructions. Importantly however, not all
semantic attributes have to be shared by all constructions. These "dimensions" are informational separation, predication, addressation and frame setting. Separation has to do with the old thetic-categorical distinction (Kuroda 1972; Sasse 1987; inter alia; cf. Chapter 6), i.e. topic-comment structures are categorical or bipartite involving two cognitive steps in their processing: (i) introducing a topic and (ii) adding a comment. Jacobs (2001: 678f., footnote 42) notes that the only dimension that is necessarily associated with topic-comment is separation and that this attribute seems to be the most important trait of what he calls "prototypes" of TC [topic-comment] constructions (we will come back to some of these notions shortly). On this view, a topic-comment construction is identified as such due to its functional and syntactic similarities with the prototype (p. 644).

In the following, I will only consider constructions that exemplify informational separation formally expressed in the bipartiteness of the whole construction involving a sentence initial part that I will call thematic TOPIC - as an umbrella term for functionally and morpho-syntactically heterogeneous constituents of a proposition. Formally such TOPICS frequently, but not necessarily, coincide with left-dislocated phrases. Metaphorically, TOPICS can be characterized as the "hitching post for the new knowledge" (Chafe 1976: 44) or the "point of departure for the clause as a message" (Halliday 1967: 212). Thus, thematic TOPIC is more like the Hallidayan notion of theme than the notion of topic discussed above. In addition, I will assume a syntactic domain of right-dislocation (henceforth postfield) characterized by two different types of intonation. I will also assume a heuristic category of entity topic, as outlined above, and investigate the relationship between the entity topic (henceforth topic) and the thematic TOPIC (henceforth TOPIC), i.e. the left-peripheral position, as well as the postfield, i.e. the right-peripheral position. In doing so, I will try to differentiate between types of TOPICS (and topics) or TOPIC constructions from a formal and functional point of view.

Before I start the empirical investigation, one word on the annotation of the corpus is in order. When doing the syntactic annotation, I found it convenient to follow the German tradition and divide a clause into three parts, an obligatory middle field starting with the verb or the nominal predicate including all arguments and adjuncts following the verb, an optional preverbal domain including all constituents pertaining to the proposition at hand that precedes the verb (prefield), and an equally optional, but much less frequent post-clausal domain, the postfield. Contrary to the other two constituents, the tail was not identified on a purely syntactic basis, but rather due to a combination of syntactic and prosodic features. This method made it possible to utilize the multiple layer search options provided by the database ELAN and cross-classify individual constituents with syntactic position, prosodic categories and information structural notions. Note, however, that the identification

---

5 In the present investigation, the dimension of separation is taken to be the most important feature of TOPIC-COMMENT structures. I will also make use of Jacobs' notion of frame-setting below. His notion of addressation is meant to supplant the traditional notion of aboutness. Despite its merits it does not allow for a unified account for all TOPICS either and has not successfully replaced the notion of aboutness in the literature. Therefore I will stick to the term aboutness in this study.
of the prefield as the pre-verbal domain, does not imply a syntactic analysis that assumes the subject to be extra-clausal or a division of the sentence into a S and a VP. In (15) some examples of such annotations are given.

(15) Syntactic annotation

a. Ihsan_FG1_07

\[
\begin{array}{c}
\\ \text{da} \\
\\ \text{feʔ} \\
\\ \text{t’abí:zi} \\
\\ \text{jaʃni}
\\ \text{DEM.SG.M} \\
\\ \text{thing} \\
\\ \text{natural} \\
\\ \text{DM} \\
\\ \text{PREFIELD} \\
\\ \text{MIDDLE FIELD} \\
\\ \text{POSTFIELD}
\\ \end{array}
\]

'That's only natural, I mean.'

b. Knee_MG1_07

\[
\begin{array}{c}
\\ \text{fi-l-hal-a:t} \\
\\ \text{il-mit?axxar\=c-a} \\
\\ \text{il?-alam} \\
\\ \text{bi-ji-b?a} \\
\\ \text{in-DEF-case-PL} \\
\\ \text{DEF-late-SG.F} \\
\\ \text{DEF-pain} \\
\\ \text{IND-3SG.M-be/become} \\
\\ \text{PREFIELD} \\
\\ \text{MIDDLE FIELD}
\\ \end{array}
\]

In advanced cases, the pain is...

\[
\begin{array}{c}
\\ \text{mawgu:d} \\
\\ \text{mustamirr / hatta ?a\=na?a} \\
\\ \text{irˤa:ha} \\
\\ \text{POSTFIELD}
\\ \end{array}
\]

always present, even in rest position.'

c. Mido_M1_01_01

\[
\begin{array}{c}
\\ \text{mi:du} \\
\\ \text{daxal} \\
\\ \text{na:m} \\
\\ \text{huwwa} \\
\\ \text{w-ik-kalb\=} \\
\\ \text{bta:ʕ-u}
\\ \text{M} \\
\\ \text{enter.3SG.M} \\
\\ \text{sleep.3SG.M} \\
\\ \text{3SG.M} \\
\\ \text{and-DEF-dog} \\
\\ \text{POSS-3SG.M} \\
\\ \text{PREFIELD} \\
\\ \text{MIDDLE FIELD} \\
\\ \text{POSTFIELD}
\\ \end{array}
\]

'Mido and his dog went to sleep.'

5.3 Thematic TOPICS

A thematic TOPIC in EA formally comes in two basic shapes, a rising contour I call leading TOPIC and another type I call linking TOPIC. In line with what has been said about leading and linking tones and tunes in 2.2.4 and 2.2.5, a leading TOPIC may either involve a rising trendline over default accents or just one leading accents (Figure 1a).

A linking TOPIC may again come in two different shapes. A downtoned linking TOPIC has a flat or almost flat, low toned horizontal contour with little prominence (Fig. 1d). A prominent linking TOPIC (henceforth leading-linking TOPIC) involves strong lexical prominences, either coded by default accents or H accents in a high-toned horizontal trendline (Figure 1b-c).
Fig. 1: Leading and linking contours: (a) leading contour, (b) prominent linking contour with default accents, (c) prominent linking contour with monotonal accents (d) downtoned low linking contour.

Furthermore, the final accent of a TOPIC may either be leading, which may involve a high boundary tone, or linking. Figures 2 and 3 schematically illustrate leading (Figure 2) and linking accents (Figure 3) again. For more details on the model see Chapter 2).

Fig. 2: 4 types of leading accents. The shaded area represents the stressed syllable: (a) inverted accent (b) rise with an H boundary tone resulting in a high rise, (c) simple rise, (d) high rise.

Fig. 3: 2 types of leading-linking accents. The shaded area represents the stressed syllable: (a) spreading of H tone (b) default accent with a high next low.

Note that the above phenomenological account of leading and linking tones and tunes does not imply categorial status for any of these instances. Rather, this classification is a pre-theoretical one. To arrive at a proper categorization of tones and tunes still requires a good deal of research to see how they differ functionally and especially how tones and tunes interact. Just one remark, I do not believe that the accents in (3a) and (3b) are functionally different. It rather seems that speakers with a very careful and accurate articulation are able to keep fundamental frequency at the same level for a certain time while more casual articulation leads to a slump in the contour. On the other hand, I am not sure whether a simple rise as in (2c) should be classified together with the high rises in Figure 2 or, as is tentatively suggested here, or together with the linking type. Finally, the difference between tunes (1b) and (1c) may well be a difference in phonological realization due to speech tempo. There is, however, a difference in prominence between the contours, depending on the excursion size of the accents. In sum, many differences between the individual tones and tunes depicted above are likely to be phonetic
and not systematic. However, I believe that at our present state of knowledge it makes sense to give a complete descriptive account of the phenomenon and to venture phonological analysis along with the descriptions. Such a tentative analysis that I would like to suggest here is the classification of *leading* vs. *linking*. Such an analysis will be presented below.

In this study I will investigate the thematic TOPICS that in the EA text corpus and rely on some of the results of the experimental data, especially those elicited with the QUIS questionnaire (Skopeteas et al. 2006). Although the unmarked case of a topic in EA is expressed by an inflectional ending - as we will see below -, such topics are not considered in the present study, as there is not much to say about the prosody of inflectional endings. The cases with postverbal subjects (see 5.2) are also excluded. This means that we will ignore 70% of all clauses occurring in the EA text corpus and concentrate on the remaining 30%, which comprise the cases listed under (16) below. These are what Jacobs calls prototypical "TC [topic-comment] constructions" (2001). It follows from such an assumption that the sentence containing an unmarked topic expression is not the one cast in a prototypical TC construction. On such a view, a prototypical TC construction is thus not the most frequent sentence type to be found in spontaneous speech. In the following subsections, I will investigate the EA corpus with respect to the different TOPIC constructions identified in the literature and investigate their information structural functions as indicating *aboutness* or *addressation* or *anchoring* and *framing*.

In doing so, I will try to make some generalizations about the prosody of these constructions. For the sake of descriptive completion, I will give a short overview on discourse markers in TOPIC position in Section 5.4. Sections 5.5 and 5.6 will be looking at morpho-lexical TOPIC marking and topics in the postfield, respectively. In addition, the use of the identified constructions will be cross-classified with cognitive status (5.7). Section 5.8 will be discussing *quantificational topics*. Finally in Section 5.9, the results of the investigation presented in this Chapter will be summarized and adduced as evidence with respect to some of the pending theoretical questions concerning topicality. Specifically, I will make a comparison between the EA data and some data of other languages described in the literature and draw some conclusions concerning theoretical and typological issues.

### 5.3.1 Entity topics as TOPICS

Research on TOPICS in European languages has shown that intonationally marked topics are mostly 'contrastive' or at least involve a topic shift. Some researchers only classify this type of constituents as 'topics', while other topics are simply part of the background (Büring 1997, 2003). Contrastive topics

---

6 Contrastive should be taken in a wide sense here. For the problematic issue of contrastivity cf. Chapter 4.5.4.
were variously referred to as *contrastive topics* (Jacobs 1997, Lambrecht 1994, Büring 1997, 2003; inter alia) or *contrastive foci* (Chafe 1976; Ladd 1980; Gussenhoven 1983; van Hoof 2003) or as topics containing a focus or probably topic and focus at the same time, or, as has become more common among more recent authors, as a topic containing a feature or a category of 'contrast' that can be combined with both, topic or focus (Vallduví & Vilkuna 1998; Molnár 1998, 2006; Krifka 2007; Hedberg & Sosa Juan M. 2008; Steedman 2008; Repp 2010).

(16) presents an overview of what has been classified as typical topic-comment constructions in the literature with a prominent (i.e. in most accounts ‘contrastive’) topic. The topic constituent is given in italics. It is crucial that the constituent is accented.²

(16)

a. *Peter* schläft. (German; Jacobs 2001)

b. *Ronald* made the hamburgers. (Chafe 1976)

c. *Évat Janós* várta a mozi elott. ‘Eva (ACC) Janos (NOM) waited for in front of the cinema’ or ‘Eve was waited for in front of the cinema by John.’ (Hungarian, É. Kiss 1995)

d. *This room*, it really depresses me. (Gundel 1974)

e. *Die Gerda*, die mag ich wirklich sehr. ‘Gerda, I really like her a lot.’ (German; Jacobs 1984)

f. *With Rosa*, Felix went to the beach. (Reinhart 1982)

g. *As for the homeless*, where did their food come from? (Maslova & Bernini 2006)

h. *Was Peters Geburtstag betrifft*, so habe ich noch kene Idee für eine Geschenk. ‘As regards Peter’s birthday, I don’t yet have an idea what to give him as a present.’ (German; Jacobs 2001)

i. *Sakana-wa* tai ga oisii. ‘Fish, red snapper is delicious.’ (Japanese; Li & Thompson 1976)

j. *Auf der Neununfünfzigsten Straße* habe ich die Schuhe gekauft. ‘On 59th street I bought the shoes.’ (German; Büring 1997)

k. *In meinem Traum* war Peter ein Krokodil. ‘In my dream, Peter was a crocodile.’ (German; Jacobs 2001)

l. *Alle Politiker sind NICHT korrupt.* ‘Not all politicians are corrupt.’ (German; Jacobs 1984)

What all the examples have in common is prominence on the topic expression. A prominent or accented subject distinguishes the sentences in (16a) and (16b) from comparable sentences with a deaccented subject that is assumed to express a continued topic. These are prosodically marked, but

² But note that Jacobs distinguishes I-topicalization constructions from multiple focus constructions. The latter would in some theories also be classified as topic-comment.

³ Some authors have also pointed to a special type of intonation that is necessarily associated with the construction (Jackendoff 1972; Büring 1997; Jacobs 2001).
syntactically, they are unmarked. A syntactically marked topic is shown in the Hungarian example (16c), in which the canonical word order is inverted to make the non-subject the topic of the clause. This option is, however, only available in languages with a relatively free word order, such as Hungarian, Russian or German. The construction is, not an option in EA, which can heavily relies on left-dislocated TOPICS as a marked option. Examples (16d-h) are instances of such left-dislocated TOPICS which are defined by their extra- clausal position, some of them also show a co-referential pronoun within the main clause. While the examples (16a-f) are all argument topics, (16g) is a special instance of a hanging TOPIC, which according to its function will be called reference point TOPIC, following Langacker (1993) and Maslova & Bernini (2006). This construction is of special importance in EA and will be discussed thoroughly in Section 5.3.1.3. (16g-h) show free TOPICS that are not syntactically linked to the main clause. In the classification adopted in the present work, (16j-k) are frame TOPICS and will be discussed separately in Section 5.3.4. Finally, (16l) is an instance of what Jacobs (1997) called I-topicalization. This construction, among other functions, serves to invert the scope of negation, i.e. instead of the having scope over the predicate (korrupt 'corrupt'), the negative marker here has scope over the quantifier, turning 'all' into 'not all' (Ladd 1980; Büring 1997; Jacobs 1997). This construction also plays no role in EA as the EA equivalent pays attention to bringing the quantifier into the scope of the negation by means of syntax (17). Generally, an inverted accent (fall-rise), which is one of the formal correlates of I-topicalization9, does not induce contrast in EA.

(17)

\[
\text{mif} \quad \text{kull} \quad \text{is-sijasijj-i:n} \quad \text{fasid-i:n}
\]

\[
\text{NEG} \quad \text{all/every} \quad \text{DEF-politician-PL.M} \quad \text{corrupt-PL.M}
\]

'Not all politicians are corrupt.'

In the following sections we will investigate whether the contrastive implicature that has been suggested for the above topic constructions in German and English is also present in EA. Coming now back to the classification of the constituents at hand as TOPICS or FOCI, I have argued in Chapter 4 that the different concepts of ‘focus’ that are assumed in the literature in fact pertain to different phenomena. The problematic classification of ‘contrastive topics’ does not arise in a model such as the one used here that makes a distinction between different concepts of ‘focus’ - FOCUS as a pragmatically structured proposition and focus as a marker of interest - and that allows for the interaction of the different information structural categories (cf. Chapter 4.5).

However, not every sentence-initial prominent constituent is necessarily a TOPIC. It is common knowledge that a single prominence on a sentence initial subject in English or German indicates subject FOCUS or a thetic utterance. A similar observation holds is true for EA. Although subject

---

9 The other one being an accent on the negative particle in this case, which is why I used uppercase to indicate prominence on *nicht*. 

281
accentuation with accompanying deaccentuation is not typical for EA, the second element is
downtoned (in subject FOCUS) and at least downstepped in thetics (cf. Chapter 6).

EA, just as other Arabic varieties, are so-called null-subject (or pro-drop) languages. That means that
the subject pronoun need not be overtly expressed, the referential information can be retrieved from
the inflectional ending of the finite verb. However, as we will see below, there are many subject
pronouns - which take the independent form in EA as opposed to object pronouns that are suffixed to
the verb - occurring in spontaneous speech. Some of them are in fact obligatory, as finite verbs are not
the only way to express a (verbal) predicate. EA, lacking a proper copula, thus conjoins a subject and a
nominal predicate asyndetically in most cases. This is where a pronominal subject has to be expressed
explicitly. Secondly, active participles are frequently used as verbs. They indicate gender and number,
but not person. I.e. in these cases an overt subject is also obligatory. But finally, there are a high
number of independent (subject) pronouns in the EA corpus occurring together with a finite verb form.
We will look at such cases below.

For the present study 1326 TOPICS from the EA text corpus have been analysed that included at least
one topic as outlined in the previous subsection and were not classified as frames (5.3.2). In the
previous subsection, we have introduced the syntactic properties of TOPIC constructions. In this
subsection we will see how these constructions are realized prosodically in EA. A crude classification
gives us two major types of TOPIC realizations: linking and leading TOPICS as illustrated above. The
results of the present investigation will hopefully shed some light on the following questions:

- What is the relationship between syntactic TOPIC constructions (as
  identified in 17) and prosodic type?
- How do syntactic type, tonal and metrical properties (phrasing) interact?
- What are the informational structural correlates of leading and linking
  TOPICS?
- How does cognitive status influence the prosodic marking of topics?
- May quantified NPs be construed as topics?

Due to technical reasons, the quantitative accounts that will be given in this Chapter will only consider
the final pitch accent within a TOPIC phrase. This admittedly problematic aspect is however mitigated
by the fact that many of the topics involved only consist of a one-word constituent and thus only carry
a single accent (cf. the examples in the following subsections). Because of these technical problems, I
also had to severely restrict the quantitative analysis of the different TOPIC types to subject topics
(5.3.3). The following example illustrates two TOPIC constituents in sequence involving two entity
topics: 'mum' and 'the food'. With the annotations guidelines used here, it was not possible to retrieve
information about the type of TOPICS involved automatically in a corpus search, nor was it possible
to relate the entity topics contained to the whole construction. In future research, I will hopefully be
able to fill that gap.
But let us now return to the investigation of the prosody of all TOPICS investigated here and first consider phrasing. Recall that in the present model, no obligatory one-to-one-relationship between prosodic phrases and syntactic constituents is assumed. On the assumption that phrasing, tonal shape and syntax are a priori independent, we will rather look at their interaction. Also recall that the present model only assumes one type of phrase, namely the intonation phrase (IP), which necessarily involves a rhythmic break that may, but need not be accompanied by a pause. Under these assumptions, we find that most TOPICS are not separated from the rest of the sentence by an intonation break. In fact, only approx. 20% of all examples show a break after the TOPIC.

As for the tones, only a subset of 842 TOPICS that were annotated in detail could be analysed quantitatively. The results indicate that phrasing correlates with the tonal properties of the final pitch accent in an expected way. While 36% of all TOPIC IPs ended in a leading tone (Figure 2), only 10% of all TOPIC phrases that are not followed by a phrase boundary showed leading tones. This goes to show that a rising tone frequently, but - and this is crucial - not obligatorily induces the occurrence of a boundary. As for the prominent linking tone (Figure 3), the distribution between IP-final and IP-medial occurrences is more equal (47% with break, 57% without break). However, if the H tone is spreads, i.e. forms a plateau, there is a much higher chance (2:1) of a boundary to occur afterwards (Table 2). Examples for the individual tones are given below: leading accent (ex. 35, Fig. 9), leading-linking accent (exs. 19, Fig. 4), low linking tone (ex. 21, Fig. 5) and closing accent (ex. 128, Fig. 28).

<table>
<thead>
<tr>
<th>Interaction of tones and phrasing</th>
</tr>
</thead>
<tbody>
<tr>
<td>tones</td>
</tr>
<tr>
<td>-----------------------------------</td>
</tr>
<tr>
<td>leading accent</td>
</tr>
<tr>
<td>leading-linking accent (spreading)</td>
</tr>
<tr>
<td>leading-linking (default) (L^H)</td>
</tr>
<tr>
<td>low linking tone (L)</td>
</tr>
<tr>
<td>closing accent</td>
</tr>
</tbody>
</table>

Tab. 2: Interaction of the tonal and the metrical components.

Also according to expectations is the low number of breaks after a low link. The 4% (corresponding to 8 cases in absolute numbers) may however be explained away as the result of other factors, such as
hesitation or the early alignment of a closing L tone in an LHL accent with subsequent spreading. This means that there are in fact no cases of a low linking contour forming an independent intonation phrase. Interestingly, a number of TOPICs were also associated with a closing accent. As can be seen in Table 2, closing accents are equally often followed by a break as not (13% vs. 12%). Closing accents on TOPICS will be dealt with in Section 5.7.1 and ignored for the moment.

In the following subsections, I will venture a qualitative prosodic analysis of the various TOPIC constructions identified in (16). On the assumption that SVO is the canonical sentence structure in EA, a preverbal subject is not classified as extra-clausal. It will be shown that there are indeed two constructions involving a preverbal subject: a canonical SP construction and a construction with a left-dislocated subject. The classification will rely on syntactic and prosodic properties of the construction.

5.3.1.1 Subject-predicate constructions

To begin with, let us first look at a canonical SP structure as it is very common in EA discourse. In the following such constructions will be called S-TOPIC constructions. The examples in (19) show (leading-)linking TOPICs whose only constituent is the lexical subject NP which is the only entity topic of the sentence.

(19)

\[
\begin{align*}
\text{a. HRuh_F0_02_03} \\
\text{il-wiliija } hilw-a \\
\text{DEF-woman beautiful-SG.F } \\
\text{H H*+ } !H^L \\
& \text{The woman is beautiful.}' \\
\text{b. Mido_M1_01_01} \\
\text{mi:du } daxal na:m } \\
\text{M enter.3SG.M sleep.3SG.M } \\
\text{H- } !H^L \\
& \text{Mido went to sleep.'} \\
\text{c. Mido_M1_01_02} \\
\text{ik-kubba:ja } kkasar-it } \\
\text{DEF-glass break.INTR-3SG.F } \\
\text{LH } !H^L \\
& \text{The jar broke.'} \\
\text{d. Mido_M1_01_02} \\
\text{il-mughtama}^f xa:jif ji-fakkar^f } \\
\text{DEF-society be.afraid.PTCP.SG.M 3SG.M-think } \\
\text{LH } L^L \\
& \text{The society shuns thinking.'}
\end{align*}
\]
All TOPIC examples under (19) end in a linking terminal and none is accompanied by an intonation break, i.e. each of the SP clauses forms a single IP. In (19a-c) the H tone spreads to the end of the TOPIC, whereas in (19d) there is no spreading. The difference between the upper and lower panels lies in the tonal property of the left boundary: the examples (a) and (b) lack the characteristic lexical rise of the default accent and the TOPIC starts at a higher frequency right away resulting in a high flat contour (H-). This lack of a rise diminishes prominence and results in a prototypical (but prominent) ‘high link’. If we have a closer look at the context in which the contours occur, we find that the examples in (19a) and (19b) exemplify a continuous topic in which the referent of the topic expression was also the primary aboutness topic in the previous sentence. In the case of (19b) il-wiliija 'the woman' refers back to the topic referent Husna, the main female character of the story, while in (19a) the proper name Mido has been used as the topic expression in a number of preceding sentences. The reason why the main character is not referred to by a zero anaphor seems to be that the preceding sentence contains background information while (19a) takes up the main storyline again.

As for the LH- and LH^L contours in (19c,d; Fig. 4c,d), they are associated with lexical NPs whose referents are also given. Nevertheless, the ‘jar’ (il-kubba:ja) in (19b) was last mentioned two clauses before and is only one of a set of possible topic candidates and has to be re-activated. Here, the narrator is describing a situation where a dog had put his head into a glass jug and fallen out of a window - one of the results being the breaking of the jar. It has been frequently noted in the literature...
that continued topics are never made prominent, and the attempts at explaining prominent topics have mostly been in terms of topic activation or contrastiveness. (19c), however, shows that in EA this is not necessarily the case. The prominence-lending rise of the accent on il-mugtamaʕ 'the society' does neither serve to activate an accessible referent (Chafe 1976, 1994), nor to newly establish a referent as the topic (Lambrecht 1994), nor is the topic contrastive in any sense. The prominence of the contour rather is a device to indicate the importance of what is being said - the articulation here is simply more emphatic than in the other examples in (19a,b). Intuitively, the leading part also strengthens the aboutness character of the topic. Note that the society is the discourse topic and the sentence topic of a number of clauses in this whole sequence of a dialogue dealing with the problems the Egyptian society has with emancipated women.

An even clearer example of a prominent topic involving a focus accent on a topic constituent is provided by (20) where il-musˤi:ba 'the disaster' is used as a quasi-synonym of il-mufkila 'the problem', the topic of the preceding clause. A taxi-driver is giving his opinion about the disastrous situation in Egypt today, saying that 'the problem is not the government, the real problem is the stupidity of the people' (il-mufkila mif fi-l-huku:ma, il-musˤi:ba fi-ʕabatˤ in-na:s). Mufkila (‘problem’) and musˤi:ba (‘disaster’) are thus forming some sort of parallel contrastive topics (although the referents are identical). The main reason for the strong prominence associated with the topic expression is the semantic weight of musˤi:ba, or in other terms, the additional information value it has compared to the semantically weak mufkila. The perceptually very clear focus accent manifests itself acoustically by high scaling, a sharp rise and relatively high duration of the whole constituent.

(20) Mar_HF_FG3_02
il-musˤi:ba fi-ʕabatˤ in-na:s
DEF-disaster in-stupidity DEF-people
L^H LH L!HL%
'The disaster is the stupidity of the people.'

Figure 5: Pitch and intensity tracks of an S-TOPIC with a leading-linking contour containing a focus of interest (ex. 20).

The corpus also contains linking TOPICS with a low linking contour as in (21). The example, illustrated in Figure 5, is a marginal case as the downtoned linking topic here is even deaccented, as
evidenced by the shortening of the phonologically long vowel /a:/ (49 msec compared to 145 msec of the same vowel in the same word and same position some clauses later). In most cases, however, the lower prominence of the topic is due to downtoning, i.e. a lower scaling with smaller excursion of the topic accent as compared to the following FOCAL accents as in (22), where the topic Husna is associated with a prominent linking accent (H-), but the following FOCUS is more prominent, rising to a higher pitch level and then falling low. Interestingly, the topic is not even continued in a strict sense. It was last mentioned two clauses earlier, i.e. it might even have been prosodically re-activated here, but in fact it is not.

(21)  
Fagr_FG3_01  
baʃ:ʃba  gih  
daddy come.SG.M  
L- LH  
'Daddy came.'

Fig. 5: Pitch and intensity tracks of deaccented subject topic in TOPIC position (ex. 21).

(22)  
HRuh_F0_02_01  
husna beːdˤ-ə w-hilw-a  
H white-SG.F and-beautiful-SG.F  
H- °L^HL !HL%  
'Husna is fair and beautiful.'

To summarize, a simple SP sentence with a topical subject is mostly associated with a linking or a leading-linking contour. We have seen that the subject and the predicate constituent are rarely separated by a break - especially, if the topic is continuous. Continuing topics in the above examples are associated with a prototypical linking contour, whereas reactivation or contrastiveness of the topic involves more prominence correlating with the rising part of the contour on the accented syllable, and even more prominence is achieved by higher scaling and longer duration. In some contexts we may also find low linking topics which are rarely devoid of proper accents. We will look at such cases below in Section 5.3.1.3.
5.3.1.2 Packaging TOPICS: scrambling, left dislocation and pseudoclefts

The next type of TOPICS I will investigate are packaging TOPICS (P-TOPICS). Maslova & Bernini (2006: 71–73); henceforth M&B define a ‘packaging topic’ as a referring expression about which an assertion is made in a sentence form that paradigmatically differs from another propositionally equivalent sentence. They point to the option of passive constructions and different linear orders exploited by many languages. M&B’s packaging topics comprise instances of word order variation inside the clause as well as left dislocation where a left-detached expression is referred to by an anaphoric expression inside the clause. Both types have in common that the topic referent is within the valency of the verb and thus a proper part of the proposition. I will differentiate between the two types here and treat dislocation as special subtype of packaging. In EA, scrambling, i.e. the first type of word order variation, is severely restricted. We have seen that the topic-comment described in Section 5.3.1.1 is the canonical sentence type in EA. It is the pragmatically neutral word order in the sense of Dryer (1995), being the ‘citation form’ of a sentence as described in Chapter 2.2.6.2, and prosodic variation alone can change the focus structure of the sentence from topic-comment to narrow FOCUS of one of the arguments. It is, however, not possible in EA to front a non-subject argument to initial position as in the Hungarian example in (16c), for instance, or like in German (23a). The EA translation of the German example given in (23b) is ungrammatical, a grammatical translation is given in (23c) showing that EA has to resort to left-dislocation instead. That is, EA word order is more rigid than German or Hungarian word order, which is to be expected because of the lack of case marking in this language.

(23)

a. 
Den Peter habe ich lange nicht gesehen
The P. have I long not seen ‘I have not seen Peter for a long time.’

b. *
but’rus ma-juft-af ba?ːa:-l-i mudda
P NEG-see.1SG-NEG remain.3SG.M-to-1SG while

c. 
but’rus ma-juft-a:-ʃ ba?ːa:-l-i mudda
P NEG-see.1SG-3SG.M-NEG remain.3SG.M-to-1SG while ‘I have not seen Peter for a long time.’

The only cases of a fronted argument in the corpus involve fronting of a PP as in (24). Such cases are rare and mostly belong to higher register, being Wusta\(^{10}\) phenomena. The interlocutors here are two hosts and a physician giving medical advice to the spectators of the popular morning show Sabah el-

---

\(^{10}\) Wusta is a term in Arabic language studies denoting a mixed code between the colloquial language and the written language. It is typical for formal speech types, such as lectures, broadcast interviews and the like.
The language of the speaker is a typical example for Badawi's 'colloquial of the educated'.

Another instance is from a written short story (25).

In (26), an example from a spontaneous colloquial dialogue, the speaker has elaborated about the different scenarios that could arise in a certain situation, all of which are threatening. And in the end she gives the advice that a plan should be found to cope with each of these scenarios.

While scrambling is usually not an option in EA, linear order variation as in Salma-examples in (12) and (13) is extensively used. As noted above, an EA sentence may begin with virtually any basic constituent. This however is only possible when non-subject arguments are in a left-dislocated position and referred to by a resumptive pronoun inside the clause, i.e. if they are extra-clausal. Lambrecht defines a dislocation construction as

a sentence structure in which a referential constituent which could function as an argument or adjunct within a predicate-argument structure occurs instead outside the boundaries of the clause containing the predicate. (Lambrecht 2001: 1050)

Such constructions are extremely frequent in EA and have been claimed to be a typical feature of colloquial speech in other languages, too. In many languages, such constructions do not normally occur in written texts (Lambrecht 1994: 185; M&B 2006: 74), but are frequently found in spontaneous speech. In Arabic in general, such constructions may also be found in the written variety (MSA). Contrary to the classical variety (CA), MSA does not heavily rely on case endings, as native speakers do not generally pronounce the case endings in reading, using the so-called pausal forms instead (case endings are only exceptionally written and have to be added by the speaker when reading MSA). Thus, scrambling is not really an option in MSA, too. The other packaging construction exploited in many languages is the passive construction. Again, this option is very restricted in Arabic varieties as...
passive voice is only used in sentences without an actor.\textsuperscript{11} We are thus left with the dislocation to mark a non-subject topic in sentence-initial position.

Although the subject is by default in the initial topic position, dislocated subject topics occur as well. The best explanation for the function of the left dislocated topic I am aware of is a cognitive one as stated by Lambrecht in the \textit{Principle of the Separation of Reference and Role (PSRR)}. Lambrecht (1994: 184) argues that dislocation "allows speakers to separate the \textit{referring} function of noun phrases from the \textit{relational} role their denotata play as arguments in a proposition" (emphasis in the orig.). The reasons for this are to be found in the processing of a message which is both easier to construct and easier to decode if the processing is done in two subsequent steps. This idea is reminiscent of Aristotle's \textit{categorical judgment} that was further elaborated on by Brentano and Marty in the early 20th century and Kuroda (1972) and Sasse (1987) (cf. Chapter 6 for more details). A categorical judgment is characterized by a \textit{logical bipartiteness}. Linguistically this bipartiteness, i.e. the separation of the procedures of \textit{naming} an entity and \textit{predicating} something on it, is indeed best fulfilled in a left-dislocation construction. Following Lambrecht (1994: 192f.), we may note that the left-dislocated NP does not have argument status itself, but the syntactic role of its denotatum is marked by pronominal topic expressions inside the clause. This distinction nicely draws in with the distinction between TOPICS and topics assumed in the present model. While the pronominal referring expressions are all topics of a clause, only the left dislocated or topicalized (i.e. left-peripheral, but intra-clausal constituents) are also TOPICS. Lambrecht only notes that the dislocated lexical NP carries the reference function. Reference and role are both essentially semantic functions. In EA, these are coded in the morphosyntax of the language, in the form of inflectional morphemes or pronominal suffixes. The \textit{pragmatic} functions of 'aboutness' or 'frame' or 'reference point' (see 5.3.1.3), however, may be associated with the pragmatic TOPIC constituent. Thus the TOPIC may be said to play a \textit{pragmatic role} or stand in a \textit{pragmatic relation} to the whole proposition as argued in Chapter 4.4. In EA, this pragmatic role is frequently associated with a leading contour, although linking contours also frequently occur. By way of demonstration we may now analyse the four sentences given in (12) and (13) above in the following way: the left-dislocated expression \textit{Salma} fulfils the pragmatic TOPIC function in all four sentences, in (12a-c) \textit{Salma} is an argument of the proposition, fulfilling the syntactic role of subject (12a), direct object (12b) and indirect object (12c), the latter two being coded by pronominal suffixed topic expressions -\textit{ha} 'her' and (12c) additionally by the indirect object marker \textit{l-} 'to' resulting in -\textit{lha} 'for her'. In (13), Salma is not an argument of the proposition, but the possessor NP of an annexation which is also coded by co-referential pronominal suffix attached to the head noun \textit{\textasciitilde{u}\textasciitilde{xt}} 'sister'.

\textsuperscript{11} This is definitely the case in EA, MSA has some analytical constructions to indicate an actor in a passive sentence, presumably modeled after constructions in European languages, such as \textit{min qibal} 'on the part of' or \textit{\textasciitilde{u}ala ja\textasciitilde{d}} 'with the hands of' that are only used in the language of the modern press, but much less so in literary texts.
In the following, I will cite some typical examples from the EA text corpus that show the many possibilities of constituent dislocation in EA, subject dislocation, direct object dislocation and indirect object dislocation. Dislocated NPs denoting a subject referent are usually not referred to anaphorically by a personal pronoun, but rather a demonstrative as in (27). In (27), the demonstrative pronoun *di* 'this, SG.F' agrees in gender and number with *ziż:da* 'increase, SG.F’

(27) Knee_MG1_14

*il-hikajt [sic] ziż:dit il-wazn / di ha:ga muhimmaʔawwi*

DEF-story increase DEF-weight DEM.SG.F thing important-SG.F very

'The story of overweight, that's a very important issue.'

All examples shown thus far are classical cases of syntactically marked dislocated TOPICS. There are, however, a number of cases where the subject topic is only intonationally detached by an intonation break, mostly exhibiting leading accent or as a leading-linking TOPIC, as in (28). Here, intonation alone seems to fulfil the highlighting function of the dislocation construction. Such TOPICS are often highly prominent, i.e. the accent associated with them is frequently a focus accent. In any case, they are necessarily separated from the comment by a phrase break. In the present case the focus accent is indicated by the alternative coding with a double peak accent given in parentheticals. Here, the topic is again the highly accessible main character of a frog story, *Mido*, who has to be re-established as a sentence topic after an episode dealing primarily with the dog, the sentence topic of the preceding four clauses.

(28) Mido_F0_01_07

*wi-bašde:n / mi:du / laʔa gizʕ fagara kibi:r [gbi:r]*

and-afterwards M find.3SG.M trunk tree big

'And then, Mido found a big tree trunk.'

A boundary is also frequent in coverb-S word order cases where it is particularly important to differentiate between a FOCAL and a TOPICAL subject (29). This is typical of narrative style in EA, as we have already seen above. If the subject were FOCAL it would necessarily be preceded by the main verb. The prosodic realization of such cases is further evidence for the analysis of such VS sentences as involving a topic.

(29) Mido_F0_01_03

*ha:wil mi:du / ji-busʕiʔ il-ʔawwil fi-l-ʔa:dʕa*

try.3SG.M M 3SG.M-look DEF-first in-DEF-room

'H tried to look around inside the room first.'

---

12 The double-peak is not absolutely clear, but the accent is in any case highly prominent.
Fig. 6: Pitch and intensity tracks of the beginning of a VS sentence with a break after the topical subject (ex. 29).

The following example (30) of a subject topic is noteworthy because the resumptive subject pronoun (this time a phonologically reduced variant hi of the independent personal pronoun hija) is brought under the scope of the negative marker ma...ʃ. It would also have been possible to say il-xuʃ:na, hija mazaditʃ ‘arthrosis, it has not increased’. I am not in a position to identify the pragmatic difference between these variants, but the fact that the subject is incorporated into the FOCUS domain might be taken as further evidence for the PSRR. It seems that while the TOPIC expression is not marked for the semantic and syntactic functions of the denotatum, but expresses the pragmatic function only, the opposite is also true. The topic expression within the clause of necessity only codes the syntactic role of the argument, but not its pragmatic function. Note that the personal pronoun huwwa (3SG.M) only is not part of the main proposition, it only serves to emphasise its validity, meaning something like ‘it is a fact’.

(30) Knee_MG1_03
walˤvahi huwwa l-xuʃ:na ma-hi-zad-itʃ
DM 3SG.M DEF-arthrosis NEG-3SG.F-increase-3SG.F-NEG
'In fact, arthrosis has not increased.'

So far, we have seen only prototypical examples of dislocated TOPICS. (31) is interesting in many respects. It shows a dislocated topic within a complement clause. Topics in embedded clauses seem to be fairly restricted in the languages of the world (cf. Bianchi & Frascarelli 2010 for discussion). They are, however, also possible in English if the contents of the clause represents the main assertion (Hooper & Thompson 1973). This is also the case in the EA example at hand.

(31) Ihsan_FG1_05
jaʔni huwa ka:n ʕaːrif inno ʕabd-in-nasˤir /
FILLER 3SG.M be.3SG.M know.PTCP.SG.M COMP A
L- H LH L\H
'That is, he knew that Abdel Nasser…. 

huwwa da z-Zaʃi:m
3SG.M DEM.SG.M DEF-leader
^LH øH LHL
... he is the LEADER.’

13 What is meant is in fact that the number of cases of arthrosis has not increased in modern times.
The complement clause is the FOCUS of the whole sentence, i.e. the relevant information here is what the referent of *huwa* 'he' at the beginning of the sentence knew, namely that *Abdel Nasser* was the leader. At the same time this clause also consists of a topic-comment clause. *Abdel Nasser*, the discourse topic of the whole passage, is also the topic of the embedded clause, coded by a leading accent. The comment clause contains the independent subject pronoun *huwwa* plus the demonstrative *da* that impose salience on the topic bestowing focus properties on it. This focus is evoking alternatives, indicating that the topic is *Abdel Nasser* and nobody else. The construction imposing salience here is also used for narrow FOCUS, when the rest of the clause is in the background. It would be possible to articulate the utterance with a downtoned accent on *zaʃi:m* 'leader' indicating that this information is taken for granted, thereby invoking a FOCUS interpretation of the subject. But in the present utterance this is not the case. The nominal predicate also carries a focus accent, indicated by the pointed hat (LHL) instead of the utterance-final default downstep. Moreover, intensity is also very high. A final piece of evidence for the FOCUS interpretation of the predicate and not of the subject in the embedded clause is the fact that it is new information whereas *Abdel Nasser* has been the discourse topic of the whole episode. It is true that *Abdel Nasser’s* leadership is common knowledge, but it is not presupposed in the present utterance, and this is what counts.

Continuing topics or topics that have been introduced in the preceding clause as a FOCUS often involve a demonstrative determiner. Example (32b) shows a leading object TOPIC separated by an IP boundary from its comment. The utterance was an answer to the preceding question of the interviewer given in (32a). Topics that are anchored by means of a demonstrative determiner, whether used deictically or anaphorically, are mostly prosodically marked, at least by accenting (mostly with a leading accent) the demonstrative particle to differentiate them from a morphologically identical construction where the demonstrative is used as an anaphoric expression within the comment (but see Section 5.7.2 for more details). While the demonstrative pronoun in (27) is associated with a low linking contour, the demonstrative in (32b) has a leading accent, and crucially, (32b) has a boundary after the demonstrative, while in (27) the IP break is before the demonstrative.

(32)

a. Taameyya_INT_02

\[ wi\text{-}b\text{-}ti\text{-}ʃmil\text{-}i\text{-}ha \quad \text{zza:j if\text{-}fite:ta di} \]

and-IND-2SG-make-F-3SG.F how DEF-Shiteta DEM.SG.F

'And how do you prepare it, *this Shiteta*?'

b. Taameyya_F5_02

\[ \text{if\text{-}fite:ta di / ?abu jasmi:na huwa b\text{-}ji\text{-}ʃmil\text{-}ha...} \]

DEF-Shiteta DEM.SG.F A.Y 3SG.M IND-3SG.M-make-3SG.F

LHL ^H L- L^HL HL- LHL

'This Shiteta, Abu Yasmina prepares it.'

The example in (33b), by contrast, remains slightly ambiguous. There is no boundary after the topic *kilu l\text{-}mo:z* 'one kilo of bananas'. However, there are two prosodic features that indicate that the
demonstrative is more likely to be part of the comment: its lower prominence, and more importantly, its tonal integration into the falling contour of the FOCUS after the leading accent on the TOPIC. Even without a boundary, the deictic use of the demonstrative could be indicated by its integration into the leading contour as in the constructed example in (33c)

\[\text{(33)}\]

\[\begin{align*}
a. & \quad \text{Goha\_F0\_09} \\
& \quad \text{ki:lu l-muzz / da b-ka:m} \\
& \quad \text{kilo DEF-banana DEM.SG.M PREP-how.much} \\
& \quad H- \quad H \quad °!HL \\
& \quad \text{One kilo of bananas, how much is it?}
\end{align*}\]

\[\begin{align*}
b. & \quad \text{Goha\_F0\_10} \\
& \quad \text{ki:lu l-muzz da b-it\'na:far\$ ?\text{irf}} \\
& \quad \text{kilo DEF-banana DEM.SG.M PREP-12 piaster} \\
& \quad H- \quad H \quad L\!HL \quad L\text{*} \\
& \quad \text{One kilo of bananas, that costs 12 piasters.}'
\end{align*}\]

c. \quad \text{constructed example} \\
\quad \text{ki:lu l-muzz da b-it\'na:far\$ ?\text{irf}} \\
\quad \text{kilo DEF-banana DEM.SG.M PREP-12 piaster} \\
\quad H- \quad H \quad L\!HL \quad L\text{*} \\
\quad \text{One kilo of bananas, that costs 12 piasters.}'

Fig. 8: Pitch and intensity tracks of TC-structure with a demonstrative as an anaphoric pronoun introducing the comment (ex. 33b).

We have seen above that prepositional arguments are rarely ever left dislocated. Most of the left dislocated PPs in the EA text corpus are instances of temporal frames (5.3.4.2) whereas spatial adjuncts are in TOPIC position rare (5.3.4.3). PPs as a type of sentence adverbials are canonical in the TOPIC position, although I did not find many examples in the text corpus. Expressions of that kind are fi-ra\d{\text{j}}i (‘in my opinion’) or li-l\text{-}\text{\text{\text{a}}} asaf (‘unfortunately’) and bi-kallo bas\text{\text{\text{a}}}:\text{t}a ‘simply’.

The rare occurrence of prepositional objects as TOPICS in EA seems to be paralleled by French. Lambrecht (2001: 1062) notes that in French PPs are common as antitopics, but they rarely occur left-dislocated. While it is not preferred in EA to use the whole PP as a TOPIC, there is a strategy to express the referent of the PP as a left dislocated TOPIC by preposing only the NP and refer to it with

294
a resumptive pronoun attached to the preposition in situ within the clause (34). I also take this as further evidence for the PSRR which is violated if the whole PP is left-detached.

(34) Ihsan_FG1_04

\[ \text{ig-gawaz ma-hadda-f wa:f\text{?} \text{\$ale: fi-l-\$e:la}} \]

DEF-marriage NEG-someone-NEG agree.PTCP.SG.M on:3SG.M in-DEF-family

'The marriage, no one in the family agreed to it.'

The leading TOPIC is applied here although the marriage has already been the topic of the previous sentences. It is followed by the all-falling FOCUS domain that has no remarkable focus accents (Figure 9). In fact, the final accent on \(\text{\$e:la} \) 'family' is very weak being the least informative part of the FOCUS. It is more informative that nobody agreed to the marriage, which is why both constituents, the negative quantifier and the verb, have high intensity and - being part of the FOCUS - exhibit closing accents with early peak alignment. Nevertheless, the TOPIC is the strongest accent in the IP with the highest peak the strongest pitch movement and the highest intensity. These phonetic correlates are matched by the auditory impression.

![Fig. 9: Pitch and intensity tracks of a left dislocated NP that has been ‘moved out’ of the PP in the main clause (ex. 34).](image)

Another case of an indirect object following a temporal connective where the proper name is coded as a TOPIC whose syntactic role within the clause is marked by the lengthening (and lowering) of the final vowel of the preposition \(\text{\$ala} \) 'on' as a remnant of the lost suffix \(-h\) (3SG.M)\textsuperscript{14}

(35) Mido_01_F0_02

\[ \text{wu-ba\text{\text{\text{-de:n}}} / mi:du kabalas \text{\$ale: n-no:m \?awi}} \]

and-afterwards M press.3SG.M on:3SG.M DEF-sleep very

\(^{4}\text{HL}^{+}H\) \(^{+}H^{2}L H^{+} LH^{2}H^{+}H^{0}H^{L} \)

'And then, Mido became very tired.' (lit.: and then, Mido, pressed on him the sleep very)

Note that in this case there is no break after the TOPIC and also no leading accent, but the topic is sufficiently indicated by the syntactic construction. Likewise, there is no break between the TOPIC NP \(il\-balad \) 'the country' and the comment \(ma-fi-ha:f fulu:s \) 'there is no money in it' in (36).

\textsuperscript{14} In EA, there is morphonological vowel lengthening before suffixation (cf. Watson 2002: 183).
Such constructions are to be considered as highly conventionalized. This is even more true for other prepositional possessive construction, among them figuring most prominently the $and-construction ($and 'at') as the one in (37a). (37a) is a grammaticalized constructions denoting possession equivalent for the English have-construction. Thus the paraphrase in (37b) is clearly odd.

The following example of a D-TOPIC contains a restrictive relative clause (38). The main proposition here is not a statement this time, but a command. The aboutness notion that also pertains to this example has to be understood in the sense of Gundel’s (1985) definition of entity topic that is used by the speaker, among other things, to evoke an action on the side of the addressee with respect to that entity.

The final example (39) is a complex phrase involving an additional embedded topic referent. Both topics, the primary one and the embedded one, are realized as TOPICS. Let us start with the D-TOPIC, which is the embedded one: il-hurrija ‘freedom’. As in (31) above, the embedded clause here is again a complement clause with a left dislocated topic, associated with a leading-linking accent.15

15 Note that the illustration in Figure 7b does not exactly mimic perception. As in many other examples, the spreading on the H tone cannot easily be detected on visual inspection of the pitch track because of the intrinsic higher frequency of the high vowel /i/ compared to the lower /a/ as well as the frequency drop due to the diminishing tension of the vocal folds after the
The higher topic here is il-masal 'the proverb' articulated in the TOPIC phrase 'the proverb I told you before'. The proverb, however, is not part of the proposition. Such topics will be classified as free (hanging) TOPICS and will be discussed in Section 5.3.1.4. The break after the TOPIC phrase may be called into question, but nothing hinges on this analysis. What is important is the general impression of a leading contour. (Figure 7a).

Another other type of relative clause frequently found in EA is the pseudo-cleft which refers to an entity that is an argument of the proposition (40). Here the TOPIC contour extends over the whole relative clause followed by the FOCUS identifying the referent of 'his friend' as Abdel Moneim Amin.

Pseudo-clefs are not restricted to subject function, but may also denote other arguments as in (41). The referent of the TOPICAL relative clause is referred to by the pronominal object suffix -u 'him' in the comment clause. Here the pseudo-cleft is not followed by an identificational FOCUS, but by the accented syllable. Such effects distort the acoustic data. This can be avoided in experimental work by carefully choosing the segmental material, but has to be taken into account when spontaneous speech data are analyzed.

As usual, I annotated a break here due to the perceptual impression that a pause could follow without distorting the auditory impression.
verbally expressed proposition *ji*-si:b-u 'he lets him go' which itself is a complete clause. The syntactic complexity of the sentence is due to the fact that the referent of the relative clause is itself an object within this clause, expressed by the object suffix on *jila:*ʔi ‘he finds’, which in turn refers back to the relative marker *illi*. The utterance is produced jointly by two sisters in a very vivid dialogic narrative about the father coming home in the middle of the night finding his wife and three children lying unconscious on the floor.

(41)  
\[
\begin{array}{llllll}
& \text{Fagr}_F0/3_03 & wu-li & ji-laʔi: & mifawwa? & fwajja & / (F3) & ji-si:b-u \\
\text{and-} & 3SG.M- & \text{awake} & a.little & 3SG.M- & \text{leave-} & 3SG.M \\
\text{REL} & \text{find:3SG.M} & & & & & \\
\text{L-} & H & L^\text{H} & LLHL- & !HL & & \\
\end{array}
\]

'When he finds one of us awake to some extent, he lets him go.' (lit. the one he finds him awake a little, he lets him go')

Finally, pseudo-clefts of the type ṭilli hasˤal ṭinn.. 'What happened (is) that...' are numerous as devices of establishing coherence, preceding FOCAL information.

As the examples in this section suggest, dislocated TOPICS are more readily articulated in a separate IP and associated with a leading contour than TOPICS in a canonical sentence structure. Again, this is by no means obligatory. This is especially true for TOPICS involving a relative clause, whether it is a restrictive relative clause modifying the topical noun or a pseudocleft. Contrary to that, the grammaticalized and highly conventionalized constructions in (37) and (36), respectively, are much less likely to be associated with a leading contour and followed by a break – they are rather typically leading-linking and phrased in one IP together with the comment.

### 5.3.1.3 Reference point constructions and multiple topics

I have argued above that the functional motivation for dislocation is the PSRR. In this section I will introduce the reference point construction that fulfils another important function of topics, i.e. coherence and anchoring of an entity in discourse. The notion of reference point, as conceptualized by Langacker (1993), is a conceptual archetype that is invoked in a whole range of linguistic phenomena of which topic constructions are just one. According to Langacker (1993: 5) cognitive reference point pertains to "the ability to invoke the conception of one entity point for purposes of establishing mental contact with another", thereby functioning as the “abstract basis for possessives” in general. An English reference point TOPIC (henceforth R-TOPIC) was cited in (16g): *As for the homeless, where did their food come from?*

Formally, an R-TOPIC is similar to a D-TOPIC exhibiting an anaphoric pronominal element referring back to the referent of the R-TOPIC. However, this referent is not an argument of the proposition, but
stands in a possessive relationship to one of the arguments. The construction is extremely common in
EA, and we have already seen an example in (13) where the referent of the R-TOPIC was Salma and
the subject of the proposition related to her was her sister. In (42), a woman from the country-side
visiting her niece in the capital tells the latest news from the village. Here she is talking about her
sister (the niece’s other aunt) who has gone blind.

(42) HRuh_F0_09_01
Sammit-ik s’afijja xal’u:a:s’ naz’ar-ha r’a:h
aunt-2SG.F S finished view-3SG.F go.3SG.M
‘Your aunt Safiyya, her eyesight is already gone.’

Unsurprisingly, the interlocutors themselves are preferred reference points, which is why first and
second person pronouns may frequently be found in that position. In (43) the speaker talks about her
software by referring to herself with the pronoun ana T which in reference point function is typically
deaccented.

(43) Mar_Shab_FG3_19
w-ana ?as’ilan is-safiwe:r bita:ʕ-i safiwe:r bi-j-kisif (!)17
and-1SG originally DEF-software POSS-1SG. software IND-3SG.M-be-ashamed
L- ^H LHI- H- H- L!HL%
‘And my software is embarrassing anyway.’ (lit.: ‘and I anyway the software of mine (is a)
software that embarrasses’)

Multiple topics are also common where each topic is the reference point for the next one.
Grammatically, the topic chain may be continued ad infinitum, but cognitive constraints pose a natural
limit to the actual occurrence of such constructions. Nevertheless, a sentence like (44b) is in principle
possible and a grammatically well-formed alternative to the canonical sentence in (44a).

(44)
a. constructed example
mara:t ?axu:-ja mudarris-a
wife brother-1SG teacher-SG.F
‘My brother’s wife is a teacher.’

b. constructed example
ana ?axu:-ja mara:i-t-u mudarris-a
1SG brother-1SG wife-3SG.M teacher-SG.F
‘My brother’s wife is a teacher.’
(lit.: ‘I, brother-my, wife-his (is a) teacher’)

In such topic chains all topics except the first one are anchored with a resumptive pronoun in the
preceding one. Although such constructions intuitively seem to be a typical trait of colloquial speech,
they also occur in Classical Arabic (Lewkowicz 1971). Thus, R-TOPICS do not only differ from D-
TOPICS semantically by not being an argument of the main proposition. The also differ with respect
to recursiveness. While, as I claim, the number of possible R-TOPICS is not constrained

17 This is obviously a mistake, it should be bijiksif ‘make ashamed’.

299
grammatically, the number of D-TOPICS seems to be limited to one (perhaps two), as illustrated in (10) above. But this issue needs further investigation.

Pronominal topics are usually accentless and not separated by a boundary from the comment part of the sentence. Multiple topic constructions with full lexical NPs, by contrast, frequently exhibit a boundary between the individual topics as in (45).

(45) Mido_F0_01_04
w-ik-kalbə btaːs-u / dimaːɣ-u mahʃiːr-a guwwa l-ʔaːz-a
and-DEF-dog POSS-3SG.M head-3SG.M stuck-SG.F inside DEF-glass-IN
LLHL H-
'And his dog, his head is stuck inside the jug.' (lit.: 'and the dog of his, head-his ...')

Multiple TOPIC constructions may also involve other types of TOPICS, especially frames and D-TOPICS as in (46) and (47).

(46) Mar_Shaw_FG3_21
faːʔawwil haːga / il- il-overhead-projector dawwat / fiː: silk
so-first thing DEF DEF-overhead.projector DEM.SG.M in:3SG.M wire
LH°L H- L- LH°H L°H’ H- L°H’
'So first, this overhead projector has a wire.' (lit. 'so first thing, the overhead projector this, in it (is a ) wire."

(47) Knee_FM1_20
waːfi aː- bidaːjít xaːlɪsˤ kalam-na / ðul-na isːsinn
and-in HESIT beginning totally speech-1PL say-1PL DEF-age
'And in the very beginning of our talk we said (that)
isˤ-sˤuːʔajjar mumkin innə huwwa jaːtaːradˤ li...
DEF-little/young possible COMP 3SG.M 3SG.M-be.exposed to/for
... young people can be exposed to ... (lit. we said the young age, possible that he is exposed to...)

Prosodically R-TOPICS are typically associated with a (leading-)linking contour as in (45) or with a downtoned linking contour if they are pronouns (43). In multiple TOPIC constructions, every TOPIC is mostly associated with its own leading or leading-linking contour (46, 47).

5.3.1.4 Free TOPICS

The D-TOPICS and R-TOPICS of the present account may be subsumed under the category of 'hanging'. M&B’s definition of hanging topics, however, is more narrow, for a sentence has a hanging topic if the “semantic (‘real world’) relation between the topic and the state of affairs described by the comment is not linguistically encoded” (2006: 73). To avoid confusion I will refer to this type of TOPIC constructions as free TOPICS.

Hanging or free TOPICS were called Chinese-style topics by Chafe (1976), is what spurred Li &Thompson’s (1976) typological classification of topic-prominent vs. subject-prominent languages. It
is well-known that in some languages, such as Chinese and Japanese, this type of topic-comment construction is the unmarked case as shown in (16i) above. Brustad’s (2000) classification of Spoken Arabic as topic-prominent rests on the assumption that Arabic dislocated TOPICS, as discussed in the above sections, are the Semitic equivalent of Chinese-style topics in Mandarin or Japanese. However, the difference between the two constructions is not only syntactic, but also functional. The pragmatic difference has also been pointed out by Chafe (1976: 50) who noted that Chinese-style topics are not characterized as “what the sentence is about”, which holds for what he calls subjects (cf. 5.3.1.1). According to Chafe (ibid.) such topics rather “limit the applicability of the main predication to a certain restricted domain.” An English example was given in (10b): As for education, John prefers Bertrand Russel’s ideas. In that sense, hanging entity topics are pragmatically closer to frames that will be dealt with in 5.1.4.

In the EA corpus free TOPICS are not very frequent. Equivalent examples of the English sentence in (10b) will be discussed in Section 5.5 under the heading of morpho-lexical topic marking. In this section we will only be concerned with unlinked NPs and unlinked pseudo-clefts with no additional topic-marking lexical items. (48a) is from a written short story, artfully read by Heidi Abdel Shahid. A new discourse topic is repeated from a preceding question, which is again repeated as part of the FOCUS within the predication. Note that it is not referred to by a pronominal topic, the whole NP is rather repeated in its in-situ position within the clause. Thus, this example is not a D-TOPIC, but rather a type of free TOPIC which only incidentally shares a referent with the main proposition. The same construction would be used in the similar sentence in (48b), where it is obvious that the cinema, or ‘going to the cinema’ as a type of leisure time activity is the frame for the predication. However, the whole intonation would be different, presumably with a focus accent on Miami, being the only new concept in the whole proposition. Note that the second occurrence of cinema in (48a) at the end of the sentence is downtoned.

(48)

a. Taxi_08_F0_01
   is-sinima / ja::: / ?ana baʔaː·l-i miljuː:n sana
   DEF-cinema INTERJ 1SG remain.3SG.M-to-1SG one.million year
   ‘Talking about cinemas, phuh, I have not been to the cinema for ages.’
   ma-raḥtaʃ sinima
   NEG-go.3SG.M-NEG cinema
   L’ L-

b. constructed example
   is-sinima / ja::: / ?ana baʔaː·l-i miljuː:n sana
   DEF-cinema INTERJ 1SG remain.3SG.M-to-1SG one.million year
   Talking about cinemas, phuh, I have not been to cinema Miami for ages.
   ma-raḥtaʃ sinima mijaːmi
   NEG-go.3SG.M-NEG cinema Miami
The framing function of the free TOPIC is better illustrated by the next example (49), an adjunct that is asyndetically preposed to the main proposition. Functionally this is an example for a 'conditional' frame (IF-TOPIC) as discussed in Section 5.3.4.4, if we accept that the semantics of the sentence is something like if the disease has reached an advanced stage.

(49) Knee_MG1_07

And in some late cases mobility is restricted' (lit.: 'And some late cases, there are restrictions)

In the final example (50) there is no relationship at all between the topic and the proposition. The speaker is commenting on something one of her interlocutors had mentioned by giving an example. She uses the topic to come back to what was said above to rectify some false assumptions on his side.

(50) Mar_Siqa_FG3_11

But the point that you are talking about here...

Thus the free TOPIC has scope over the whole paragraph. It establishes the link between the previously discussed issue of how a person who is completely devastated can recover, being referred to as in-nuʔtˤa 'the point' and the following metaphoric example, thereby also fulfilling an anchoring function we have identified for R-TOPICS. Formally the TOPIC contains two entity topics - 'the point' and 'you'. The second clause starts with the a topic shift to ana 'I', a classical D-TOPIC, the second potential entity topic expressed by the definite NP 'the train' is clearly not a topic. It is only the subject of an idiom that means that a person is dashed to the ground after having been 'rolled over by the train back and forth (i.e. twice). It is clear that the train is no matter of concern, the only topic here is 'me'. Note that in this case the word order of the clause is VS to diminish the likelihood of a topic-comment construal of the sentence.

The prosodic picture that emerges from the investigation of free or unlinked TOPICS shows a high preference for intonation boundaries and leading contours at the end of the TOPIC constituent. In fact, all examples exhibited this prosody. This is in line with the lower cohesion between this type of
TOPIC and the main proposition. One factor that induces a boundary is doubtless the propositional nature of clauses as in the final example (50) for instance.

5.3.1.5 Pronominal TOPICS

Above, we have seen that pronouns may be coded as TOPICS in EA. It is clearly not the case that independent subject pronouns only occur when they are contrastive as suggested by Jelinek (2002). In this section, I will give a short overview of pronominals in TOPIC position and discuss an additional function of pronouns that is related to reference point TOPICS, but differs from these both formally and pragmatically.

But let us look at independent pronouns and their use as S-TOPICS or D-TOPICS first. In Section 5.3.1.3 we have already seen that pronouns are frequently deaccented. (51) is the first sentence of a recipe and a typical example of an accentless pronominal subject in preverbal, but not in left-dislocated position. As can be seen in Figure 10, the pronoun is prosodically proclitic to the verb hanišmil 'we will make'. Another interesting aspect that can be seen in this example is the division of the whole utterance into a TOPIC and a FOCUS. Note that the FOCUS here is narrow, identifying the dish that is going to be prepared: bamja bi-l-lahma ('Bamia (ladyfingers) with meat'). Given that the interviewer had asked for a recipe of the speaker's choice, it is clear that the presupposition of the sentence is 'we will make X today'. This is linguistically reflected in the word order, the time adverbial innahar'da 'today' is in initial position and the new FOCAL information is in final position. The second indication for the TOPIC-FOCUS partition comes from prosody. The reader may have noticed that in the examples above the turning point between the leading/linking tune and the closing tune was always at the border between TOPIC and FOCUS. In (51), this is also the case with the exception of the first FOCAL element Bamia. The reason is that the name of the dish is treated here as a 'prosodic compound' by means of the Integration I contour (cf. 4.5.5). Nevertheless, the first part is made prominent by a focus accent, further enhanced by the following break. Thus, I take the NP + PP constituent to be the narrow FOCUS of this sentence and the part preceding it is the TOPIC. This suggestion will be discussed in detail in a separate subsection (5.3.2).

(51) Bamia_F3_01

\[
\begin{array}{cccccc}
\text{ihna} & \text{ha-ni-šmil} & \text{innahar\textquotesingle}da & \text{bamja} & / & \text{bi-l-lahma} \\
1\text{PL} & \text{FUT-1PL-make} & \text{today} & \text{Bamia} & \text{PREP-DEF-meat} \\
L & LH & L^{\text{\text{\text{\text{\text{H}}}}} - ^{\text{\text{\text{\text{\text{L}}}^\text{\text{\text{\text{\text{H}}}}}^\text{\text{\text{\text{\text{L}}}}}}} & ^\text{\text{\text{\text{\text{L}}}^\text{\text{\text{\text{\text{H}}}}}^\text{\text{\text{\text{\text{L}}}}}}} & H^\text{\text{\text{\text{\text{L}}}^\text{\text{\text{\text{\text{H}}}}}^\text{\text{\text{\text{\text{L}}}}}}} \\
\end{array}
\]

'Today we will make Bamia with meat.'
The subject topic is, however, not accentless when it marks a topic shift and/or is contrastive as in the following example of parallel contrast, which is readily identified as three subsequent TC-structures. By the way, the use of the pronoun here is obligatory because the predicate is nominal, i.e. there is no finite verb indicating agreement with the subject. The somewhat peculiar contents of these clauses calls for an explanation: Here Marwa Rakha who we are already familiar with by now is again a guest in one of the talk shows and plays a little game with the three hosts (a young man and two young women) to illustrate a point she wishes to make. So each of the hosts has to imagine himself/herself being one of the named things (carrot, egg, nescafé). These metonymically used nouns constitute three parallel contrastive FOCI. Each of the bipartite SP clauses has a leading subject pronoun and a FOCUS made prominent by a focus accent. The first two of the FOCI are high-rising to indicate importance and continuation and the third one is low-falling to indicate assertion and finality (see also Section 4.5.2 for details on the accents).

(52) Mar_Siqa_FG3_15

\[\text{inta} \ quad \text{gazārˤa} / \quad \text{wi-ʔinti} \quad \text{bidˤ-a:ja} / \quad \text{wi-ʔinti} \quad \text{neskafe:}\]

2SG.M  carrot-IN and-2SG.F  egg-IN and-2SG.F  nescafé
LH    L*H^>   LH    L*^H>   LH    !H*L%

'You are a carrot, and you are an egg, and you are a Nescafé.'

As we can see in Figure 11, the TOPICAL subject pronouns are all rising with leading accents.

Fig. 10: Pitch and intensity tracks of an utterance with a pronominal subject inside the clause.

Fig. 11: Pitch and intensity tracks of an utterance with a pronominal subject inside the clause.
However, independent pronouns as TOPICS can also have another function. In this case they are not necessarily arguments of the proposition they precede, as has already been pointed out in 5.3.1.3 in the discussion of R-TOPICS. The following example is ambiguous between two readings. The first reading is that *huwa* 'he' is used as an anaphor referring back to the main discourse topic the late Egyptian writer *Ihsan Abdel Quddous*. The interviewee is telling the hosts of the talk show that *Abdel Quddous* had entrusted all his personal documents with her, his assistant, because she was supposed to write down his memoirs. In the preceding six clauses, the referent of *huwa* has not been the primary topic. It is therefore quite probable that the speaker should choose to re-activate him as a topic. In this interpretation, the TOPIC domain contains two entity topics *Ihsan Abdel Quddous* and the records documenting his life (*kullo ha:ga fi-haja:t-u* 'everything in his life'). This is most probably the preferred reading.

(53) Ihsan_FG1_07

*huwa*  
3SG.M all/every thing  
*kullo*  
3SG.M in-life  
*ha:ga*  
3SG.M be-3SG.F  
*fi-haja:t-u*  
3SG.M be-1SG with-1SG  
*ka-t*  
3SG.M  
*m\:\:a:-ja*  
L\:-LHL  
H^\#:L  
L\:-LH  
H  
L\::HL%  

'He, everything in his life ... was with me.'

In the second, less preferred reading, *huwa* is not an anaphor. It could also be interpreted as non-referential used in a purely discoursal function, indicating the factuality of the following information.

In the following example, the pronoun is used in exactly this function. Here *in-na:s* 'the people' is one of the topics, coded as an R-TOPIC, the second potential topic is *awareness*, which may be analysed as a topic because of the leading accent associated with it,\(^{18}\) and the third topic is the issue of arthrosis (*hika:jit il-xu\u00fa:na*). There is no potential entity to which *huwa* could be referring, thus *huwa* cannot be a topic expression.

(54) Knee_MG1_03

*huwa*  
3SG.M DEF people  
*in-na:s*  
3SG.M DEF begin  
*ibtada*  
3SG.M become  
*ji-b?a*  
3SG.M more  
*mu:drik-ha*  
3SG.M be-3SG.F awareness  
*ʔaktar*  
3SG.M be-1SG more  
LH  
L^\#:H  
L\:-LH  
L\:-LH  
H  
L\::HL%  

'As it stands, people's awareness of arthrosis has increased.'

*li-hkajt*  
DEF for-story  
*il-xu\u00fa:na*  
DEF arthritis

I have chosen to translate *huwa* with the English expression 'as it stands' to indicate reference to how things are according to common knowledge. What the choice of the third person pronoun in such situations seems to indicate is that this is not (only) the opinion of the speaker, nor is it supposed to be related to the hearer, but it rather is a commonly accepted fact or at least a fact suggested by others, not the interlocutors. Such R-TOPICS thus seem to have an evidential flavour to them. A further example is (30) above, where the pronoun *huwa* also does not agree with any referential expression in

---

\(^{18}\) Note that there are also indications that suggest an analysis of 'awareness' as belonging to the FOCUS. These are word order and the relative 'newness' of the information.
the context. In that case, it also seems to fulfil the same 'evidential' function. If this analysis is correct, we may expect that first and second person pronouns also occur in this function. And this is indeed the case. As speakers mostly speak of themselves and, as most of what they say is, of course, related to their own view of the world, first person pronouns are also frequently occurring in this function. Owens et al. (2010: 31), in their quantitative study of overt and null subjects in three peninsular Arabic varieties note that "[e]pisodes are sometimes delimited by an overt 1.SG pronoun, often with a characteristic verb such as *ana ataḍakkar* 'I remember.' This observation seems to point in the direction of the here proposed function. Utz Maas (p.c.) informs me that Moroccan Arabic also frequently preposes 1SG pronouns before an utterance, which cannot be regarded as ordinary (aboutness) topics.

The following passage from the same talk show nicely illustrates the hedging effect produced by the repeated use of the first person plural pronoun *ihna* 'we' preposed before every statement the doctor is giving about his and his fellow-physicians' current way of thinking about the therapeutic measures that are preferred in most cases of knee arthrosis.

(55) Knee_MG1_12

```
ihna  ig-gira:ha  da  ?a:xir  malgaʔ  li:-na  /
 1PL  DEF-operation  DEM.SG.M  last  refuge  to-1PL
'As we see it, operation, that's the last resort for us.'

ihna  'abʕan  il-wiqa:ja  xajr-un  min-il-filaːg  /
 1PL  naturally  DEF-prevention  good-NOM.INDF  from-DEF-therapy
Of course, prevention is better than therapy.

ihna  /  bi-n-ʔuːl  /  illi  bi-j-samm-uː-ha  ir-riskfakturs  ...
 1PL  IND-1PL-say  REL  IND-3-name-PL-3SG.M  DEF-risk.factors
We think that what they call risk factors...

humma  talat  hag-aːt
3PL  three  thing-PL.F
they are three things.'
```

Contrary to the pronominal subject topics all three instances are prosodically prominent and associated with a leading accent. Particularly the third and final instance is high rising and followed by a break which mitigates against the construal as an aboutness subject topic of an informative proposition 'we say', rather the whole phrase *ihna bi-n-ʔuːl* is functionally discoursal.

The example of a first person singular pronoun in (56) illustrates the relationship between this type of R-TOPIC and the reference point definition given in Section 5.3.1.3 above as a type of possessive construction. But note that we have seen that pronominal R-TOPICs are mostly deaccented. In this example, the speaker seems to assert the fact that she is talking about herself. Her use of the R-TOPIC with a closing contour is clearly emphatic and seems to be indicating her own importance in the story she is telling.
Second person pronouns are used to pose or answer a question. Here, the use of the polite form is particularly frequent with less familiar, older or socially superior interlocutors. The following two examples from the corpus are both uttered by a magician who is a guest in a talk show answering questions posed by the male (57) and female (59) hosts. Both examples clearly show that the polite forms hadˤritak (M) and hadˤritik (F) have no argument status in the proposition. In (57), the aboutness topic is il-lišba di 'this game'. The speaker is demonstrating one of his tricks and showing and explaining the objects he is using.

(57) Sahr_MG2_12

hadˤrit-ak baʔa il-lišba di Ŧiba:ra Ŧan POL-2SG.M SEQ DEF-game DEM.SG.F expression PREP

'Now look, this game is essentially ...

šas'a:jä fi:-ha Ŧarbaš kuwar stick in-3SG.F 4 ball.PL

a stick with four balls attached to it.'

(58) Sahr_MG2_02

walˤla:hi hadˤrit-ik min zama:n DM HON-2SG.M from time

'A long time ago.'

Example (58) is the answer to a question about the time the speaker's interest in magic had begun. The answer to the question is min zama:n 'since a long time, a long time ago'. Thus, both instances of the referring expressions cannot be classified as aboutness topics, but rather fulfil a phatic function. The phatic function may also be discerned in instances of preposed pronominals where the referent is an argument of the proposition. (59) is from an Egyptian film cited in Brustad (2000) and shows the typical combination of the familiar form of the independent pronoun together with the polite form.

(59) (from Brustad 2000: 391)

inta hadˤrit-ak ti-ʕraf-ni 2SG.M POL-2SG.M 2SG.M-know-1SG

'Do you know me, sir? (lit. you, your grace, you know me?)

Example (60) is from a recipe where the addressee was instructed how to make Taameyya (better known as Falafil).

(60) Taameyya_F5_01

inti ha[dˤ]rit-ik law Ŧand-ik baʔa kabba 2SG.F POL-2SG.F if at-2SG.F SEQ masher

'If you then have a masher...'

307
The following passage from a very lively TV discussion contains a number of pronominal hedging expressions in TOPIC position that all express that the speaker is only giving her own point of view about the situation under discussion. The passage nicely illustrates the different accents associated with the pronouns: the first one is closing expressing assertion; the second one leading-linking for continuation, whereas the third pronoun is again accentless being no carrier of an attitude itself - an asserting closing accent is associated with taqdiːrī 'my estimation' with ana ‘I’ following it as an antitopic.

(61)  Mehwar FG2_01_03_02
kida / liʔannu ʔhna / jaʃni ʔana fi raʃj-i tˤahʕan /
so because 1PL FILLER 1SG in opinion-1SG naturally
°HL% LH°L °HL% L- LH- °HL- L*H
That's why! Because we ... of course in my opinion...

taqdiːr-i ʔana / wighet nazʕar-i /
estimate-1SG 1SG side view-1SG
LHL%- LH- °HH this is my estimate .... my viewpoint is...

Summarizing the use of independent pronouns as TOPICs, we may state that they are used in different ways. First, they frequently have the function of continuous topic, which is to some extent unexpected, as EA is a null-subject language. In this function, their prosody is usually linking with low prominence. When the topic is shifted or even in explicit contrast to other topics they are prominent, and associated with a leading(-linking) contour. And third, they may serve as a reference point as discussed in 5.3.1.3 above. In all of these functions, they are usually not separated by an IP boundary from the subsequent clause. A fourth type of pronominal TOPICs, which could be regarded as a subtype of R-TOPIC, also has the function to relate the main proposition to the 'referent' of the TOPIC expression. However, this referent is frequently nonspecific. In the case of first and second person pronouns, the referents may be identified. But when the pronoun is third person singular, it does not refer to a specific person, but rather to a nonspecific general other, indicating that what is being said is either common knowledge or at least not only or explicitly the opinion of the speaker. Thus, this type of pronominal TOPICS seems to have an evidential flavour to it and may be used as a hedging device. If the pronoun serves the fourth function, there is usually no explicit grammatical relation between the pronoun and a constituent of the clause, IP boundaries are frequent and the tonal contour is usually closing.

5.3.2  Relational presupposition as TOPIC

As we have seen so far, the distinction between topics and TOPICS proves rather useful for distinguishing between different types of left-peripheral TOPIC constructions and the notion of an
entity aboutness topic that can occur freely in any position or is not expressed overtly at all. In this
section, I will present additional evidence that makes the notion of TOPIC necessary to account for a
consistent prosodic phenomenon in EA utterances that cannot be explained if we stick to the narrow
concept of aboutness topic that necessarily pertains to an entity.

Let us start with the results of the experiment reported in 4.4.1.1. There we saw that in response to the
question 'Where is Nagwa?' the utterance 'Nagwa has gone to X' will mostly show a leading or
leading-linking contour to the end of the verb with an optional break at the end of the contour, as if the
speaker wants to say something like 'Nagwa has gone to: X'. As already pointed out in 4.4.1.1, this
makes sense because the verb can be assumed to be coding a presupposed action, Nagwa's not being
here implies her having gone somewhere. This leaves us with a narrow FOCUS domain consisting of
the variable X and nothing else. An example is given in (62), illustrated in Figure 12.

(62) NTOPCVCF01
    nagwa ra:hit malga? il-musinn-i:n
N go-3SG.F refuge DEF-aged-PL.M
LH- oL^H LH L!H L%

'Nagwa has gone to the old people's home.'

Fig. 12: Pitch and intensity tracks of a TOPIC phrase including the verb (ex. )

The leading contour here is associated with the sequence nagwa ra:hit 'Nagwa has gone' and the
turning point of the contour is at the end of the verb, the following locative object is associated with a
closing contour. Steedman (1991: 60) notes that in an answer to a contrastive question as illustrated in
(63), the "intonation contour, which conveys the contrast between the previous topic concerning Alice
and the new one concerning Mary, imposes the perceptual grouping indicated by the brackets." (emphasis added).

(63) (= Steedman's ex. 1)
  a) I know that Alice likes velvet. But what does MARY prefer?
  b) (MARY prefers) (CORDUROY).
      [L+H* L H% H* LL%]

I have added to the example a phonological notation according to Steedman's model. Steedman argues
that the two intonational constituents are characterized by a "theme tune" and a "rhem tune"
(Steedman 1991, 2000) and explicitly notes that the above intonation contour divides the sentence into an "open proposition and its complement" (Steedman 1991: 261). Apart from differences in details of the phonological analysis, Steedman's proposal is identical to the one advocated here. Contrary to many other formal models developed to describe the interface between information structure and intonation, Steedman recognizes the importance of tones and tunes and their relation to information structure and potential divergence from surface syntactic structure, as the following quote illustrates:

> These two tunes, L+H* LH% and H* LL%, thus distinguish two informational units within the sentence. Their functional roles of defining 'what the utterance is about', versus 'what the speaker says about it', have been discussed under a bewildering variety of nomenclatures, most of them unformalized. (Steedman 1991: 275)

Translating these observation into the present model means that the open proposition coincides with the functionally and formally identified TOPIC of the proposition, which is tantamount to saying that TOPIC is the counterpart of FOCUS. This view is not commonly held (cf. e.g. Lambrecht 1994 for a different view), but I believe that the proposal made by Gundel (1976, 1999) is essentially the same. Gundel (1976: 30) notes that the different partitioning of a sentence into presupposition-focus as suggested by Chomsky (1970) and Jackendoff (1972) "can be accounted for directly in terms of topic-comment structure, i.e., in terms of the distinction between what the sentence ... is about ..." (emphasis added). It is commonly accepted that for an entity to be a topic, it must carry an existential presupposition. If an entity is expressed by a pronoun or a definite NP, for instance, its existence is necessarily presupposed. What Gundel suggests, is that in a simple predicate FOCUS sentence, the subject is existentially presupposed within a particular situation (time and place). Such a subject-predicate sentence may, for instance be preceded by time and place adverbials as in example (64) (= Gundel's example (102); Gundel 1976: 40).

(64)  As for (Washington) yesterday, Archie rejected the proposal.

As noted by Gundel, the presupposition is that something happened yesterday in Washington, which is the event of Archie rejecting the proposal. Given a question such as **What can you tell me about the Watergate situation, the news from Washington?**, I assume that, in principle, there are two possible pragmatic presuppositions: (1) only the situation in Washington yesterday is pragmatically presupposed or (2) the pragmatic presupposition including Archie's presence. In this case, our world knowledge tells us that Archie was part of the situation. Anyhow, if the answer in a similar question would be **One person rejected the proposal**, only the reading under (1) would be possible. However, the proper noun Archie suggests that **Archie** carries an existential presupposition and thus reading (2) is more likely. Being more likely, however, does not mean that the reading is determined by these

---

19 This not to say that Steedman's model makes no reference to syntax. On the contrary, what he tries to develop is a model within the framework of Combinatory Grammars that makes reference to a "unified notion of structure" (1991: 263) which is nevertheless different from syntactic surface structure, but is claimed to account for intonational structure and syntactic structure at the same time. We will not be concerned with this aspect of Steedman's model here.

20 Presumably Archibald Cox, the special prosecutor in the Watergate scandal, who rejected a proposal by President Nixon.
facts. Gundel does not differentiate between existential presupposition and pragmatic presupposition. In the model presented here the difference is exactly the speaker's choice. While existential presuppositions only relate to individual concepts, a pragmatic presupposition relates to information. It is crucial that such pragmatic presuppositions may vary according to the speaker's choice. That is, existential presuppositions and discourse factors such as the cognitive status of individual referents do not uniquely determine a pragmatic presupposition. Rather, this is left to the speaker's choice, i.e. whether he wants the predication to be about a certain referent (in this case Archie) or only about the situation, i.e. the time and place of the event. To stress the relational character of pragmatic presuppositions and their function as the TOPIC of a sentence, I call them relational presuppositions. On this view, TOPIC and FOCUS are complementary categories related to each other, which is contrary to Lambrecht's suggestion that topic and focus (= our FOCUS) are necessarily co-extensive with syntactic constituents. For Lambrecht there are only three types of focus structures: Sentence Focus, Predicate Focus, and Argument Focus (Lambrecht 1994: 221ff.). As Lambrecht (ibid.) argues, this classification is based on the formal properties that distinguish semantically identical "allosentences". Such a formal property is suggested to be the focus accent that in the unmarked case of a Predicate Focus necessarily occurs somewhere in the focus (FOCUS) domain (p. 304). As we have repeatedly pointed out, such an assumption does not make sense for EA. In EA, however, we have the formal property of contour that may distinguish otherwise identical sentences ("allosentences"), varying with the specific contexts.

Returning now to the example in (64): In EA, as I claim, the different readings of the sentences would be expressed differently. Under the second reading, Archie would be an ordinary topic and the proper noun would be associated with a leading(-linking) contour. Under the first reading, the proper noun would be associated with a closing accent, inviting the addressee to construe the subject as belonging to the new information. In that case, the only relational presupposition would be here is that there was a meeting yesterday in Washington concerning the Watergate Affair.

The following example (65b), again from one of the talk shows with Marwa Rakha, is a good illustration for a relational presupposition TOPIC and its prosodic realization. In this example, the context explicitly expressed here, as the preceding sentence given in (65a) already tells us that the only relationally new information we are expecting is where the problems come from.

(65) Mar_Shab_FG3_02

a. pretext
wu-$\lambda$mâ$a$:q im-mu$\lambda$ami:$q$ il-mâ$a$:kil kull-aha lli
and-depth.PL DEF-depth.PL DEF-problem.PL every/all-3SG.F REL
And the deeper reason for all the problems that...
bi-t-gâ:i-l-i $a$-l-fe:sbuk gajj-a mine:n
IND-3SG.F-come-to-1SG on-DEF-facebook come.PTCP-3SG.F from.where
are sent to me on facebook, where do they come from?
The deeper reasons all come: ...

mi-s-siqa
from-DEF-confidence

bi-n-nafs
in-DEF-self

°LHL
!HL%

from self-confidence.'

(65) is a particularly clear example, not only because we are provided with the context, but also because of the emphatic leading contour that is highly prominent, thereby increasing the tension and arousing the speaker's interest in what is going to come. This effect is still enhanced by the break after the verb and before the PP constituent. Figure (13) illustrates the leading-linking contour on the TOPIC with a final leading accent and a FOCUS domain falling to low, thereby marking assertion. Note also the high intensity that is maintained to the end of the TOPIC domain, with a further peak on the first noun of the FOCUS.

Fig. 13: Pitch and intensity tracks of an utterances with a relational presupposition TOPIC followed by a low falling FOCUS.

Examples (62) and (65) contain verbs that may be subsumed under the category of semantically weak or underspecified verbs, such as 'come', 'go', 'make', etc. And indeed, it is true that such verbs are often included in the TOPIC domain, being used in many frequently used collocations and the information they convey is therefore likely to be presupposed rather than asserted. But also semantically richer verbs may appear on either side of the partitioning line between the two pragmatic halves. In (66a), the verb itˤtˤallaʔu 'they got divorced' has a high information value, and what is more, it is not accompanied by a noun that could carry the main information. Nevertheless, the information is presupposed and included in the TOPIC domain. In another context, we could also imagine the verb to be asserted and associated with a closing accent. (cf. the constructed example in (66b)). The first sentence of (66a) provides the context for the next sentence. The speaker is talking about the famous Egyptian writer Ihsan Abdel Quddous and his childhood. She is explaining why Ihsans personality developed the way it did, noting that he did not live with his mother. We also know that his mother
had not died from the context before. So, the only reasonable assumption is that his parents were separated.21 This is what we are told in the second sentence, but we also get to know that the divorce took place even before the boy was born, which constitutes the only information that could not have been accommodated before the sentence was uttered. The constructed example in (66b) is assumed to have a different context. Here the sentence is supposed to directly explain certain traits of the personality of the referent of huwa 'he' in the second clause. In that case the information of the divorce itself is equally new and relevant.

(66)

a. Ihsan_FG1_06

(huwaʔawwalenmaʕaf-ʃ maʕ waldir-u) / 
3SG.MfirstNEG-live.3SG.M-NEGwithmother-3SG.M

('First, he did not live with his mother'.

waldir-u w-waldir-u tˤallaʔ-u / 
mother-3SG.Mand-father-3SG.Mget.divorced.3-PL 
LLHL²H L²H⁴L

'His mother and father got divorced,

wu-huwa f-batˤnə waldir-u 
and-3SG.Min-wombmother-3SG.M 
L²H LH L-

when he was still in his mother's womb.'

b. constructed example

waldir-u w-waldir-u (/) tˤallaʔ-u / 
mother-3SG.Mand-father-3SG.Mget.divorced.3-PL 
leading/linkingclosing

'His mother and father got divorced,

wu-huwa f-batˤnə waldir-u 
and-3SG.Min-wombmother-3SG.M 
when he was still in his mother's womb.'

Figure 14 illustrates the linking character of the whole TOPIC constituent containing linked default accents, the FOCUS domain is falling with a focus accent on the first item huwa which also has the strongest intensity.

21 In Muslim societies, the children, especially boys, frequently remain in the custody of the father after a divorce.
Fig. 14: Pitch and intensity tracks of an utterance with a relational presupposition as linking TOPIC.

Predicates introduced by the existential quantifier \( fi \): are often phrased in a separate IP (67). In the present example many items are introduced by \( fi: \), and the phonological weight of the FOCUS enhances the likelihood of a boundary, fulfilling the function of the colon in written texts. Usually there is no break between \( fi: \) and the FOCUS, in which case the existential marker almost always carries a high tone leading to the following FOCUS. In (69), \( fi \) 'in' (the etymon of the existential marker) is used as a preposition with a resumptive pronoun referring back to the topic \( il\text{-}mawdˤuʕ \) 'the issue'. Similar constructions are found with other prepositional phrases.

(67)  
\[
\begin{array}{l}
fi\text{-joːm} \quad \text{min-zaːt} \quad l\text{-ijaːm} \quad \text{kan} \quad fi: \quad /
\end{array}
\]
\[
\begin{array}{l}
in\text{-day} \quad \text{from-self.SG.F} \quad \text{DEF-day.PL} \quad \text{be.3SG.F} \quad \text{EXIST}
\end{array}
\]
‘One day there was:
\[
\begin{array}{l}
farraːn \quad \text{wu-sˤajjaːd} \quad \text{wu-gizamaːti} \quad w-husna
\end{array}
\]
baker and-hunter and-shoemaker and-H
a baker, a hunter, a shoemaker and Husna.’

(68)  
\[
\begin{array}{l}
il\text{-mawdˤuʕ} \quad fi: \quad / \quad aːzawaːja \quad \text{tasˤwiːr} \quad /
\end{array}
\]
\[
\begin{array}{l}
def\text{-issue} \quad \text{in:3SG.M} \quad \text{ah} \quad \text{angle.PL} \quad \text{photography}
\end{array}
\]
\[
\begin{array}{l}
\text{LH} \quad \text{L} \quad \text{L} \quad \text{°H} \quad \text{L} \quad \text{H}
\end{array}
\]
‘In this trick, are the fotos taken from different angles?’

As in EA verbs and other lexical categories are generally not deaccented, the linking (or even leading) character of a verb that is not part of the FOCUS is mainly expressed in the tonal layer and only rarely by reduced prominence. EA thus radically differs from the well-studied West-Germanic languages, but far less so from many other languages, e.g. languages of the Romance family. But the conceptual origin in the different languages may be the same. While in EA, TOPICS including the verb of the sentence mostly receive a linking contour with lexical prominences, in English and German the linking contour is not accompanied by strong prominences, i.e. deaccented. If the topic is made prominent, the resulting tune is a hat pattern. In EA, however, this hat pattern is paralleled by a leading-linking + closing tune, the main typological difference between the languages being the lack of deaccenting of content words in EA. Another difference is that in West-Germanic languages verbs may be deaccented, whether they belong to the FOCUS domain or not. This suggests that deaccenting
of verbs is strongly conventionalized in these languages. In EA, linguistic material may also be articulated with reduced prominence, but this is only related to the cognitive status or rather the way the speaker feels about it - whether he takes a certain information for granted or not, and not to argument structure. We will return to that issue in Section 5.7.

By way of conclusion, let me note that the explanation of TOPICS in terms of relational presupposition as has been suggested here also has the advantage of providing a unifying criterion for aboutness TOPICS involving entity topics (5.3.1) and frame TOPICS (5.3.4). Recall that in the beginning of Section 5, we have come to the conclusion that aboutness is not a good common denominator for the different types of topics suggested in the literature. Especially frames do not fulfil an aboutness function in the narrow sense. A characterization of thematic TOPIC as relational presupposition, meaning something like “with respect to what we already know”, however, holds for frame TOPICS as well.

5.3.3 Some figures

At the current state of annotation and search options it is unfortunately not possible to give a comprehensive quantitative account of the prosody of the different TOPIC types discussed in this section. The only reasonably accurate statements can be made concerning S-TOPICS. Therefore I conducted a small count of subject NPs and independent subject pronouns and cross-classified them with the four tonal categories. Note that this cross-classification could also only be conducted on a subset of the data excluding modifiers. As is well-known, adjectival modifiers are not especially frequent in spontaneous speech, but in EA nominal modifiers in an annexation construction or the analytical possessive construction with the possessive marker (genitive exponent) bitaːʃ are frequent. These could not be captured by the search procedure. However subjects with a possessive pronominal suffix could be included. Furthermore, this count refers to non-subject pronouns and nouns with a demonstrative determiner separately, wherever they occurred within a TOPIC phrase. The results are shown in Table 3.

The cross-classification shows the preponderance of the leading-linking accent (50% of the cases), followed by the linking type. Among the interesting aspects of the count is the behaviour of pronouns. Although in both categories, the leading-linking accent is the most common one (almost 50%), we also find a fair number of linking tones on pronouns (37% for subjects, 27% for non-subjects). Pronouns are given by definition and low prominence is thus an expected feature. The quantitative results thus support what has been said about pronouns in Section 5.3.1.5, namely that the prosodic shape of the pronoun correlates with its function as a continued topic expression or a contrastive or at least shifted topic.
Another interesting aspect is the relatively high number of closing accents in a TOPIC. The number is higher in non-subject pronouns and lexical subjects than in subject pronouns. The assertive accent on lexical NPs will be discussed below in Section 5.7.2.

The leading accent is the least frequent in this count. This can be explained by the fact that NPs occurring in D-TOPIC or R-TOPIC constructions are excluded from the statistics. Compared to the overall proportion of leading tones, the demonstrative determiner shows a rather high proportion of leading accents which comes as no surprise. Less up to expectations is the equally high number of linking tones. These results clearly show that demonstrative determiners in EA may well be unaccented, contrary to Brustad's (2000: 113) assumptions.

As noted above, a quantitative account of the different TOPIC types cannot be given yet. This must await further investigation. But at present, a crude estimation could be obtained indirectly which suggests that the qualitative observations made above indeed seem to hold. This preliminary estimate suggests that R-TOPICS and D-TOPICS including pseudoclefts have a much higher proportion of final leading accents than S-TOPICS. Future research will have to investigate this aspect and crucially also the tonal character of the whole contour associated with long TOPIC domains and their relationship with the contour terminal, i.e. the final pitch accent.

Finally, let us turn to the distribution of the linking tone. As already noted, long downtoned or even deaccented stretches of speech are not all too frequent in spontaneous dialogue. All topic expressions annotated with a linking tone (L-) - which means that they were actually classified as deaccented - were 132 instances. Unsurprisingly, the overwhelming majority of these are prononinals which are inherently given. The small print figures in the third columns relating to the pronouns give the proportion of demonstrative pronouns among them. The strong ban against deaccentuation of content words produces the small number of 10 items of low-toned deaccented lexical NPs. But note that this
is not the total number of downtoned expressions, as NPs with low prominence may also occur in the rising or falling portions of a rising-falling contour or they may be articulated in a reduced pitch range. At closer inspection of the low-toned NPs that were included in the count it turns out that many of these are nouns in the construct state, i.e. first nouns of an annexation. This fact underlines that such forms are treated as quasi-compounds with a single pitch accent on the second noun. The minimal number of other occurrences comprise continuous topics, but also two case of shifted topics and nouns with low semantic content such as il-hika:ja 'the story' (meaning the issue, used as a kind of dummy topic).

5.3.4 Frame TOPICS

As noted above, I assume that frames are functionally different from aboutness TOPICS, following Chafe (1976) and Jacobs (1997). The definition of frames as a separate TOPIC category is implicit in Chafe's topic definition as "the spatial, temporal or individual framework within which the main predication holds" (Chafe 1976: 50; emphasis added). One such example was already discussed in (10b) above. There I suggested that 'education' cannot be considered an aboutness topic as it is unlikely that Strawson's suggestion of 'increasing our knowledge about' the entity would hold for this type of TOPIC. The function here is rather to "delimit the frame" within which the predication holds. In the case of (10b) 'prefer Bertrand Russell's ideas' is predicated of John and the framework within which this predication is true is the realm of 'education'. In the following paragraphs, I will not be concerned with such conceptual frames here (they have already been mentioned in passing in the previous section), but I will predominantly discuss the other two types identified in Chafe's definition, namely spatial and temporal frames. In the EA corpus, I analysed 287 frame TOPICS. The count was done manually because the automatic search could not have identified the different frame types and the different contours associated with them. As can be seen in Table 4, most frames are temporal. Spatial frames are far less frequent. Syntactically, frames may consist of adverbials, prepositional phrases and full clauses, such as the temporal frames introduced by conjunctions such ba:šdø-ma 'after' and źa:blo-na 'before'. Temporal and spatial adverbial phrases may also be positioned at the right periphery or in the middle of the sentence. Such instances, however, do not count as frame TOPICS. In the following study I differentiate between leading, linking and closing TOPICS and indicate phrasing separately. Table 4 gives an overview of the instances found in the corpus, classified as temporal, spatial, conditional clauses (IF-TOPICS, see 5.3.4.4) and other.
The results in Table 4 give rise to a number of observations: (1) Over 80% of all frame TOPICS are temporal. We will, however, see below that we may interpret some of these differently. (2) In line with the hypothesis that TOPICS usually have leading or linking intonation, it is only a very small number of frames that have a closing accent at the end (8%). (3) Frames may or may not be phrased separately in an independent IP, the number of breaks is only slightly higher than the number of non-breaks (142 vs. 122). At the same time, leading frames are more prone to be phrased separately, especially if they are full clauses. The higher number of breaks after a linking clausal frame also reflects the tendency to use separate IPs for clausal units. On the other hand, non-clausal constituents, such as adverbials are more readily phrased together with the main clause, if they have a linking intonation. (4) Temporal and spatial frames strongly tend to have a linking intonation. As the figures show we have 175 linking frames and less than half that number of leading frames (89). It is noteworthy to mention here that a leading frame was classified according to the whole contour. If the contour was rising, it was classified as a leading frame and if the contour was more or less horizontal or slightly falling, the frame was classified as linking. On the other hand, the small number of closing frames was classified according to the final accent. But note that many of the examples were very short containing only one accent. In the following, we will therefore undertake a separate investigation of the final accents the results of which are described in the next section.

5.3.4.1 Frames with leading and closing final accents

From the 289 frames investigated, only 40 were found to be ending in a leading accent (cf. example (69), Figure 15). Four of these were inverted accents (L^H') and twelve had two adjacent H tones (H^H') with the second H being a boundary tone or a second accent. The rest were simple rises (L^H').

<table>
<thead>
<tr>
<th>FRAMES</th>
<th>linking</th>
<th>leading</th>
<th>closing</th>
<th>TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>break</td>
<td>no break</td>
<td>break</td>
<td>no break</td>
</tr>
<tr>
<td>temporal frames</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>non-clausal</td>
<td>36</td>
<td>63</td>
<td>18</td>
<td>14</td>
</tr>
<tr>
<td>clausal</td>
<td>30</td>
<td>23</td>
<td>36</td>
<td>2</td>
</tr>
<tr>
<td>spatial frames</td>
<td>6</td>
<td>6</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>IF-TOPICS</td>
<td>2</td>
<td>2</td>
<td>8</td>
<td>2</td>
</tr>
<tr>
<td>other</td>
<td>1</td>
<td>6</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>TOTAL</td>
<td>75</td>
<td>100</td>
<td>67</td>
<td>22</td>
</tr>
</tbody>
</table>

Tab.4: Frame TOPICS.
Let us now look at the type of frames associated with these final rising accents. Firstly - as already suggested by the figures in Table 4 - what I called IF-TOPICS, i.e. conditional clauses, are frequently characterized by a leading contour and/or a leading final pitch accents as in (69) (Fig. 15).

(69) Mar_Shab_FG3_17

\[
\begin{array}{cccc}
\text{fa} & \text{law} & \text{il-bani\^a:dam} & \text{dawwat} \\
\text{SEQ} & \text{if} & \text{DEF-human.being} & \text{DEM.SG.M} \\
\text{L} & \text{LH} & \text{L} & \text{L^H} \\
\end{array}
\]

'If this person is a fighter...'

Fig. 15: Pitch and intensity tracks of an IF-TOPIC with high rising terminal.

Another type of leading frame TOPIC has been included in the category of temporal frames above. But at closer inspection, these frames are not really 'temporal' in the sense that they specify the time of the proposition as for example in the frame typically used in fairy tales fi-jo:m min il-\^ajja:m 'once upon a time' or imba:rih 'yesterday' or bi-l-le:li 'in the evening' and the like. We will have a closer look at such frames below in Section 5.4. The semantics of such frames is rather sequential or circumstantial. They frequently express a relation between a point in time or in the sequence of chronological events that constitutes the starting point for the following proposition (70-72). A special case of such frames is the word ba\^de:n 'and then' - especially frequent in narratives - which exhibits different types of intonational tunes (cf. Section 5.4). Ba\^de:n, in fact, is frequently not temporal at all, but rather behaves like a discourse marker introducing a new episode which need not at all be temporally subsequent to the preceding one and thereby marking a specific and point in discourse rather than in time.

(70) Mido_F0_02_13

\[
\begin{array}{cccc}
\text{ra\^h} & \text{\^awwil-ma} & \text{r^\^a:gil} & \text{fa:f} \\
\text{SEQ} & \text{first-SUB} & \text{DEF-man} & \text{see.3SG.M} \\
\text{L} & \text{LHL} & \text{LHL} & \text{\(^H\text{H}^H\)} \\
\end{array}
\]

'As soon as the man saw the frog...'

(71) Mar_Siqa_FG3_08

\[
\begin{array}{cccc}
\text{\^awwil} & \text{ha:ga} \\
\text{first} & \text{thing} \\
\text{H-} & \text{L^H} \\
\end{array}
\]

'The first thing is...'
In (72) the sequential marker introduces the final step in a recipe for Bamya ('ladyfingers').

Similarly, temporal clauses introduced with kull-(am)ma 'whenever' (74) or lamma 'when' specify the actual reference point for something to happen; most of the lamma-sentences can be felicitously translated with 'as soon as' (75; Fig. 16). By the same token, circumstantial clauses constitute TOPICS that specify the point of departure (76).

![Pitch and intensity tracks of a clausal frame with a leading contour (ex. 75).](image)

Fig. 16: Pitch and intensity tracks of a clausal frame with a leading contour (ex. 75).

As he set about drinking...’
I hope I am not stretching the iconic explanation for the use of linking vs. leading intonation too far, when I suspect that a typical temporal frame, especially one that only identifies a vague time frame, does not motivate a high rise. This is contrary to a temporal reference point that, similarly to an individual one, gives rise to leading intonation, suggesting a meaning that is also implicit in the 'aboutness' intonation discussed in the preceding section. It seems that the speaker is trying to draw the hearer's attention to what is coming next, implying something like 'Watch out now, what's coming next!'. The same is also true for conditional clauses (Section 5.3.4.4). Many of the temporal frames of the type investigated here, in fact come close to conditional clauses in the sense that the comment following the frame TOPIC is somehow a consequence of the situation or event expressed in the TOPIC. In examples (74) and (75), for instance, the proposition in the comment is dependent on the proposition expressed in the TOPIC. In (74) the jumping of the frog is not only temporally related to the waiter's attempt to catch it, but the two activities also have a cause-effect relation. By the same token, in (75) the temperature of the chilli pepper (il-f'ilfil) that has been roasted is the necessary precondition for it being preserved in a jar.

A final category that partly exhibits leading intonation is the family of discourse marking elements, many of which also show a closing pattern in the corpus. They will be given a separate treatment in Section 5.4 below.

Coming now to the closing pattern: the results of the investigation show that whenever a closing accent is used on the final item an assertive flavour is associated with the message. Moreover, such frames are in fact parenthetical, i.e. they are following another frame or a topic expression that carries a forward-looking leading/linking intonation (77, 78) or the frame is not phrased separately and followed by a leading/linking item as in (79).

(77)  Mar_Shah_FG3_08

\[
\begin{array}{llllll}
\text{fi:} & \text{bana:t} & / & \text{min} & \text{u-humma} & \text{s'wajjar}^{4}\text{a} & / & \text{tixi:n-a}^{22} & /
\end{array}
\]

\[
\begin{array}{llllllll}
\text{EXIST} & \text{girl.PL} & \text{from} & \text{and-3PL} & \text{little-SG.F} & \text{fat-SG.F} \\
\text{LH} & \text{LH-} & \text{L} & \text{°H} & \text{L!HL} & \text{LHL%}
\end{array}
\]

'There are girls - from the time on when they were little - (they hear nothing but you are too) fat, ...'

(78)  Mar_Shah_FG3_20

\[
\begin{array}{llllllll}
\text{ʔawwil} & \text{jo:m} & \text{fi-g-gamša} & / & \text{ʔawwil} & \text{jo:m} & \text{xa:lis}^{3} & /
\end{array}
\]

\[
\begin{array}{llllll}
\text{first} & \text{day} & \text{in-DEF-university} & \text{first} & \text{day} & \text{totally}
\end{array}
\]

\[
\begin{array}{llllll}
\text{LH} & \text{LH-} & \text{°H} & \text{LH} & \text{°HL} & \text{!HL}
\end{array}
\]

'On my first day at university - on the very first day!'

---

22 The pronunciation tixi:na without velarization of the alveolar stop is a typical sociolinguistic characteristic of (young) high-class women.
In (77) the downstep accent asserts the fact that some girls have been told all of their lives that they are ugly and will not be able to find a husband, even when they were still little kids. Similarly, the speaker asserts in (78) that it was her very first day as a university teacher when something awkward happen to her in class. Finally in (79), the speaker closes the series of events with a closing accent on fi-l-ʔa:xir 'in the end', the leading intonation being associated with the following topic il-garso:n 'the waiter' without a phrase break in between.

(79) Mido_F6_02_04
fi-l-ʔa:xir il-garso:n / in-DEF-last DEF-waiter
LHL L^H'
'At last the waiter...'

The examples shown in this section are ample evidence for the Bolingerian view that it is not possible to tease apart the grammatical and the attitudinal in intonation. The type and strength of accents thus do not only depend on the information structure, let alone the syntactic structure of the linguistic material they are associated with, but at the same time reflect the speakers attitude as well as phatic aspects. In the following sections I will present the different frame types that can be differentiated according to their semantic type.

5.3.4.2 Temporal frames

As we have seen above, the majority of frame TOPICS are temporal. In Section 5.3.4.1 we were concerned with 'temporal' frames that indicate a specific reference point for the comment to happen, either as a starting or end point. In this section, we will be discussing more vague time frames that for the most part specify a certain period within which the event expressed in the proposition is taking place. Such frames are normally characterized by a linking contour, the final accent of which also tends to be of the linking type, i.e. the H tone is mostly continued until the end of the phrase (80, 81). The tonal contour after the stressed syllable may equally well be slightly falling or slightly rising - perceptually this does not make a strong difference (cf. Bolinger's description of the B accent, Bolinger 1986: 152). In any case, the terminal is what Delattre (1966) called a "continuation mineure".

(80) HR_F0_02_02
jum-ḥa tlamme:-na ʕaleː w-raḥham-na:-ha min-ʕalʔa
day-3SG.F gather-1PL on:3SG.M and-rescue-1PL-3SG.F from-slap
LH- LHL H L^H L!HL%
'That day we held him back and saved her from a slap in the face.'

(81) Fagr_F3_08
taːni joːm / itʕasʕal bi-kull il-ʕeːla
second day contact.3SG.M PREP-all/every DEF-family
L- H-
'Next day, he phoned the whole family.'
Fig. 17: Pitch and intensity tracks of two temporal frames with a linking contour (ex. 80, ex. 81).

(82) Bamya_F3_03

\begin{verbatim}
  w-baśdə  kida  ha-n-ʔammaš-ḥa / w-baśdə  kida  ha-n-qiːb...
  and-after so  FUT-1PL-fry-3SG.F and-after so  FUT-1PL-bring
  L-  H-  L\(^H\)  L-  H  LH

'And then we fry it... and then we bring...'
\end{verbatim}

In (82) the adverbial baśd kida 'after that' refers to a temporal sequence without implying a new episode or a special event that is going to follow, contrary to (72) above, which is from the same text, where the speaker used a leading contour to introduce the final step in the process of making Bəmə.

(83) Taxi_F0_01_02

\begin{verbatim}
  baśdə  ʕafa  tijjəːm  ʔul-tə  l-il-hagg-a ...  
  after 10 day.PL say-1SG to-DEF-pilgrim-SG.F
  L-  H  LH  "HL"  L\(^H\)-

'After ten days I said to my wife ...'
\end{verbatim}

(84) Taxi_F0_08_01

\begin{verbatim}
  w-ʔawʔaːt  ʔiːnj-a  kun-na...
  and-time.PL second/other-SG.F be-1PL
  L-  H-

'And at other times we ...'
\end{verbatim}

(85) Sihr_MG2_08_01

\begin{verbatim}
  wi-halijan  b-a-s'annaʃ  aːːːi  ʔalʔaːb-i  bi-naʃ-i
  and-at.present IND-1SG-produce HESIT game.PL-1SG PREP-self-1SG
  LH  LH  LH  L\(^*\)-

'And at present, I produce my tricks on my own.'
\end{verbatim}

All examples so far illustrate non-clausal frames: adverbs or nominal adjuncts. All examples except (82) phrase the TOPICAL element and the main proposition together in one IP. As may be seen in Table 4 above, this is the most common type of temporal frames in the corpus. By contrast, (86) illustrates a clausal frame which is separated from the following main clause by an intonation break, but nevertheless exhibits a linking contour.

323
'And when you open your heart ...'
Fig. 19: Pitch and intensity tracks of two typical introductions to a narrative.

To summarize, temporal frames are the most frequent type of frame TOPICS in the EA text corpus. Functionally, temporal frames are characterized in the Chafean sense as the 'temporal [...] framework within which a predication holds'. We have seen that temporal frames mostly show a linking intonation, but there also are a number of cases that exhibit leading intonation. Such cases were shown in Section 5.3. As it seems, a leading contour and especially a leading final accent indicate a certain reference point for the main proposition to start or end at or to happen before or after another point in time implicit in the situation. Some 'temporal' frames could also be shown to be not only temporal, but involving a conditional meaning.

The typical temporal frame, however, has a linking contour. Whether there is a break or not, is again dependent on the speaker's choice and realizational factors such as speech tempo. It is likewise influenced by the phonological weight of the constituent and its semantic and syntactic status, i.e. whether it is an adverb or a clause. As we have seen, clauses are generally more likely to be phrased separately.

5.3.4.3 Spatial frames

Spatial frames are quite rare in the corpus, i.e. spatial expressions that are sentence initial are rare. As noted above, I only investigated sentence initial occurrences of temporal and locative expressions for the analysis of frame TOPICS. This investigation yielded an imbalance between these two types of adverbials (233 temporal frames and only 14 spatial frames). It goes without saying that this imbalance is not one of temporal and locative expressions in general. In Table 5 we may see the distribution of such adjuncts as they occur in the different positions within the utterance. Recall that in the annotation, an utterance - if complete - was divided into an obligatory middle field, an optional preverbal preffield and an optional postfield. The following statistics do not differentiate between locative expressions that are adjuncts and those that are selected by the verb, however.
As expected, the number of locative and temporal expressions is approximately the same, but as Table 5 clearly indicates, they differ enormously in their preferred positions within the sentence. Temporal adjuncts tend to be in preverbal position (63%), and only half of those within the clause are positioned clause-finally. Occurrences within the postfield are comparably rare (5%). This result stands in stark contrast to the distribution of locative expressions. Here the opposite is true: only 7% of all locative expressions are in TOPIC position and the overwhelming majority is integrated in the main clause (74%), two thirds of which occur clause-finally. Compared to preverbal locatives, the number of post-clausal ones is rather high (15%). This distribution explains the striking difference between temporal and spatial frames shown in Table 4 above.

(90) Taxi_F0_08_01
hina:k fi-ja:ril’i ig-ge:j /
there in-street DEF-army
L^H LHL ^H
There in Gesh-Street ...'

(91) Makarona_F0_02
fi-n-nahja t-tanj-a /
in-DEF-side DEF-second-SG.F
LH^ L^H^ 'On the other side...'

(92)

a. Taxi_F0_11_04
fi-f-suyl la:zim maz’har-sa ji-ba kwajjis
in-DEF-work necessary appearance-1SG 3SG-M-be/become good
LHL H LH L H^L
'At work, my looks have to be good.'

   b. Taxi_F0_11_05
fi-be:t-i / wu-fi-l-hitta /
in-house-1SG and-in-DEF-piece/place
L-LH- L-LH-
'At home, and in the neighbourhood...'

The example in (90) shows a typical spatial frame with clear locative semantics. It has a leading contour, but no leading final accent as the contour starts falling towards the first low tone within the following IP. Example (91), by contrast, is not really locative. 'On the other side' may rather be
understood metaphorically as 'on the other hand' in this context. Here the final accent is a rising one, an inverted accent L*H, which indicates the starting point of a new episode - in this case preparing the pasta in addition to the pasta sauce for the Makaro:na fi-l-furn (the Egyptian equivalent of the Greek Pastizio) whose making the speaker described in the previous episode.

The final example (92) shows a spatial frame with a closing accent in (92a) and typical linking accents (92b) (see Figure 20). The second example follows the first one in the text, the first frame fi-f/uyl 'at work' is contrasted with the following fi: be-t-i wi-fi-l-hitta 'at home and in the neighbourhood'. The reason for the closing accent may partly be the reintroduction of the work as a topic (the previous sentences had been about the speaker's morals), but the main reason most probably is the contrast between the two constituents (see 4.5.4).

Fig. 20: Pitch and intensity tracks of two linking spatial frames in sequence.

In sum, the low number of spatial frames does not permit a clear statement as to its preferred intonation. However, the number in Table 4 together with the observations made on the few examples suggests that the preferred intonation is the same as the one of temporal frames, namely a linking contour.

5.3.4.4 IF-TOPICS

In Table 4 of Section 5.3.4.1, conditional clauses as TOPICS were counted separately. Of course, the number of examples is small, but again the distribution of leading and linking contours permits the hypothesis that IF-TOPICS, in contrast with temporal and spatial frame TOPICS, are preferably leading. We have already seen that some of the temporal frames are in fact semantically conditional (cf. examples (74) and (75)). In this section, we will come across other types of conditional phrases that are not explicitly marked as such. If we add these two types to the number of explicit conditional clauses, we get results that are slightly more robust.
The following paragraph contains a number of conditional clauses in sequence. A magician is illustrating one of his tricks involving four differently coloured balls (white, yellow, black and red) tied together by means of a thin rope which is partly hidden inside a hollow pole. From each side of the pole there are two balls dangling down. When he pulls one ball, another ball on the opposite side goes up, but at the same time pulling one ball may also move the other one on the same side. I will give the whole sequence here; complete annotation is restricted to the first sentence, in the following sentences only the IF-TOPICS are annotated. Four of them are illustrated in Figure 21.

(93) Sihr_MG2_12
kull ik-kuwar murt‘abit‘-a b-baºd‘ / jaºni masalan
all/every DEF-ball.PL linked-SG.F PREP-part FILLER for.example
‘All balls are linked to each other. For example, ...

ni-fidd il-hamr-a / il-be:d‘a ti-tºlaº /
1PL-pull DEF-red-SG.F DEF-white-SG.F 3SG.F-go.up
LH LºH-H L H LºHL%
if we pull the red one, the white one goes up.

law fade:-na il-be:d‘a / il-hamr-a / ti-tºlaº / tºab
if pull-1PL DEF-white-SG.F DEF-red-SG.F 3SG.F-go.up ok
L LH LH
If we pull the white one, the red one goes up. Ok...

ni-fidd is=ºsªfi-a / ha-n-laº:ha maºj-a /
1PL-pull DEF-yellow-SG.F FUT-1PL-find-3SG.F go.PTCP-SG-F
LH L*H-H
if we pull the yellow one, we find that it moves...

maº il-be:d‘a / il-itnen do:l li-wahd-uhum / maºj-i:n
with DEF-white-SG.F DEF-two DEM.PL to-one-3PL go.PTCP-PL
with the white one. Those two alone go...

maº baºd‘-i:hum / tºab ni-fiddo bard’u l-hamr-a /
with part-PL ok 1PL-draw also DEF-red-SG.F
L H 0ºH 0ºHºH*ºH

0
Time (s) 0.7168 0.6121
(nishidd il-Hamra)

Together. Ok, if we again pull the red one,

il-be:d‘a ti-tºlaº
DEF-white-SG.F 3SG.F-go.up
the white one goes up.

(a)

(b)
Example (93) permits the following observations: (1) All IF-TOPICS have leading contours, but only the last one is rising to the very end. We have already seen several times that final elements in a series of leading contours are rising higher than the ones before. (2) The final accent in the second conditional (Figure (21c) and (21d)) is classified as an inverted accent (L\*H). Inverted accents are typical for repetitions. In terms of structure, the third and fourth frames are both repetitions, the final one is even repeating the same contents, the pulling of the red ball (cf. lines 2 and 6 in ex. 93). (3) Finally, it may be noted that only one of the three conditionals is marked explicitly by the conjunction *law* 'if', while the other two instances are simple declaratives with a conditional meaning.

Such IF-TOPICS in disguise are not rare, (94) gives another example of that kind. Here it is a PP that can be interpreted as an IF-TOPIC in the context at hand. The expression 'without paper and pencil' in fact means: 'If I don't have paper and pencil'.

(94) Mar_Shab_18

\[
\begin{array}{l}
\text{minye: } \text{wara}\{?\} \text{ w-\text{alam... } } / \\
\text{without paper and-pencil} \\
\text{LH } \text{°H } \text{L°H^H} \\
\text{Without paper and pencil...}
\end{array}
\]

Again, the low number of if-TOPICS does not permit any clear statements about the preferred intonation type. However, the small number of examples suggests the hypothesis that leading contours and separate phrasing are again preferred.

---

23 It has, however, be noted that the H-tone in the final accent in the first if-TOPIC (Fig. 21a) is only marginally higher, which points to the gradience of scaling phenomena. It must be taken into account that the global pitch range of speaker MG2 is very narrow and thus the scaling differences are generally not very large.
5.4 Discourse markers in TOPIC position

Finally, some elements that typically occur in TOPIC position look like temporal frames, but at closer inspection are not really frames, or at least they are on the borderline of this category. We have already seen some examples in the section about temporal frames (5.3.4.2) where I distinguished them from real frames on functional grounds. In this section I will show some additional examples and discuss their occasionally peculiar prosody. The other elements that occur in the left periphery of the sentence are sentence adverbials and connectives which are primarily used in the narratives. This hotchpotch category was referred to as 'other' in Table 4. Like real TOPICS, such elements do not only serve to introduce the following proposition, they are also used to establish coherence within the text.

The most frequent connective elements in the narratives of the EA text corpus are *il-muhimm* 'the important (thing)', *faqʔa* 'suddenly', *fwajja* 'a little' (meaning 'after a while') and *baʔde:n* 'afterwards' or the variant *baʔdo kida* 'after this'. The latter have already been alluded to in 5.3.4.2 above. We will see that they are frequently not used in a genuinely temporal sense, but only serve the discourse function of starting a new episode.

Interestingly, we find a variety of intonational variants with these elements. Let us begin with instances of *il-muhimm*. Most instances in the corpus come from one speaker who is not a very gifted storyteller. What is important, however, that we can find a difference in the prosodic shape of the markers as used by this speaker and also in the rare instances from another speaker, whose interpretation is a very tentative one. We find all three different types of contours associated with the marker. The default accent in (95) does not seem to have any special information structural effect. The leading type (L^H), exemplified in (96), seems to be associated with a topic shift. Here the topic shifts from Mido to the dog, the other salient participant of the story, where Mido and his dog are looking for their mate, the frog. At the same time, a new episode begins describing the dog looking at a tree. In the episode before, Mido had been frightened by a squirrel and run a way.

(95) Mido_M1_01_02

*m-muhimm* m:idu d-daːjiʔ...
DEF-important M PASS-annoy.3SG.M
LH LH^L °HL
'Anyway, Mido was really annoyed.'

(96) Mido_M1_01_03

*m-muhimm* / *il-kalb*
DEF-important DEF-dog
L^H LH-
'In the meantime, the dog...'

There is also a number of instances that have a closing intonation as shown in (97) (without a break) and (98) (with a break). These instances are not necessarily associated with a change of the topic.
They, however, clearly mark an episode change. In (97), Mido is the discourse topic of both episodes. The first episode devoted to introduce a little boy named Mido, his friends, a dog and a frog, and his room. In the second episode some background information is given stating that Mido had put the frog into a jar overnight. The sentence topic of the sentence preceding the discourse marker was the room. Thus, we can assume a topic change to have happened. It is, however not the case that the discourse topic is changed, being Mido in both cases. The closing intonation seems to have a cataphoric character, making a full stop before moving on to a different theme.

The same interpretation presumably holds for example (98) where a taxi driver is driving a woman who gets on the taxi fully veiled in a burka, and during the trip strips off layer after layer until she ends up wearing a mini skirt, black tights and a tight blouse. The present situation is after the first round of undressing when the burka has been removed revealing a simple head scarf.

(97) Mido_M1_01_02
im-muhimm mi:du kan ha:rítíí idíí-duííí / fi kubba:ja
DEF-important M be.3SG.M put.PTCP.SG.M DEF-frog in glass
L^HL H L HL L^H- H^ L%
'Anyhow, Mido had put the frog into a jar.'

(98) Taxi_F0_11_02
im-muhimm / laʔi-t-ha muhaggab-a
DEF-important find-1SG-3SG.F veiled-SG.F
LHL LH^ L H^HL
'To make a long story short. I found her wearing a head scarf.'

Another expression that is typically used in narratives is fagʔa 'suddenly, lit. 'sudden'. In accordance with its original emphatic meaning, we find it associated with double accentuation and a leading contour in (99).

(99) Taxi_F0_11_04
fagʔa / ik-kalb wiʔíí barʁʔa
suddenly DEF-dog fall.3SG.M outside
LH^H* H !HL !HL
'Suddenly, the dog fell out [of the window].' (lit. fell outside)

The most frequent of the connectives is baɗe:n 'afterwards'. The following examples show the range of intonational variation associated with this item. The examples also reveal that baɗe:n is not only used in its temporal meaning, but that the contextual use is much more varied, suggesting its pragmationalization as a (sequential) discourse marker. (100) is a good example of baɗe:n meaning something like 'furthermore'. Its combination with the temporal dilwaʔtí 'now' and the context exclude a temporal interpretation.
Similarly, *baːdeːn* is simply used, when the speaker does not know how to go on with his story. In one example Mido had been looking out of the window, his dog standing beside him and the dog fell out of the window on the ground. After a short pause (450 msec), the speaker continues with ‘*wu-baːdeːn* / he may have hurt himself’ (Figure 22b). Again, a temporal interpretation does not make sense here.

Fig. 22: Pitch and intensity tracks of four instances of the discourse marker *baːdeːn* in TOPIC position.

By contrast, an example of a truly temporal sequential marker is depicted in Figure (22a). The marker is associated with a simple default accent that is fully integrated into the IP containing the main clause. Likewise, a temporal interpretation is not excluded in the other two instances illustrated in Fig. (22c, d). However, it is more than simple continuation of the story that is indicated by the marker. Both instances are surrounded by long pauses and both involve a major break in the story. From a prosodic point of view it is interesting to see that the variation lies between a linked accent (default) over a simple rise ending with a linking contour (a leading-linking accent), mostly accompanied by a break, to more complex tunes involving a fall(-rise) with a high beginning (Figure 22c,d). Looking at the four pictures more closely, reveals additional variation in prominence. The intensity values indicate that the
example in (Fig. 22a) involves only one accent in the expected location, the second syllable of bāṣid:ē:n, whereas the accent is retracted to the first syllable in the example in Figure 22c. Other examples in the corpus exhibit double accents and the fourth example in Figure 22d even shows a triple accent (H*L*H*-) associated with one word. Such variation is clearly attitudinal and once more points to the necessity to draw an a priori distinction between the tonal and the metrical component, and particularly also between the prosodic component and syntactic structure.

5.5 Morpho-lexical topic markers

A common test for topics is the as for-test (Gundel 1976). Constructions such as as for or concerning or as far as X is concerned are used as topic marking devices in English. Their Arabic equivalent is usually supposed to be amma - ūa. This construction, typical for Standard Arabic, is not frequently used in EA. Amma is a typical wusta phenomenon and may thus be considered a loan from Standard Arabic when it is used. Consequently, it is only found in texts that are especially prone for Colloquial-Standard code switching such as learned discussions. Not surprisingly, the only incidence in the EA text corpus is in one of the written stories. Amma, however, does not seem to be a general topic marker, its occurrence seems to be restricted to special contexts. Pashova (2003: 16) notes that in her corpus of written Standard Arabic amma is only used when the topic is in contrast to previous topics or in listing or a series of topics. This is also my observation about its use in MSA. Another explanation would be that amma is associated with a switch in discourse topic. The only EA example in my corpus has this function. The narrator has been talking about films and described his favourite American films with Charles Bronson starring. He now turns to his favourite Egyptian film (101):

(101) Taxi_F0_08_02

amma ʔahla film ə masˤr-i bi-n-nisba:l-i / TOP most.beautiful film Egypt-NISB.SG.M PREP-DEF-relator/to-1SG

‘As far my favorite Egyptian film is concerned,

ka:n / sawwaʔ il-utubi:s be.3SG.M driver DEF-bus

this was 'The busdriver'.

There is no real EA equivalent to this construction, apart from the possibility of saying bi-n-nisba l- 'in relation to' which is morphologically identical with the corresponding MSA expression. EA, however, seems to have another particle that frequently occurs between a topic and a comment. This is the particle baʔa which etymologically comes from a verb baqija 'remain'. The EA verb baʔa itself does not mean 'to remain', but 'to become' or simply 'to be'. The particle baʔa can, however, not be considered a topic marker. First, it is optional, second its main function, in accordance with the original verbal meaning is something like indicating a sequence, and third, it also occurs after FOCAL constituents. Prosodically, being a particle, it is frequently unaccented, but it is possible to accent it as in (107). The particle clearly imposes salience on the constituent it is related to. Insofar, it would not
be surprising if it were typically used to indicate contrast or at least topic shift. In the following example (102), this latter function is indeed present. The speaker, after a digression to other issues, comes back to the main discourse topic of the whole discussion: 'self-confidence'. In many of the examples baʔa is used in the same function as English 'now' in the well-known example of the 'wizard' who is re-introduced as a sentence topic after a passage about his wife and sons as 'Now the wizard, he lived in Africa.' (cf. Lambrecht 1994: 177).

(102) Mar_Siqa_FG3_12

<table>
<thead>
<tr>
<th>il-hika:ja</th>
<th>bau'-t</th>
<th>is-siga</th>
<th>bi-n-nafs</th>
<th>baʔa</th>
</tr>
</thead>
<tbody>
<tr>
<td>DEF-story</td>
<td>POSS-</td>
<td>DEF-conf</td>
<td>PREP-DEF-self</td>
<td>PRT</td>
</tr>
</tbody>
</table>

'Now, coming back to this whole issue of (this) self confidence di
DEM.SG.F and-so
and the like..' But note that topic activation is not a necessary condition for the use of baʔa. (103) shows an utterance where the topic da 'this (M)' is referring back to a film that was already the topic of the preceding sentence. Instances of baʔa together with an anaphoric referential expression such as (demonstrative) pronouns are relatively frequent.

(103) Taxi_F0_08_02

<table>
<thead>
<tr>
<th>da</th>
<th>baʔa</th>
<th>juf-t-u</th>
<th>ji:gi</th>
<th>/ ʕaʃar marr-a:t</th>
</tr>
</thead>
<tbody>
<tr>
<td>DEM.SG.M</td>
<td>SEQ</td>
<td>see-1SG-3SG.M</td>
<td>approximately</td>
<td>10 time-PL</td>
</tr>
</tbody>
</table>

'This one I saw maybe ten times.'

In (104) we see an example of a new topic, an indefinite plural noun that is referred to in the subsequent sentence, indicating the topic status of the referent in the following proposition. Functionally it is a spatial frame. The sentence could be paraphrased as in (104b), only that such a paraphrase would require borrowing the construction baʕdˤ il- 'some' or something similar from MSA.

(104)

a. Taxi_F0_09_02

<table>
<thead>
<tr>
<th>simiʕ-ta</th>
<th>haːdˤit-ak</th>
<th>ʔinnɔ</th>
<th>fi:</th>
<th>bila:d</th>
<th>baʔa</th>
</tr>
</thead>
<tbody>
<tr>
<td>hear-2SG.M</td>
<td>HON-2SG.M</td>
<td>COMP</td>
<td>EXIST</td>
<td>country.PL</td>
<td>SEQ</td>
</tr>
</tbody>
</table>

'Have you heard that there are countries, ...

<table>
<thead>
<tr>
<th>ʕadad</th>
<th>is-sitt-aːt</th>
<th>fi:-ha</th>
<th>/ ʔaktarˤ bi-kitiːr giddan</th>
</tr>
</thead>
<tbody>
<tr>
<td>number</td>
<td>DEF-woman.PL</td>
<td>in-3SG.F</td>
<td>more PREP-much very</td>
</tr>
</tbody>
</table>

in which the number of women is much higher

<table>
<thead>
<tr>
<th>ʕadad</th>
<th>ir-riggaːla</th>
</tr>
</thead>
<tbody>
<tr>
<td>number</td>
<td>DEF-man.PL</td>
</tr>
</tbody>
</table>

from number | DEF-man.PL |

than the number of men.'

b. constructed example

<table>
<thead>
<tr>
<th>simiʕ-ta</th>
<th>haːdˤit-ak</th>
<th>ʔinnɔ</th>
<th>fi</th>
<th>baʕdˤ il-bila:d</th>
</tr>
</thead>
<tbody>
<tr>
<td>hear-2SG.M</td>
<td>HON-2SG.M</td>
<td>COMP</td>
<td>in</td>
<td>country.PL</td>
</tr>
</tbody>
</table>

'Have you heard that in some countries, ...
The particle is frequently inserted in syntactic positions that are not normally filled by a lexical or grammatical element, such as within an annexation as in and (106a) and between a noun and a demonstrative determiner as in (105). Furthermore, example (106a) illustrates that it may occur within the FOCUS domain. Obviously its position makes it possible to impose salience to the important part of the FOCUS, the two concepts ḥubb 'love' and tanazulaːt 'concessions', suggesting that the function of baʔa is somehow similar to that of a focus accent. The paraphrase in (106b) would clearly also be possible and, to the best of my knowledge does not have any significant difference in meaning.

(105)  
Mar_Shab_FG3_03  
\[ \text{ma}-\text{huwwa} \quad \text{kull} \quad \text{baʔ} \quad \text{dawwad} / \]  
PRT-3SG.M  all/every  SEQ  DEM.SG.M  
'Well, all this...'

(106)  
a.  
Mar_Shab_FG3_03  
\[ \text{ma}-\text{hi} \quad \text{miʃ} \quad \text{ḥikajt} \quad \text{baʔ} \quad \text{ḥubb} \quad \text{wu-tanazul-aːt} \]  
PRT-3SG.F  NET  story  SEQ  love  and-concession.PL  
'Look, but this is not a matter of love and concessions.'

b.  
\[ \text{ma}-\text{hiba} \quad \text{baʔa} \quad \text{miʃ} \quad \text{ḥikajit} \quad \text{ḥubb} \quad \text{wu-tanazul-aːt} \]  
PRT-3SG.F  SEQ  NEG  story  love  and-concession.PL  
'Look, but this is not a matter of love and concessions.'

The final example I will give shows baʔa with high prominence and a leading accent. In this case, baʔa does three things. First, it reactivates a referent that has been introduced at the beginning of a long episode. Second, it marks the referent in-niskafihaːja 'the nescafé' it is associated with as standing in contrast to two other referents (a carrot and an egg) and third it marks it as the final topic of this series to be talked about. The contour is illustrated in Figure 23.

(107)  
Mar_Siqa_FG3_16  
\[ \text{in}-\text{niskaf}\text{-haːja} \quad \text{baʔa} \quad \text{tafaːśal}^{24} \quad \text{maʃ} \quad \text{al-mawqif} / \]  
DEF-nescafé-IN  PRT  interact.3SG.M  with  DEF-situation  
L-LH  L*H  LH*L  H*L  H*L%  
'Now the nescafé, it reacted to the situation.'

---

24 The emphasis of the comment clause is still enhanced by being borrowed from MSA. Note that the verb does not agree with the subject in gender.
In sum, we can say that the sequential marker *baʔa* is used in EA in a highlighting function. It very frequently occurs after topics making them especially prominent, thus indicating topic shift to a final element in a list and contrast. Its syntactic position is quite unrestricted although it usually is prosodically enclitic and therefore not found in initial position. It may, however, occupy syntactic positions in which grammatical and lexical elements usually are not allowed, such as within an annexation or between head and demonstrative determiner. Finally, *baʔa* is not restricted to topics, it may also occur within a FOCUS domain. There it may also occur in *front of* or *after* expressions whose denotata seem to carry the main FOCUS information. For these reasons, it seems to me that *baʔa* is a kind of focussing device on a par with a focus accent.

### 5.6 Postfield constructions: antitopics and topic addition

It has frequently been observed that the right-dislocated periphery of a sentence may host different constituents called *antitopics* and *afterthoughts*. Whereas these are frequently used synonymously (e.g. by Moutaouakil (1989: 124) for Classical Arabic), some researchers make a difference between antitopics that are low in prominence and afterthoughts that add some information to the predication (Ziv & Grosz 1994; Lambrecht 1981, 1994, 2001; Fretheim 1995; Birner & Ward 1996, Birner 1998). As Birner & Ward (1998: 146) write in their survey of information status and non-canonical word order in English, their “examination of naturally occurring data indicates that right-dislocation *not only permits, but in fact requires*, the dislocated NP to represent information that is given in some sense.” (emphasis added). I adopt this assumption in the present study.

Postfield constructions containing a topic are not particularly frequent in the EA text corpus. I counted only 18 instances of topics occurring in postfield positions. These come in two (or three) types, as we will see. Examples (108) and (109) are typical antitopics containing a lexical NP. Antitopics are quite frequent in questions, thereby the FOCUS is preposed and more immediately accessible to the hearer. The high rise on the antitopic in (108) is due to question intonation. Note that (108) has two topic
referents, *il-lišba*, the game' and *ism-aha* ‘its (lit. her) name’, the second one being articulated in initial position and the second one in final position.

(108)  Sihr_FM3_11

<table>
<thead>
<tr>
<th><em>ism-aha</em></th>
<th><em>il-lišba</em></th>
</tr>
</thead>
<tbody>
<tr>
<td>name-3SG.F what DEF-game</td>
<td></td>
</tr>
<tr>
<td>L*-</td>
<td>*L-</td>
</tr>
<tr>
<td>'What's it called, this game?'</td>
<td></td>
</tr>
</tbody>
</table>

Example (109) is a typical case of antitopic with the topic expression deaccented because it is taken for granted, being the discourse topic and having been the sentence topic of the preceding sentence.

(109)  Mar_Shah_MM3_02

<table>
<thead>
<tr>
<th><em>ʔasa:s-u</em></th>
<th>/ <em>s-siqu</em></th>
<th><em>fi-n-nafs</em></th>
<th>/</th>
</tr>
</thead>
<tbody>
<tr>
<td>basis-3SG.M DEF-confidence in-DEF-self</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LHL%</td>
<td>L-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>'It's its basis, self-confidence.'</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

In the next example, the topic is referred to by a personal pronoun in TOPIC position and then picked up by means of the full proper noun in the tail (110). The example is illustrated in Figure 24 below.

(110)  Ihsan_FG1_05

<table>
<thead>
<tr>
<th><em>fa:..</em></th>
<th>/ <em>huwa</em></th>
<th><em>ziːil</em></th>
<th><em>minn-u</em></th>
<th><em>abd-in-nasːir</em></th>
</tr>
</thead>
<tbody>
<tr>
<td>so...</td>
<td>3SG.M becoming.3SG.M from-3SG.M AN</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L-</td>
<td>LH*-</td>
<td>^L^H HL</td>
<td>HL-</td>
<td></td>
</tr>
<tr>
<td>'So, he was angry with him, Abdel Nasser.'</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

It has been pointed out in the literature that topics in right-dislocated position are generally more accessible than left-dislocated topics or that they even have to be already at the centre of attention or 'in focus' (Gundel et al. 1993). We have seen in the above sections that left-dislocated topics may be activated in that position. Lambrecht (2001: 1074) explains this difference in terms of a universal *iconic ordering principle*, according to which TOPIC and postfield positions correlate "with the relative pragmatic salience of the topic referent at utterance time." He notes that the TC-order is used to signal the "announcement or establishment of a new topic relation between a referent and a predication", whereas the reverse order "signals continuation or maintenance of an already established relation." (ibid.; emphasis in the orig.). These remarks are particularly important, as they make clear that it is not only the information or cognitive status of a referent that counts, but rather his expected pragmatic relation to the proposition at hand. Such information is included in the *Givenness Hierarchy* as proposed by Gundel et al. (1993) (see Section 5.7 for more details).

The above observations seem to be true for EA as well. In most cases, the antitopic is a repetition of an expression from the preceding clause. In (108), the referent of *il-lišba* is also active, as the whole episode deals with the explanation of the 'game'. However, only in (109) the topic referent was already
the topic of the preceding clause, whereas in (110), *Abdel Nasser* had been in the FOCUS of the preceding proposition. However, his occurrence as a topic of the proposition at hand is highly expected, indicating his reaction to what has been said in the preceding sentence. In (111), *il-xufuna* 'arthrosis' is the topic of the whole conversation. We are told that people immediately think they are suffering from arthrosis whenever their bones and joints ache. The present utterance is the continuation of a preceding sentence stating that not every pain in the knee joint should be interpreted as arthrosis. There may be many reasons for pains in the knee joint (and arthrosis is the most commonly occurring one). Arthrosis was also mentioned in the preceding clause, although in FOCAL position.

\[(111)\text{Knee_MG1_19}\]
\[
\begin{array}{ll}
\text{ʔaktar-ha} & \text{fiжу-ш-an} \\
\text{most-3SG.M} & \text{distribution-ACC.INDF} \\
\text{LH^L} & \text{DEF-arthrosis} \\
\end{array}
\]

'It is the most common of these, arthrosis.'

Pronouns may also occur as antitopics. Pronouns are inherently given and thus their use as antitopics is also in line with the claim that antitopics have to be active in discourse. The following example exhibits a double antitopic. The speaker has told a joke and now makes the remark that his interlocutor does not seem to have known this joke. The reason for this construction is also stylistic, confer the parallel construction of the two consecutive sentences.

\[(112)\text{Mar_HF_MM4_02}\]
\[
\begin{array}{ll}
\text{ma-ʃirif-ti-ha:-f} & \text{inti di} \\
\text{NEG-know-2SG.F-3SG.F-NEG} & \text{2SG.F DEM.SG.F} \\
\text{LHL-} & / \\
\end{array}
\]

'You didn't know it, this one.' (lit. 'You didn't know her, you, this one.')

\[
\begin{array}{ll}
\text{ʔul-tə-ha:-l-ik} & \text{ana} \\
\text{tell-1SG-3SG.F-to-2SG.F} & \text{1SG} \\
\text{LHL-} & / \\
\end{array}
\]

'Now I have told it to you.' (lit. 'I have told it to you, me. ')

I suspect that the reason for an antitopic construction may be to prepose the FOCUS and thereby probably also to diminish the aboutness relation, contrary to a hanging TOPIC that arguably enhances the aboutness relation between the topic and the comment. It has been suggested that antitopics are adding topic information or correcting a performance error. But as Lambrecht (2001: 1076) convincingly argues, this is not likely to be the explanation for antitopics, which are much too frequent and conventionalized in the languages of the world, e.g. in some Romance languages, such as French (Lambrecht 2001) and Catalan (Vallduví 1992), to be regarded as ad-hoc constructions. One major reason against the correction hypothesis is the typical low prominence of antitopics. The accessibility of referents coded as antitopics is also evident in examples such as (113). Here, the information contained in the independent pronouns is superfluous, so to say, as the person marking is already done within the verb phrase. As far as many of the other examples are concerned, the antitopic is just a
continuous topic and - at least if a verb is present in the predicate - may also be left out without a remarkable loss in pragmatic information.

In the remainder of this section, I will introduce two more topic constructions occupying postfield position that differ from antitopics exactly in this respect. These constructions add information and given that their information value is rather high, at least in the second type discussed below, their prominence is not low. The first example (113a) combines a TOPIC with a postfield construction. The conversation is about cinemas in the old days. *It-tazkara 'the ticket', is a topic expression with a low, but prominent contour. This construction seems to be conventionalized for the expression of a certain whole-part or type-token relationship (cf. also 113b). An abstract representation can be formalized as follows: $\text{TOPIC} + \text{predicate} + \text{partitive topic}$.

(113)

a. *Taxi_08_04*

$\begin{array}{llllll}
\text{kan} & \text{ki} & \text{bir-ha} & \text{sit}^{\text{ח}} & \text{a}: & \text{far}^{\text{ח}} & \text{?rf} & \text{wu-nus}^{\text{ח}} & \text{s}^{\text{ח}} & \text{it-tazk}^{\text{ח}} & \text{ra}.
\end{array}$

be.3SG.M big-3SG.F 16 piaster and-half DEF-ticket

'The most of expensive of them [the cinemas] were 16 piasters and a half per ticket.'

b. *constructed example*  
(context: *bi-ka:m il-lahma di* 'How much is this meat?')

$\begin{array}{llllll}
\text{il-lahma} & \text{di} & \text{bi-talati}^{\text{ח}} & \text{h} & \text{gine:} & \text{h} & \text{il-ki}^{\text{ח}} \text{l}^{\text{ח}}
\end{array}$

DEF-meat DEM.SG.F PREP-30 pound DEF-kilo

'This meat costs 30 pounds per kilo.'

The second construction was found in the narratives, but variants of it may also occur in dialogue. It is, however, a very salient feature of the narratives that is used by all speakers. The topic that is named at the beginning of the sentence, as usual, is then amended in the postfield. This additional information is made highly prominent, usually by focus accents of the closing type, with the final accent exhibiting either a leading accent as in (114) or a leading-linking accent for simple continuation as shown in Figure (24b) below. I will call this type of construction *topic addition*. Despite the use of highly prominent closing accents, topic addition must not be mistaken for the FOCUS, not being the 'new' information to be conveyed. Its prosody also differs from FOCUS prosody. To detect the difference we have to pay attention to the whole tune, i.e. the trendline of the whole contour. Contrary to FOCAL contours, the contour associated with topic additions is linking (Figure 24b). Pragmatically, this addition does not constitute propositional information on its own, it receives its informational character only from its status as standing in a topic relation to the proposition. The following (114) is a typical example of that construction. The construction may be represented abstractly as (S) Pred (S) / anaphoric pronoun + NP$_1$ (+ NP$_2$...).

(114) *Mido_M1_01_01*

$\begin{array}{llllllllll}
\text{mi}^{\text{ח}} & \text{du} & \text{daxal} & \text{na}^{\text{ח}} & \text{m} & \text{huwa} & \text{wi-k-} & \text{kalb}^{\text{ח}} & \text{bta}^{\text{ח}} & \text{u}
\end{array}$

enter.3SG.M sleep.3SG.M 3SG.M and-DEF-dog POSS-3SG.M

LHL H L$^\text{H}$° HL HL $^\text{H}$°H$^\text{H}$°

'Mido went to sleep, he and his dog.'
This construction is more like what has been called an afterthought. But again, the explanation cannot be a repair mechanism used to add material that has been forgotten. The construction is completely conventionalized and must have another explanation. One possible candidate is heaviness. It is common knowledge that phonologically heavy constituents tend to come last. However, TOPICS can be very long in EA. An example of a TOPIC containing four individual entity topics in parallel is the following (115).

(115)  Mido_F0_01_08
wu-baʃde:n / ba:ba / wu-ma:ma / wu-mi:du / wu-ʔuxt-u / ra:h-u...
s-othen / and-mummy / and-M / and-sister-3SG.M / go.3-PL

'And then, daddy and mummy and Mido and his sister went....'

The explanation might be found in cognitive processing constraints and a preference for discourse topic continuity. It is probably conceptually easier to process one topic, especially if the topic is the one expected from the preceding clauses and add additional participants that are also topics of the same predication in an additional separate constituent in the tail. Evidence for this assumption comes from the occurrences of the two-sided constructions in the narratives. In these cases it is always Mido, the main character of the frog stories who is coded as a primary topic in the left periphery of the sentence. There is also one example of direct speech and it is again Mido who is speaking and saying ana haʔab maʕa:, ana wi-k-kalbə btaːʃi I will play with him [the frog], me and my dog.' Contrary to that, it would have been rather odd if the example in (115) had been articulated as (116) in the given context, because baːba ‘daddy’ is not a continued topic in this case, and probably also because Mido is the primary discourse topic.

(116) constructed example
and-then / he / and-mummy / and-M / and-sister-3SG.M

'And then daddy went..., he and mummy and Mido and his sister.'

However, the whole issue of antitopics and tail constructions clearly needs further investigation, as the examples in the corpus are not enough to draw decisive conclusions. In sum, the postfield constructions that involve topics may be classified into two clearly distinct categories: an antitopic construction with a highly accessible referent and topic addition adding information to the topic that cannot be taken for granted. In these cases only one of the topic referents is mentioned sentence-initially - normally the most salient of the referents (the discourse topic) and the others are added in postfield position. The third construction that was identified is an in-between-case involving a sentence-initial topic whose referent is a certain type of something or a hyper-concept and an antitopic in tail position that is inferable from the primary topic and represents a token or a sub-concept of the first mentioned entity. Such 'antitopics' are somewhat more prominent prosodically than typical antitopics as they clearly add information, even though this information is inferable. Their prominence status as halfway between the low prominence of antitopics and the high prominence of the topic
addition is again evidence for the assumption made here that prosodic prominence varies according to the information value, in our terms the **pragmatic properties**, of the constituents under concern. Figure 24 illustrates a typical *antitopic* (24a) and a typical *topic addition* construction (24b).

The cross-classification shows the preponderance of the leading-linking accent (50% of the cases), followed by the linking type. Among the interesting aspects of the count is the behaviour of pronouns. Although in both categories, the leading-linking accent is the most common one (almost 50%), we also find a fair number of linking tones on pronouns (37% for subjects, 27% for non-subjects). Pronouns are given by definition and low prominence is thus an expected feature. The quantitative results thus support what has been said about pronouns in Section 5.3.1.5, namely that the prosodic shape of the pronoun correlates with its function as a continued topic expression or a contrastive or at least shifted topic.

The cross-classification shows the preponderance of the leading-linking accent (50% of the cases), followed by the linking type. Among the interesting aspects of the count is the behaviour of pronouns. Although in both categories, the leading-linking accent is the most common one (almost 50%), we also find a fair number of linking tones on pronouns (37% for subjects, 27% for non-subjects). Pronouns are given by definition and low prominence is thus an expected feature. The quantitative results thus support what has been said about pronouns in Section 5.3.1.5, namely that the prosodic shape of the pronoun correlates with its function as a continued topic expression or a contrastive or at least shifted topic.

![Pitch and intensity tracks](image)

Fig. 24: Pitch and intensity tracks of an antitopic (panel a) and a case of topic addition (panel b).

### 5.7 Topics and pragmatic properties

In Chapter 1, the concept of *pragmatic properties* was introduced. It was suggested that pragmatic properties are of two types, pertaining to the cognitive status of referents in the interlocutors’ mind, mostly referred to as *cognitive status, information status* or *givenness* in the literature, and the property
of background vs. foreground, which is essentially the way a speaker perceives the salience of the individual concepts. On such a view, cognitive status is not directly reflected linguistically, but filtered through the mind of the speaker who chooses to reduce or increase the prominence of individual concepts or pieces of information according to his own perspective. Nevertheless, cognitive status has a strong bearing on the articulation of these concepts. Much of what has been said about this topic and could also be said about it in EA, is related to the morphosyntactic coding of referents. Again, I will have to leave that issue aside here and concentrate on prosodic realization. As is well known, the prototypical topic is given, or at least accessible (cf. Chapter 1). Moreover, topics are also frequently continued. Consequently, the typical topic expression is either a null-anaphor or an unstressed pronoun (Givón 1983; Lambrecht 1994; inter alia). EA is no exception here. As we have seen in 5.2 above, most of the verbal predicates have no explicit subjects, i.e. the topic is coded in the inflectional ending of the verb, which is the least prominent type of topic articulation possible. But we have also seen that independent pronouns are quite frequently used as topics in EA, in which case they are mostly accentless (see Table 6 below).

We have already seen in 4.3.1 that lexical items may also be downtoned, although rarely completely deaccented, whether the expressions occur in a TOPIC or in a FOCUS domain. Downtoning within a FOCUS will not interest us here, however confer the discussion in 4.3.1 and example (68a) above where walditu 'his mother' is a repetition from the preceding clause and the repeated word may even be regarded as deaccented.

In 4.3.1, we have also seen that downtoning of TOPICS, even if they are long, is indeed possible in EA, when the TOPIC already has the highest degree of salience in the interlocutors' consciousness, which is the case if it is a repetition from the preceding question. Example (117), illustrated in Figure (25) shows such an answer that fully repeats the question. Such utterances are typical for experimental data, but occur much less frequently in spontaneous speech. The reason for that is simply that in experimental data, the speakers are requested to repeat whole sentences, whereas in spontaneous dialogue there is no need to articulate uninformative chunks of speech.

(117) 43ARZ-D03FCT-04M1200
   (context: Is the boy holding the man's hand?)
   laʔ [il]-walad ma$:ik ḥ:i:d il-bint
   no [DEF]-boy hold.PTCP.SG.M hand DEF-girl
   'No, the boy is holding the girl's hand.'
Repetition, however, does not inevitably lead to downtoning as the following example (118), illustrated in Figure 26 shows. The example is from a TV discussion, where the host has presented a certain detail about the issue under discussion and is now asking of the discussant to comment about it.

(118)  Mehwar_FG2_01_03_02 I introduced

\(\text{ta}\tilde{\text{a}}\text{qi:b-ik }\tilde{\text{e}}?\) comment-2SG.F what
LH L/HL%
'What is your comment?'

\(\text{ta}\tilde{\text{a}}\text{qi:b-i} \tilde{\text{t}}\tilde{\text{i}}\text{-t'ammin} \)
comment-1SG PASS-calm.IMP.SG.M
LH LHL%
'My comment is: Don't worry!'

Here the repeated item is associated with a leading-linking contour, i.e. a default accent, which raises the expectations of the hearer and shows more involvement and more eagerness to answer the question than a low-toned contour. If the speaker had used a low flat linking contour, she would have downplayed the question adding to it a notion of taken-for-grantedness. Such examples again show that there can never be a one-to-one correspondence between an information structural category and prosodic coding, because of the multifunctional nature of prosody, particularly its extralinguistic functions.

A frequent type of topic expression is a noun denoting a given or accessible referent followed by a demonstrative. The topic may either be a verbatim repetition of an item in the preceding clause (119) -
mostly contained in the FOCUS, but it may also be referring to some information given in the clause before as in (120), where the topic expression /-ʃakwa 'complaint' sums up the the information given in the preceding clause. The speaker tells us about girls complaining to her on her facebook site about their partners.

(119) Knee_MG1_09
[ind in this surface] cracks develop by time...

and these cracks...

(120) Mar_Shab_FG3_02
and all that complaint....

Such TOPICS are frequently leading TOPICS, very frequently accompanied by a phrase boundary, but again a phrase boundary is not obligatory as illustrated by (122). Leading TOPICS, whether with or without the demonstrative, are also used for activation, i.e. they are associated with an accessible referent which is reactivated by the prominent leading accent as in example (46) above, where 'overhead projector' is accessible, i.e. it has been mentioned before, but not within the last five IPs.

Although, the final accent in this example is leading, the whole contour associated with the TOPIC is a linking contour. This suggests that it is not necessary to invoke a focus accent for activation of accessible items. Contrary to that, new referents that are introduced by a TOPIC construction rather than a FOCUS construction necessarily involve some higher prominence as will be seen in 5.7.1 below. The following example exhibits another - more prominent - type of leading accent, a fall-rise, which is much less frequent in EA than the simple rise. I have the intuition that L*H’ is used here to indicate something like 'coming back to our main issue of concern...', in the sense of Gussenhoven's (1983) characterization of the fall-rise as indicating "selection of a Variable from the Background".

The context of this utterance is a question related to the main issue of 'self-confidence', but not directly about it. One of the hosts is giving a longer explanation and expecting Marwa's comment about it. The latter, now relates what has been said before to the main theme of the discussion.

(121) Mar_Shab_FG3_06
This issue of [self-]confidence - I heard you [when you were talking before ...]
TOPICS involving the repetition of lexical elements are a kind of genre-specific feature of narrative texts. A complex proposition is divided into two and more chunks forming a logical sequence, in line with what Chafe (1987, 1994) calls the one new idea constraint. First a concept is introduced as a FOCUS and taken up as a TOPIC in the next sentence as in (122).

(122) Goha_F0_01

\[
\begin{align*}
\text{fi-marr} & \quad \text{min-il-marr-a:t} / \quad \text{kan} & \quad \text{fi:} & \quad \text{ra:gil} & \quad \text{ism-u} & \quad \text{guha} \\
\text{LH*} & \quad \text{LLH} & \quad \text{L} & \quad \text{H} & \quad \text{G} & \quad \text{L*-}
\end{align*}
\]

'Once upon a time, there was a man called Goha.'

In the next example (123), the whole contour is leading with a leading final accent, introducing a new topic - the vendors. This new referent is anchored in the discourse by the deictic reference to Cairo which was mentioned in the preceding clause, using the locative adverbial hina:k 'there' (Figure 27).

(123) Goha_F0_03

\[
\begin{align*}
\text{il-bajja9i:n} & \quad \text{hina:k} / \quad \text{bi-i-yall-u} & \quad \text{kulla} & \quad \text{ha:ga} \\
\text{DEF-vendor-PL} & \quad \text{there} & \quad \text{IND-3M-make.expensive-PL} & \quad \text{all/every} & \quad \text{thing} \\
\text{LH} & \quad \text{L} & \quad \text{L} & \quad \text{HL} & \quad \text{HL} & \quad \text{HL%}
\end{align*}
\]

'The vendors there sell everything expensive.'

Fig. 27: Pitch and intensity tracks of leading TOPIC contour introducing new topic.

The final example of this section shows the repetition of a clausal TOPIC containing the topic referent ħurrijja 'freedom', which is already given because it has been the discourse topic of the whole paragraph. The analysis of this example is not without complications. Looking at the context shows that everything that is being said in (124a) is in fact given except the idea that 'freedom comes all of a sudden' and that it is 'only natural' that freedom should turn into chaos under these circumstances. This would suggest that the dependent clause after the topical ħurrijja 'freedom' should be classified as FOCUS belonging to the first clause. And indeed, I believe that this is a possible analysis, i.e. coming all of a sudden is the FOCUS related to the topic ħurrijja. The leading accent with a boundary tone is also a typical FOCUS accent as we have seen thus far. But this cannot not be the whole story. It has to be noted that this FOCUS has a weird syntactic coding for a FOCUS constituent. But if we allow for
embedded information structural categories, this fact is not problematic for the analysis. We can analyse the whole temporal phrase including the left dislocated entity topic as the TOPIC for the following FOCUS, the predicate phrase *tˤabiːʕi* associated with a very prominent closing accent (a pointed hat) and the COMP clause which has a topic relationship to the FOCUS as an antitopic, whose prominence is reduced (verb and object a integrated in a one non-prominent falling contour). The other option is to say that the dependent clause is a TOPIC containing a focus of interest on 'sudden' because of the unexpectedness and newsworthiness of the information, but that the whole information within the TOPIC is being presupposed from the start and the only FOCUS domain in this complex sentence is the one on the predicate of the second clause. The difference between the two analyses is essentially only whether the long TOPIC contains a FOCUS or not. However, this issue is not important for the argument at hand.

(124) Mar_HF_FG3_04

(Context: MM4: Now, this freedom, can it be abused? - FG3: Anything can be abused, anything. MM4: Ok, so you... do you think that this freedom may turn to chaos? FG3: It's a very thin line, a thin line. And I believe (that)...

a. clausal TOPIC - first occurrence

\[ (inn) \text{il-ħurrijja} \text{lamma} b-tiː-gi fagʔa / \]

COMP DEF-freedom when IND-3SG.F-come sudden

\[ ^L \text{LH} \text{L}^\text{^H^L} \text{H}^\text{^H} \]

..(that) freedom, when it comes all of a sudden...

\[ tˤabiːʕi \text{ nna-ha tiʔilib fawdʔa } / \]

COMP DEF-freedom NEG-someone-NEG IND-3SG.M-give-3SG.F

\[ ^L \text{LH} \text{L}^\text{H} \]

it's only natural, that it will turn to chaos.'

b. clausal TOPIC - repetition

\[ [i][l]-ħurrijja [lam]na b-tiː-gi fagʔa / \]

DEF-freedom when IND-3SG.F-come sudden

\[ ^L \text{H} \text{^L} \text{H}^\text{^H} \]

'Freedom, when it comes all of a sudden,

\[ ha-ja-saːʔ istiylal-ha tˤabʕan \]

FUT-3SG.M-abuse.PASS use-3SG.M naturally

\[ ^L \text{L}^\text{^H} \text{L} \]

it will be abused, of course.'

Now let us return to the original motivation for citing this example. Looking at the tonal annotation and Figures (28a) and (28b) below, we see that the first instance of the clausal TOPIC is coded by a leading contour. The prominence associated with the leading type corresponds to the newness of the information, at the same time the starting point character of the clausal TOPIC is enhanced. In the repetition in Figure 28b, on the other hand, we see that the TOPIC has a linking contour altogether. In both cases, the FOCUS starts with a prominent focus accent, but only in the second example the FOCUS is more prominent than the TOPIC.
5.7.1 Presentational TOPICS, topic activation, contrast and ... emphasis

In the preceding sections we have mostly looked at TOPIC types that were either characterized by a linking or a leading contour. This is what we expect of the prosodic realization of a TOPIC. There are, however, a number of cases in which the topic expression is associated with a closing contour indicating assertion rather than aboutness. In Chapter 4.3.1 we have seen that one type of focus accent is the closing accent, mostly involving a full rise-fall. Elements that are associated with such focus accents are highlighted because of the special interest the speaker has in them. As the closing accent has an assertive character it may be used to introduce new referents or to highlight concepts that are not part of a structured proposition, but precede them as unlinked TOPICS (5.3.1.4). I will call such TOPICS presentational; examples are given in (125) and (126). Concerning the second type of these TOPICS, Chafe (1976: 52) notes such words in isolation may receive "the intonation of a complete sentence", which suggests that they may also be analysed as a FOCUS. It is, however more difficult to analyse an ordinary subject topic in a canonical SP sentence as a FOCUS. In an investigation of a corpus of read in British English, Wichmann (2000) makes the following observation about the functions of what she calls the "citation contour" on topics\(^{25}\), with is equally appropriate for EA.

From the above examples we can see how the citation contour can be used strategically to highlight a topic shift. The contour combines the extra high onset, a high pitch reset typical of topic / paragraph beginnings, with a falling tone, often downstepped, i.e. beginning relatively low in the speaker's range, and falling to a low endpoint. This applies, as we have seen, to titles and to metatextual elements such as story openers, but also to integrated sentence constituents which have no discourse-organising role in themselves. (Wichmann 2000: 38)

Where the EA data differ from Wichmann's description of the British contour is the exact tonal shape and alignment of the fall. In EA, the contour is mostly a full rise-fall, a pointed hat. But it rarely has a high onset (cf. (128) below for one such example) and the accented syllable is typically associated

---

\(^{25}\) Wichmann, however, refers to discourse topics rather than sentence topics.
with high pitch. The topic shift function mentioned by Wichmann is also a prominent function of this accent in EA. In addition, or rather in combination with the topic shift, the contour also may indicate contrast as in (127).

The first example is from one of the talk shows with Marwa Rakha. Here Marwa suggests a topic which, in the sense of Chafe's *premature subjects* (Chafe 1976: 51f.) is referred to by a topic expression in the following clause. The referent is *brandnew* and only barely inferable from the context, which is about problems between the sexes among young people in Egypt. An analysis of this example as a FOCUS also has a lot to recommend it.

(125) Mar_Shab_FG3_04

```
il-mafa:kil  bita:ʕ-it  il-gawa:z  il-surf-i /  
DEF-problem.PL  POSS-3SG.F  DEF-marriage  DEF-custom-NISB  
LHL-  LH   L!HL%
```

The problems of common-law marriage. 26

```
ma-hijja  di  bardˤu  min  mawdˤu;ʕ  is-siga  bi-n-nafs
DM-3SG.F  DEM.SG.F also from issue DEF-confidence PREP-self  
... these also pertain to the issue of self-confidence.
```

The second example (126), however, is less likely to be a FOCUS. Here, the speaker introduces a new, but inferable topic *is-sa:ħir* ‘the magician’ in an interview with a magician. The host asks his guest what being a magician is all about.

(126) Sihr_FM3_03

```
is-sa:ħir  huwwa  min
DEF-magician  3SG.M who
LHL  !H  L*.
```

‘What is a magician like?’
(lit. ‘a magician, who is he?’)

In the following example, the topic shifts from one given referent to another and then to a third one. The topics in this example are clearly contrastive.

(127) Mido_F0_02_05

```
ba:ba  kan  la:bis  badla  gami:l-a  ʔawi /  wi-ma:ma
Daddy  be.3SG.M  wear.PTCP.SG.M suit  beautiful.SG.F very and-mummy
LHL-  
```

‘Daddy was wearing a very beautiful suit, and mummy...’

```
ka[ni]-t  labs-a  fustə:n  gami:l  gami:l  ʔawi /  wu-ʔuxt-u
be-3SG.F  wear.PTCP.SG.F dress  beautiful beautiful very and-sister-3SG.M
LHL-
```

was wearing a very very beautiful dress, and his sister...

---

26 Common-law marriage has become a real social problem in Egypt. Because young people often lack the money for getting married, it has become common practice in some social groups to 'legalize' a relationship by a 'common-law marriage' which only needs two witnesses. However, this type of marriage does not offer security to the girl and her offspring, if she gets pregnant, because it has happened that the boy simply denied that marriage. Common-law marriage is also not respected socially.
However, closing accents with simple topic shift are also common. It is interesting that most topic shifts expressed by a closing contour occur in the frog stories. The narratives all involve a limited number of animate referents that are the main characters of the story. So the topic shift between them seems to have something inherently contrastive, as these referents are forming a closed set of alternatives among which the speaker chooses his topic. Topic shifts between inanimate referents also occur, but these also pertain to situations where different items are being predicated about, such as a glass, a plate, and a fork. Other scenes involve the frog, the waiter and some guests, or in another story bees, a dog, and a boy. Brunetti et al. (2011), in a study investigating marked TOPIC constructions in three Romance languages, English and German, also report that particularly one of the frog stories ('Frog goes to dinner') which I also used for my data collection, provides a story that induces a great number of topic shifts. And indeed, many of the closing accents on topics occurred in this story in the EA text corpus.

In addition, a closing accent may also be associated with a shift in discourse topic or the beginning of a new episode. As we have seen in the quotation from Wichmann (2000: 38), new topics involve higher pitch at the beginning of the IP. This can also be observed in the EA data. In the following example (128), the sentence topic does not even change, but a new episode begins. The first sentence is some background information and then the scene changes with the second sentence, returning to the main story line. The speaker uses a pitch reset to indicate that change. The first L also has a relatively high frequency, a low L would have further enhanced the prominence of the accent. Figure 29 illustrates the pitch reset at the beginning of the respective IPs.

(128)  Mido_M1_01_02

\[
\begin{align*}
\text{mi:du} & \quad \text{kan} & \quad \text{id-dufdˤiʕ} & \quad \text{fi-kubba:ja} & / \\
\text{M} & \quad \text{be.3SG.M} & \quad \text{put.PTCP.SG.M} & \quad \text{DEF-frog} & \quad \text{in-glass} \\
\text{^L^HL} & \quad \text{L} & \quad \text{HL}- & \quad \text{LH!H-} & \quad \text{H^L%} \\
\end{align*}
\]

'Mido had put the frog into a glass jar.

\[
\begin{align*}
\text{mi:du} & \quad \text{daxal} & \quad \text{naːm} \\
\text{M} & \quad \text{enter.3SG.M} & \quad \text{sleep.3SG.M} \\
\text{^L^HL} & \quad \text{HL} & \quad \text{H}^-
\end{align*}
\]

Mido went to sleep.'
Fig. 29: Pitch track showing a pitch reset between phrases to indicate the beginning of a new episode.

The last function of the closing accent - not only in TOPICS, but even more so in FOCI - simply is to emphasize the whole proposition or to ascertain that the event described has really happened or is really happening, as in the following example. I previously analysed the first sentence as a thetic utterance. But I now believe that this analysis is wrong. Rather, this is just a case of the narrative VS structure discussed above in Section 5.2. The example is illustrated in Figure 30.

\[(129)\] Goha_F0_07

\[
\begin{array}{ccc}
\text{wi-nizil} & guha & q-qahira \\
\text{and-go.down.3SG.M} & \text{G} & \text{DEF-Cairo} \\
\text{LHL} & \text{HL} & \text{LH}^L \\
& 'And Goha went to Cairo...'
\end{array}
\]

\[
\begin{array}{ccc}
\text{wu-ʔaṣad} & ʔala & ʔahwa \\
\text{and-sit.3SG.M} & \text{on} & \text{coffee(shop)} \from \text{DEF-coffeeshop.PL} \\
\text{HL} & \text{LHL} & \text{!HL} \\
& 'and sat down in one of the coffeeshops.'
\end{array}
\]

Fig. 30: Pitch track of two assertive propositions from a narrative.

5.7.2 Some figures

In the EA text corpus, every denotation was coded for one of the three information status categories: given, accessible or new (see Section 4.2 for more details). Participants of the situation were always
coded as given and other referents were coded as given if they had been mentioned within the preceding five IPs. The first count that was undertaken, however, looks at the distribution of all given, accessible and new denotations that occurred in a TOPIC or a FOCUS domain throughout the EA text corpus (Table 6).

<table>
<thead>
<tr>
<th></th>
<th>TOP</th>
<th>FOC</th>
<th>TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>GIV</td>
<td>1091</td>
<td>1535</td>
<td>2626</td>
</tr>
<tr>
<td>ACC</td>
<td>565</td>
<td>1894</td>
<td>2459</td>
</tr>
<tr>
<td>NEW</td>
<td>139</td>
<td>1637</td>
<td>1776</td>
</tr>
<tr>
<td>TOTAL</td>
<td>1795</td>
<td>5066</td>
<td>6861</td>
</tr>
</tbody>
</table>

Tab. 6: Given, accessible and new referents in the EA text corpus.

The numbers in Table 6 are as expected. We see that from a total amount of 6861 denotations that were coded for information status, three quarters occur in FOCUS domains and one quarter in TOPIC domains (including frame TOPICS). Again three quarters of all denotations were annotated as 'given' or 'accessible' and only one quarter as 'new'. Also as expected, given referents or denotations were especially frequent in TOPIC position, even though the absolute numbers of given items are, of course, still higher in FOCI. Accessible denotations behave according to the overall distribution (23% in TOPICS and 77% in FOCI). It also comes as no surprise that new referents were overwhelmingly introduced in FOCI (92%).

The next Tables (7) and (8) illustrate the distribution of the accent types over the three information status categories from two different perspectives. This complex search could only be conducted on a small subset of the data, which explains the discrepancy between the figures in Tables (7) and (8) and Table (6). For the detailed counts in Tables 7 and 8 only a subset of the topics could be evaluated due to limitations in the search options of ELAN, given the annotation used. That means that only topical entities (almost only subject topics) could be investigated that were coded as TOPICS, not those that occurred within TOPICS, i.e. only words that were co-extensive the TOPIC could be taken into account. Topics that fall into that category are definite and indefinite NPs, personal pronouns and demonstrative pronouns. Unfortunately, relational presupposition TOPICS and all types of longer hanging TOPICS could not be captured by this count.

<table>
<thead>
<tr>
<th></th>
<th>GIV</th>
<th>ACC</th>
<th>NEW</th>
<th>TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>LOW LINK</td>
<td>100</td>
<td>12</td>
<td>0</td>
<td>112</td>
</tr>
<tr>
<td>LEADING-LINKING</td>
<td>241</td>
<td>116</td>
<td>6</td>
<td>363</td>
</tr>
<tr>
<td>LEADING</td>
<td>77</td>
<td>36</td>
<td>0</td>
<td>113</td>
</tr>
<tr>
<td>CLOSING</td>
<td>68</td>
<td>27</td>
<td>3</td>
<td>98</td>
</tr>
<tr>
<td>TOTAL</td>
<td>486</td>
<td>191</td>
<td>9</td>
<td>686</td>
</tr>
</tbody>
</table>

Tab. 7: Distribution of accent types broken by information status categories.
Tab. 8: Distribution of information status categories broken by accent types.

<table>
<thead>
<tr>
<th></th>
<th>LOW LINK</th>
<th>LEADING-LINKING</th>
<th>LEADING</th>
<th>CLOSING</th>
<th>TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>GIV</td>
<td>100</td>
<td>21%</td>
<td>241</td>
<td>50%</td>
<td>77</td>
</tr>
<tr>
<td>ACC</td>
<td>12</td>
<td>6%</td>
<td>116</td>
<td>6%</td>
<td>36</td>
</tr>
<tr>
<td>NEW</td>
<td>0</td>
<td>0%</td>
<td>6</td>
<td>0%</td>
<td>0</td>
</tr>
<tr>
<td>TOTAL</td>
<td>112</td>
<td>16%</td>
<td>363</td>
<td>53%</td>
<td>113</td>
</tr>
</tbody>
</table>

We see that in these 'narrow' TOPICS, the distribution of given/accessible vs. 'new' further shifts towards the given/accessible type and within this category towards the given type. This is in line with the common assumption that topics are usually given. 'New' referents, constituting a vanishingly low proportion of 1% (9 items), are particularly infrequent. Now, let us consider the distribution of accent types across these three categories. It is particularly instructive, although not unexpected, that 89% of low links occur on a given item (Table 7). If we look at the correlation from another perspective, however, only 21% of all given items are associated with a low linking contour (Table 8). The clearly largest group is the leading-linking contour (frequently a default accent) with more than 50% of the total amount of occurrences, supporting its default character. It is the most frequent accent in all three information statuses. New referents either have a leading-linking accent or - as expected - a closing accent, whereas neither low linking nor leading accents occurred. The surprising result of this count is the relatively high number of closing accents on given items (69% of all closing accents). Although closing accents on a whole were not very frequent (14%), the high amount of assertive accents on a given word is puzzling to a certain extent. We will have to look at these occurrences in more detail. If we compare these results with the ones listed in Table 3 above, we see that closing contours occur equally often on pronouns and on lexical subjects. As I suggested in 5.3.1.5, a closing accent on an independent pronoun may indicate that this pronoun denotes a special type of reference point, but subject pronouns with closing accents do occur. As pronouns are inherently given, this explains almost half of all occurrences of closing accents on given items. The second group associated with closing accents are lexical subject NPs. This rather unexpected result is probably due to a serious drawback of the type of annotation used in the corpus, namely the fact that topic shifts cannot be accounted for. A distinction between the cognitive status of referents that were continued topics (Givón 1983) or at least explicitly mentioned in the preceding clause and those that have occurred earlier in discourse would most probably have been better suited to test predictions about the occurrence of strong and weak types of prosody, i.e. leading and closing accents on the one hand and linking tones on the other. Such a framework is provided by the Givenness Hierarchy (Gundel et al. 1993).27 As a study by Gundel et

27 The Givenness Hierarchy assumes that the following cognitive statuses in focus > activated > familiar > uniquely identifiable > referential > type identifiable may have an explicit linguistic coding in languages, where in focus necessarily means that the concept has already been in focus in the preceding utterance, i.e. roughly equivalent with being the subject (or
al. (2010) on four different languages, among them Tunisian Arabic studied by Amel Khalfaoui, shows, the two highest categories in the Givenness Hierarchy (in focus > activated) were distinguished in all investigated languages. The results for Tunisian Arabic suggest that null-coding and independent personal pronoun equally require the referent to be in focus, while demonstrative determiners only require activation and a definite NP only requires the referent to be at least uniquely identifiable. As the categories of the Givenness Hierarchy stand in a unidirectional entailment, any form requiring a certain cognitive status may also be used to code all higher cognitive statuses. At present, I am not in a position to verify this for EA. The Givenness Hierarchy was developed to investigate morpholexical devices, but I suspect that it could also be used to make predictions about a preferred prosodic realization. Note, however, that I do not think, that cognitive status on its own can explain the prosodic coding, just as the Givenness Hierarchy does not explain all morpho-lexical codings. However, such a study would be most interesting and must be left for future research.

Now, let us look at the contexts in which the closing accents found on given topics occur. First, 30 of the 51 lexical subjects with a closing accent (see Table 3) occur on a given expression. At closer inspection, it turns out that most of these (23 cases) involve a topic shift, mostly with some measure of contrastiveness. A handful of examples are explicitly contrastive as was shown in example (127). One example is an antitopic with low prominence of the falling accent, another one is a postverbal topic with a very special intonation and a third one is a possible thetic utterances which was (perhaps wrongly) coded as a topic-comment clause. Only two examples cannot be readily explained.

Summing up the small-scale quantitative account of the correlations between prosody and information status, we may note that information status is only indirectly reflected in the tonal coding. The most frequent accent type is the default or leading-linking accent across all information statuses. Personal pronouns, which are referring to given referents by definition, are very frequently accentless, but content words are rarely accentless. It is particularly noteworthy that an expressions referring to a given entity, for example, may occur with all accent types. The same is true for accessible denotations. The only exception seems to be new referents. At least in the small-scale count that was conducted here, the expressions denoting new referents were never associated with low links or leading accents. It has been noted that the crude annotation of 'given-accessible-new' is probably not the best type of annotation to be used. In future research, I would attempt a more fine-grained annotation scheme using Gundel et al.’s (1993) Givenness Hierarchy.
5.8 Quantificational Topics

In this section we will look at quantified phrases and whether they can be analysed as topics of their propositions. In the literature, various constraints on topicality have been proposed, typically concerning the referential properties of the potential topics. As Gundel & Fretheim (2004) note, "[v]irtually the whole range of possible referential givenness conditions on topics has been suggested, including presupposition, familiarity, specificity, referentiality, and focus of attention." Gundel (1976, 1999), for example, claims that only definite quantified phrases and nonspecific indefinites may be topics. According to Gundel (1988), indefinites that cannot be interpreted referentially are excluded from topic position. Evidence for this comes from the fact that they may not be constructed as hanging TOPICS as the following English example shows:

(130) *A window, it's still open. (from Gundel 1988: 213)

As the referent of the window cannot be identified by the hearer, he will also not be able to evaluate the truthfulness of this proposition. A very important grammatical correlate of cognitive identifiability is definiteness although the correlation is "at best an imperfect one" (Lambrecht 1994: 79). However, as Lambrecht (p. 169) notes, the more the subject of a sentence is identified as the topic in a language, the less acceptable sentence-initial indefinite NPs will be. It has been claimed for Arabic that indefinite NPs are banned from sentence-initial position. The following example from Mughazy (2010: 107) seems to support this claim. However, this ban on indefinite subjects is not absolute. Woidich (2006: 182f.) list a number of constructions with an indefinite NP as the subject of an SP clause. Such clauses are especially common in jokes, beginning, for instance, with wa:hid sˤiʔi:di....‘A Saʔdi...’ We will have more to say about such instances at the end of the section.

(131) (= Mughazy's ex. 1)28

(*rˤa:gil) irˤ-rˤa:gil mistanni:-k barˤrˤa
(a man) DEF-man waiting.PTCP.SG.M-2SG.M outside
"(A man) is waiting for you outside."

The exact felicity conditions for topics are still a very much contested issue. The likelihood of an entity to be construed as a topic is expressed in the Topic Acceptability Scale as proposed by Lambrecht (1994: 165), based on Prince's hierarchy of cognitive statuses (Prince 1981). In the following, I will take this scale as a point of departure and try to establish the felicity conditions for EA topics at the end of this section.

---

28 Transcription and glosses are changed to match the conventions used in this book.
We will start with two cases of quantified phrases that are uncontested cases of topics or non-topics. The first one involves a numeral with a definite marker and a demonstrative determiner (133). The TOPIC is referred back to by a subject demonstrative pronoun in the main clause, thus it is a classic case of D-TOPIC. Topic tests with 'as for X' or 'talking about X' also work well.

(133)  
Mar_Siqa_FG3_15

What we may readily exclude from being a topic are non-specific indefinites such as 'nobody', 'anybody', 'somebody'. First, the topicality tests do not work with these (134).

(134)  
*amma  hadd  or:  amma  ma-hadd-of  or:  amma  ?ajja  hadd
TOP  somebody  TOP  NEG-somebody  TOP  any  somebody
'as for someone' or 'as for noone' or 'as for anybody'

Second, these non-specific indefinites are necessarily in the scope of negation when the proposition is negated, which is already indicated by the word for 'nobody' itself. The only possible negation of the positive sentence is (135a), while (135b) is infelicitous.

(135)  
?aj(ja)  hadd  mumkin  ji-t'sallim
any  somebody  possible  3SG.M-learn
'Anybody can learn.'

a.  
 mif  ?aj(ja)  hadd  mumkin  ji-t'sallim
NEG  any  somebody  possible  3SG.M-learn
'Not anybody can learn.'

b.  
*?aj(ja)  hadd  mif  mumkin  ji-t'sallim
any  somebody  NEG  possible  3SG.M-learn

But note that the negated sentence becomes significantly better, if some information about this non-specific person intervenes between the subject constituent and the negated predicate (136).

(136)  

a.  
?aj(ja)  hadd  /  law  iddi-t-l-u  l-furs'a  /  mumkin  ji-t'sallim
any  somebody  if  give-2SG.M-to-3SG.M  DEF-possible  3SG.M-learn
'Anybody, if given the chance, can learn.'
Gundel (1999a) and Gundel & Fretheim (2004) propose an interesting solution for such cases. Gundel claims that left-dislocated constituents of the nonspecific type may also be regarded as topics, given that they are in her terms "relationally given" (my "relationally presupposed"). Consider the following English examples that are originally from Prince (1998).

(137) (from Gundel & Fretheim 2004, ex. 25 and 26)
   a. *Most middle-class Americans*, when they look at the costs plus the benefits, they’re going to be much better off.
   b. *Any company*, if they’re worth 150 million dollars, you don’t need to think of ....

Gundel's solution is that the topic here is only the entity itself, not the quantified entity as a whole. She notes that such phrases "often have a partitive reading" and have "the same presuppositional effects as definite NPs" (Gundel & Fretheim 2004). This reading is paraphrased in (138).

(138) (from Gundel & Fretheim 2004, ex. 25' and 26')
   a. (As for) *Middle-class Americans*, when most of them look at the costs plus the benefits, they’re going to be much better off.
   b. (As for) *Companies*, if any one of them is worth 150 million dollars, you don’t need to think of ....

I will adopt Gundel's proposal and bring further evidence from Egyptian Arabic to support it. One type of evidence comes from universal quantifiers. From a functional point of view, a universally quantified phrase such as the one in (139) is difficult to construe as a topic.

(139) **Knee_MG1_08**

```
beginner kull wa:hid sand-u taxajul mu`ajjan
also all/every one at-3SG.M imagination/idea specific
L°H L°H L°H L°H L°H
L°H L°H
```

'Also everyone have their own ideas.'

Although the construction is formally similar to a hanging TOPIC, having a resumptive pronoun suffixed to the preposition *sand* 'at' referring back to the subject phrase, this alone cannot be regarded as a sufficient condition for the topichood of the referent. I rather believe that the PP here is a grammaticalized possessive construction and should therefore not be regarded as compositional, which in turn means, that the suffixed pronoun does not necessarily fulfil the same function as in D-TOPIC constructions. If we use a verb phrase instead of the PP, the outcome is rather odd ('*kull wa:hid fa:ʃ-u sa:mi* 'everyone, Sami saw them'). But now consider the following sentence (140).
Every mother and every father, when they raise...

On the view I adopt here, these instances of universally quantified phrases are indeed to be regarded as topics. Note that in EA, there are three constructions involving the universal quantifier kull: kull + SG.INDF has the meaning of 'every' as in (142) and kull + PL.DEF means 'all'. The third construction is kull + SG.DEF, meaning 'the whole'. The following examples with the word kita:b 'book' serve to illustrate these constructions: kull kita:b (SG) ('every book'), kull il-kutub (PL.DEF) 'all books' and kull il-kita:b (SG.DEF) 'the whole book'. It is noteworthy that there is no construction kull + PL.INDF *kull kutub to express the meaning of 'all books'. This already points to the fact that the referent of a universally quantified plural is necessarily specific which is consequently marked in EA by the definite marker. Thus, examples like the one in (141) are proper topics.

Now let us return to the above examples. In (142) paraphrases for (140) and (141) are given that show that both cases indeed involve generic topics. Note, that in EA, and in Arabic in general, generic NPs are also marked as definite.

The interesting thing about these paraphrases is that they exemplify exactly the construction that is typically used in EA in such cases. That is, in EA, universally quantified referential NPs, meaning 'all
X’ are very frequently expressed with an R-TOPIC. Examples in the corpus are legion. I will cite two of them, one with an independent pronoun (143) and the other one with lexical NPs as R-TOPICS (144).

(143) Ihsan_FG1_05
 jašni ihna kull-inà fiuf-na;
 DM 1PL all/every-1PL see-1PL:3SG.M
 L- H- °HL L'HL%
 'That means, we all have seen him.'

In the following example, two new referents are presented as topics of the propositions coded as presentational topics - the first IP introduces the first referent, 'girls' problems', and is associated with a closing accent, being completely new to the discourse. The second IP contains the reference to a new referent that is however inferable from the first one, namely 'boys' problems'. The phrase has a leading contour to express continuation, aboutness and parallel contrast with the first expression. These two are followed by a quantifying expression concerning all incoming mails containing young people's problems, which the first and the second referent are part of, and finally all this is referred to by the pronominal suffix -u on kullu 'all of this'. This whole long sequence is coded as a TOPIC of the main proposition containing the predication that all of this is due to [the lack of] self-confidence.

(144) Mar_Shab_FG3_02
 mafà:kl il-bana:t / wu-mafà:kl il-wula:d /
 L^H L*HL LH L°H°
 'The problems of the girls and the problems of the boys...

 kull il[l]ij bi-[jì]-lr-rimi $and-i fi-l-inbox /
 every/all REL IND-3SG.M-PASS-throw at-1SG in-DEF-inbox
 L°H LH °H L°H°H
 everything that is dropped in my inbox...

 kull-u / bi-sabab is-siqa fi-n-nafs
 every/all-3SG.M PREP-reason DEF-confidence in-DEF-self
 H°H° LH L°HL L°H°
 ... all that is a problem of self-confidence.'

In such phrases, prosody indicates the bipartiteness of the proposition and the topichood of the dislocated element. While short dislocated constituents are less likely to be phrased separately as shown in (143), longer constituents are often followed by an IP boundary. The tonal properties are usually rising, frequently of the leading type as in both examples above.

The next point I would like to discuss here, is the existential quantifier construction with fi:, which is usually used to express entity-central thetics. I will, however, venture an analysis of one type of fi:-constructions as a TOPIC construction. Drawing on the same partitive-reading assumption proposed by Gundel, the following sentence may be assumed to be about a certain subset of the hyper-set 'girls'.
In (145), the indefinite NP 'girls' is preceded by the existential quantifier, but intuitively, the proposition is about the girls. In EA, the existential marker is the usual way to say 'some X' or 'several X' as we have already noted in the discussion of example (104). A possible paraphrase would be something like il-bana:t fi: min-hum... 'the girls, among them there are some...'. We will return to that example shortly.

The issue becomes more complicated, when such a partitive paraphrase is not straightforward. (146), like (139) above, involves the indefinite pronoun wa:ħid 'one' which does not even refer to a specific person. But now again consider the function fulfilled by the dependent clause introduced by lamma. Obviously, the information provided in this clause is used to anchor the unidentifiable referent in the interlocutors' universe of. The 'someone' is put in a specific situation and thereby anchored in time and space. Note that this promotes the topic on the Topic Acceptability Scale given in (132). As also pointed out by Lambrecht (1994: 39), some existential constructions do not have the function of asserting the existence of their referents, but rather of presenting them. If we look at the two examples cited here (145, 146), the validity of this observation for these cases is immediately obvious. It is not the existence of girls that is being asserted, but they are introduced to the discourse as new referents about whom a predication is made in a subsequent clause. Thus, the function of the existential quantifier here rather is to present the new referents and to promote them on the Topic Acceptability Scale. In (145) an existential interpretation does not make sense in the first place. An existential interpretation of fi: wa:ħid (EXIST one) would indeed be possible, however with a different prosody. If the main prominence is on fi:, the sentence has the meaning of 'there IS one' (a counter-assertion) and with main the prominence on wa:ħid it means 'there is ONE (= not two or more'). Judging from the English translation here, we note that the meaning, though not the form, is again the same as in the other examples cited here. That is a partitive reading is also possible and the sentence could be paraphrased as in-nas, lamma wa:ħid minhum ... 'the people, when one of them ...'.

(146) Mar_Siqa_FG3_15
fi: wa:ħid / lamma b-ji-t-hat'i‘ə f-mav[aː]qif
EXIST one when IND-3SG.M-PASS-put in-situation.PL
LH ^L°H’ L- LH LH
'Some people, when they are confronted with difficult situations...

30 The speaker here actually uses this pronunciation with a non-velarized /t/ which is typical of young upper-class women.
Additionally, we find other cues that support a topicality analysis for the referents. One striking feature is prosodic. In both cases, the NP referring to the potential topic is associated with a leading-linking contour and an intonation break afterwards. It could, however, be argued that the leading-linking contour is only there to indicate continuation. Another, perhaps more important feature is semantic. We have seen that in both cases the existential clause is not a full proposition, at least not one with the same meaning as in the pragmatically structured proposition expressed by the whole sentence. This fact mitigates a compositional account that could be applied to story openings such as the one in example (147).

(147) Mido_M1_01_01

kan fi: walad ism-u mi:du
be.3SG.M EXIST boy name-3G.M m
'There once was a boy named Mido.' (lit. ... his name [was] Mido)

In this example, the proposition could be interpreted as involving two separate propositions, an assertion of existence 'there was a boy' and a predication about that boy with the topic expressed by the suffix -u 'his' on the noun ism 'name'. However, there is reason to believe that in this example the fi: is also presentative and used for topic promotion (cf. also Lambrech't's analysis of narrative introductions in Lambrecht 1994: 177f.). The third feature is syntactic. In line with the semantic dependency between the different constituents, we can also assume the whole utterances in (145) and (146) to have a higher macrostructure. The first case is the less complex one. Here the comment or the predication is tixi:na .... s'amr'a 'fat ... dark-skinned'. The predicate phrase is elliptic in the sense that it lacks a clause-internal topic expression, such as 'they' or 'you'. The paraphrase of (145) would be something like: '(As for) girls, some of them have been told from the time on when they were little that they are fat or dark-skinned'. Syntactically, this means that the NP bana:t 'girls' is the topic of the proposition and the comment is separated from the TOPIC by a parenthetical circumstantial clause embedded in a temporal adjunct.

The second example (146) involves a topic NP 'one', followed by a dependent clause, specifying that topic 'when he is put in a difficult situation', followed by three appositional clauses (we will come

sˤaʔ-b/a / interview it-raʕafadˤ / qis'siʔ hubb
difficult-SG.F interview PASS-refuse.3SG.M story love-
°H- LH L°HL LH L°H
an interview, where they have been refused - a lovestory...

u-ba:zˤ-ʔit / ðʔajj musˤi:ba w-ħasˤal-it / and-fail-3SG.F any calamity and-happen-3SG.F
L°HL LH L°H L°H°H
that has come to an end ... any calamity that has happened...

b-ji-hʔa rijak-fin-u zajj il-gazˤarˤ-a
IND-3SG.M reaction-3SG.M like DEF-carrot-IN
L°-H° LH° HL° HL%
their reaction is like that of the carrot.'
back to these below), and finally the comment clause with a resumptive pronominal suffix -u 'his' referring back to the topic 'his reaction will be...'. Especially this last example makes it clear that the construction we are confronted with is a marked TOPIC-construction, i.e. a hanging TOPIC, in that specific case an R-TOPIC. Now, returning to the idea of topic promotion as suggested by Lambrecht (1994). Lambrecht's insightful outline of topic promotion identifies two types: the presentational type (the existential constructions) and left-dislocation (as in all types of hanging TOPICS). In the light of the EA data, this is no coincidence. In Lambrecht's framework the first type is not counted as a topic construction, whereas the second type is. However, I believe that the cases discussed thus far show that EA presentative constructions may serve to mark topics, given that a partitive reading is possible. Such an assumption is in line with the clear aboutness semantics of the topics at hand and their formal expression. The upshot of these examples is that indefinites in general and quantified indefinites in particular that are somehow anchored in discourse may well be construed as topics of their propositions. This anchoring makes them to a certain degree identifiable, which in turn means that it is perhaps only the final category of the Topic-Acceptability-Scale that is excluded from topichood.

To conclude this section, I will have a little more to say about indefinite NPs as topics and subjects. We have seen that an indefinite NP may well serve as a topic in EA if it is properly anchored and thus to some degree identifiable. The following example shows such an NP that functions as a free TOPIC or unlinked frame TOPIC in the utterance at hand. The second, third and fourth clause would clearly not be acceptable if they were not anchored by the preceding clause that divides the day into three shifts (cf. the constructed example). This enhances the identifiability of the shift and thereby promotes it on the Topic Acceptability Scale, so that a subsequent reference to one of the shifts by use of an indefinite NP becomes felicitous.

(148) Taxi_23_F0

\[b-a\text{-rassim} \quad i{l}-j{o}:m \quad tala\text{i}t\text{-a\text{t}} / \quad w\text{ar}dij\text{a} \quad b-a\text{-fi\text{a}y\text{a}}}\]
IND-1SG-divide DEF-day 3 shift:PL.F shift IND-1SG-work
'I divide the day into three shifts. One shift, I work...

\[\text{\#a-t-taksi} / \quad \text{wi-war}dij\text{a} \quad u-\text{\#}u\text{d} \quad fi:-ha \quad m\text{a\text{c}} \quad \text{ma\text{ra}\text{-}\text{t}}-\text{a}]\]
on-DEF-taxi and-shift 1SG-sit in-3SG.F with wife-1SG
as a taxi driver, and one shift I stay with my wife...

\[w-\text{\#a}\text{ja\text{\-}\text{a\text{\-}\text{a}}} / \quad \text{wu-war}dij\text{a} \quad u-\text{\#}u\text{d} \quad a-s\text{\#a\text{-}\text{d} \quad fi-n-ni\text{l}}\]
and-child.PL-1SG on-shift 1SG-sit 1SG-fish in-DEF-Nile
children and one shift, I sit fishing in the Nile.'

Without the pretext, such a construction with the indefinite NP would not be interpretable (149).

(149) constructed example

(context: What are your plans for this week?)

\[#\text{war}dij\text{a} \quad b-a\text{-fi\text{a}y\text{a}} \quad \text{\#a-t-taksi} / \]
shift FUT-1SG-work on-DEF-taxi
One shift, I will work as a taxi river,
Wi-wardijja / ha-ʔud fi:-ha maʕ mara:tu:-i w-ʕaja:-l-i / and-shift FUT.1SG-sit in-3SG.F with wife-1SG and-child.PL-1SG
and one shift I will stay with my wife and children

Wu-wardijja u-ʔud a-sˤʔa:d fi-n-ni:l
on-shift FUT.1SG-sit 1SG-fish in-DEF-Nile
and one shift, I will sit fishing in the Nile.'

At the beginning of this chapter, we have seen that marked word order of an indefinite subject plus predicate is possible in EA, although it particularly occurs in a limited number of constructions, such as the joke introductions mentioned above. Woidich (2006: 183) notes that in some SP sentences the indefinite subject may be interpreted as an elliptic existential clause. The EA corpus contains a number of such instances. The problem is only how to interpret them. In the examples from the QUIS corpus, the constructions are used to describe a picture. A typical example is (150). Two interpretations are possible. Either the first IP is interpreted as hosting a full clause with an indefinite subject ‘a man’ on which ‘is carrying a little boy’ is predicated, or it is interpreted as a NP with a non-restrictive relative clause functioning as a FOCUS answering the question ‘What do you see on this picture?’

(150) 35ARZ-C05GNC-04M1200
man carry.PTCP.SG.M boy little and-EXIST bicycle in.front.of-3SG.M
'A man (that) is carrying a little boy, and there is a bicycle in front of him.

I strongly favour the second interpretation. Many of the other utterances of the same picture description task explicitly put the indefinite noun in FOCUS position of a short identificational clause such as in (151).

(151) 35ARZ-B04GNC-04M1200
di ʔasˤriji:t za-rʕ / maħtˤu:tˤ a[a][la]-lʔardˤ
DEM.SG.F pot plant put-SG.F on-DEF-floor
'This is a plant pot, put on the floor.'

The same is true for the next example which is also from the QUIS corpus.

(152) 25_ARZ_A01GNT_00F1300
ʔasˤfu:r-a / tˤajr-a / lon-ha ʔazra?
bird-IN flying.PTCP-SG.F colour-3SG.F blue
LLH^H > LH^H > LH!HL%
'A bird - that/it is flying - its colour is blue.'
Fig. 30: Pitch and intensity tracks of an utterance containing three clauses in three separate IPs.

The example is interesting because of its prosody. Each of the two initial IPs exhibits a marked leading accent. We have already seen that such accents are typical of FOCAL constituents in Chapter 4.5.1, where we suggested that this type of leading accent is a focus accent that typically occurs in FOCUS phrases that are not utterance final. We have also noted that this type of accent is an innovation among the younger generation of Egyptians. Also compare example (52) and the contour depicted in Figure 11 above, where there is no doubt that the three indefinite NPs are FOCAL. As we have seen above, a subject topic is not typically associated with a leading accent, especially not with a focus accent involving a boundary tone. Based on these facts and the general assumption that an unanchored brand-new referent is not a felicitous topic, I consider these cases as constructions involving multiple FOCUS phrases that correlate with an equal number of syntactic phrases.

5.9 Summary and typological implications

In the literature on topicality, a wide range of phenomena has been subsumed under the term ‘topic’. Constraints on topicality have been formulated with regard to referentiality, identifiability and activation state, and to position. Some researchers have limited the occurrence of topics to the left-peripheral position, whereas others assume a topic to be unconstrained with regard to position. A narrow definition of ‘topic’ refers only to left-dislocated constituents, the so-called ‘Chinese-style’ topics (Chafe 1976; Li & Thompson 1976). Early work by Gundel (1976) and Reinhart (1982) has paved the way for a more comprehensive view on topic and its semantic definition in terms of ‘aboutness’ as it is usually held today. While some approaches only rely on aboutness for their functional definition of topic, others make a distinction between this function and the function of frame-setting (Chafe 1976, Jacobs 2001).

In this study, I have suggested two categories that are not mutually exclusive: entity topics and thematic TOPICS, referred to as topic and TOPIC, respectively. This distinction was motivated by the need to account for the formal property of bipartiteness in EA sentences that is prosodically encoded.
It could be shown that the left-peripheral prosodic constituents coincide with semantic constituents that fulfil the function of aboutness or frame-setting and thus cover a wide range of phenomena which have been associated with topic in the literature. On the assumption that a referential expression that serves as an aboutness topic of a proposition can occur in various forms (null-coding, affix, pronominal, lexical NP) and in various positions the unambiguous identification of ‘the’ topic of a sentence becomes difficult, as a sentence may contain more than one topic expression. Moreover, such a definition does not allow for the identification of a stable form-function correlation in terms of prosody. A thematic TOPIC is restricted to the left-peripheral domain of a sentence and its most common prosodic shape is that of a leading or a leading linking contour. A common instance of TOPIC consists of only one entity topic, prototypically a subject constituent. Frame TOPICS may be adverbials and nominalized verbs may also serve as TOPICS. However, a TOPIC may also cover longer stretches of speech such as different types of dependent clauses, e.g. circumstantial clauses and conditional clauses that typically constitute frame TOPICS. Furthermore, a TOPIC does not have to coincide with a syntactic constituent, as is mostly assumed in the literature. A special case of TOPIC pertains to what I called relational presupposition. It could be shown that the presupposition of a proposition expressed by the subject and the verb is realized as a thematic TOPIC. Thus, we may assume that relational presupposition and TOPIC are co-extensive in EA, which in turn suggests that the TOPIC is the formal counterpart of rhematic FOCUS.

The second purpose of this chapter was to identify different types of TOPIC constructions in EA in terms of their semantic, syntactic and prosodic properties. To begin with, thematic TOPICS were distinguished according to their semantico-pragmatic property of aboutness, anchoring and frame-setting. Aboutness TOPICS by definition involve a referential entity topic, whereas frames only rarely do so. The detailed investigation of the TOPICS in EA yielded essentially four different types of the first kind: a clause internal type I called S-TOPIC, with the subject functioning as the topic of the sentence, and two left-dislocated types: the D-TOPIC and the R-TOPIC (see Maslova & Bernini 2006). In both cases, the topic constituent is referred to by a resumptive pronominal element within the clause. The main formal distinction between the R-TOPIC and the D-TOPIC is in terms of argument structure. While the D-TOPIC refers to an argument of the proposition, the R-TOPIC does not. Functionally R-TOPICS specify a reference point (Langacker 1993) to anchor the proposition in discourse. R-TOPICS are extremely frequent in EA, involving lexical NPs or independent pronouns that mark the possessor of another topical entity within the clause. An important characteristic of this construction is its recursivity. While D-TOPICS are restricted to one (perhaps occasionally two) entities, R-TOPICS are recursive - in principle ad infinitum. Following Lambrecht (1994), I assume that the function of both types of TOPICS is to differentiate between the reference function of an expression and the role its denotatum plays within a proposition (Lambrecht’s Principle of Separation of Role and Referent (PSRR)) to facilitate cognitive processing in production as well as perception.
Finally, left-peripheral hanging TOPICS may also be *unlinked*, with no formal grammatical relation between the left-peripheral element and a constituent within the clause. These were called *free TOPICS*. Functionally, S-TOPICS and D-TOPICS are always characterized by an aboutness relation to the proposition, whereas R-TOPICS and free TOPICS more typically serve to anchor a proposition in the context and to delimit the frame within which the predication holds.

The opposite periphery of a sentence, the *postfield*, can host two types of topics in EA. One is the familiar type of antitopic (Chafe 1976, Lambrecht 1994) that necessarily refers to given, highly salient referents in discourse and accordingly exhibits a low flat contour with very little prominence. The other type is a highly prominent constituent following the main proposition, which I termed *topic addition*. Topic addition is a phenomenon especially characteristic for narratives where one topical referent is singled out as the main topic of a proposition and further participants are identified as additional topics in the postfield. These referents crucially have the same syntactic role within the proposition as the TOPICAL referent.

In addition, the study paid special attention to the behaviour of independent pronouns. Independent pronouns are attested in all different kinds of TOPIC types. One special type of pronominal R-TOPICS was said to have the pragmatic function of naming the source of a following statement or opinion or emphasising the relation between what is being said to the entity referred to by the R-TOPIC. All three personal pronouns participate in this function. First and third person pronouns are especially frequent, relating the utterance to the speaker or to a generalized other, whereas second person pronouns are mostly used in questions or answers. A pragmatic function typically related to such elements is hedging.

TOPIC domains frequently include clausal elements, especially restrictive relative clauses that further specify the topics embedded in the higher domain. Another type of longer TOPICS that involves more constituents than the entity topic itself is the relational presupposition TOPIC referred to above. It seems that the formal marking of this type of constituent is not restricted to EA. Steedman (1991, 2000) identified a special ‘thematic’ tune for such relational presuppositions whose overall tonal shape is very similar to that of EA TOPICS.

The study of *frames* concentrated on the description of temporal and spatial frames. It was shown that spatial frames are extremely rare in the corpus and that locative information is mostly coded as a clause-internal adjunct or occasionally in extra-clausal postfield position at the right periphery. In addition to these two types, another type of TOPIC (termed IF-TOPIC), consisting of a conditional clause as part of the presupposition, was identified, which is prosodically very similar to other TOPIC types and whose function could also be subsumed under the category of frame-setting. Moreover, the study revealed that certain expressions (*baːdeːn* ‘afterwards’) that at first glance look like temporal
frames are in fact connective elements that share with other types of frame-setting TOPICS their identification of a reference point from which the event in the following proposition starts out. In addition, a small number of morpho-lexical constructions were identified, one of which involves a sequential marker (*baʔa*) that is frequently associated with topics or marks the border between the TOPIC and the FOCUS domains.

In the final two sections, the important question of the relation between information status or cognitive status and topicality was investigated, touching upon one of the major concerns in recent research, i.e. the distinction between continuing topics, shifting topics and contrastive topics. At the present state of the investigation, there is no indication that shifted topics and contrastive topics could be differentiated prosodically in EA, as claimed by some researchers. At present, I rather embrace the view that contrast is just a special kind of emphasis that is frequently, but not obligatorily, marked by a focus accent - whether in TOPIC or FOCUS position (see 4.5.2) and that a shifting topic also involves some measure of contrast. There, however, remains a host of unresolved questions, such as the question of the size and explicit mention of the *alternative set* that is referred to, which may prove to be fruitful for identifying an independent category of contrast. So, the lumping together of different functions of *imposing salience* (Mulkern 2003) is maybe not the final solution and clearly more research is needed in that respect. In the absence of explicit evidence to the contrary, I follow Bolinger for the time being in the contention that contrast is not “a category of grammar but the result of the general cognitive processes referred to as ‘conversational implicatures’” (Bolinger 1961: 291).

The final section was especially devoted to the much debated issue of quantificational topics, which is notionally related to the question of information status. It was argued that quantificational topics do occur in typical TC constructions such as D-TOPICS and R-TOPICS as long as they are identifiable to a certain degree, i.e. anchored in the discourse model of the interlocutors. Gundel’s proposal that nonspecific topicalized indefinites (‘most X’, ‘some X’) can be construed as topics, assuming that it is in fact the entity itself (which is identifiable because of its being generic) that is the aboutness topic and the quantification may be considered as part of the predication (Gundel 1999b), was adopted and formal evidence for this assumption could be adduced. EA, in fact, has a construction that exactly corresponds to the analytic paraphrase suggested by Gundel. It was, however, also shown that formal requirements such as definiteness or restrictions on the cognitive status of topics (e.g. unique identifiability) other than the condition that they have to be anchored (Prince 1981; Ward & Prince 1991) do not hold.

The prosodic analysis of the individual topic/TOPIC types identified above shall now be summarized with reference to some suggestions made in the literature for other languages. The comparison is difficult to a certain extent because the phonological model used here differs from the standard assumptions in that it does not identify discrete phonological pitch accent types that are only
associated with the stressed syllable, but pays more attention to the contour that follows the accented syllable. Moreover, the different proposals in the literature also differ in the way they define the accent types they assume. This is due to the difficulty posed by the gradience of the alignment parameter and the unsettled issue of what a starred tone actually is as well as the relatedness of certain accent types (e.g. L+H* and H*) (see the discussion in Chapter 2.2.4). In a very interesting paper, Hedberg (2006) discusses seven different analyses of the prosody of topics in the literature and shows that they not only differ terminologically, but that the categories they presume are also notionally different. One such complication for the comparison I am going to undertake here lies in the different assumptions about pragmatic constituency between my approach and the other approaches, as there is no category that matches my TOPIC. Some comparability is, however, warranted, as I mostly analysed TOPICS that coincide with a topic. In the remaining cases I picked the final pitch accent of the constituent like Hedberg and Sosa (2008) did for FOCUS domains in their investigation of English spontaneous dialogue.

To summarize the EA results, it was found that topics have a different prosody whether they occur in the left periphery or in the right periphery of the sentence. Let us first look at the less problematic cases of right peripheral topics. As already noted, we have identified a right-peripheral domain, the postfield, that may host two different topic types: antitopics and topic addition. The difference between these two is invariably associated with a prosodic difference. While antitopics are associated with a low linking contour, topic addition is highly prominent and mostly associated with closing accents, but default accents also occur. The flat antitopics have their counterpart in many other languages (Lambrecht 1994: 204; 2001: 1071; see Frascarelli & Hinterhölzl 2007 for Italian; Uhmann (1991) for German; Lambrecht 1981 for French; Astruc-Aguilera & Nolan 2007 for Catalan). Some of the authors note that antitopics are by definition unstressed or unaccented (Lambrecht 1994), while others only refer to their low flat intonation and suggest some degree of prominence, e.g. Frascarelli & Hinterhölzl (2007) note that antitopics in Italian are characterized by an L* tone.

According to the latter, the same contour also characterizes continued (familiar) topics earlier in the sentence in Italian as well as in German. Hedberg and Sosa (2008)’s results for English yield 80% deaccented occurrences of familiar topics, while 10% were characterized by an accented low linking contour (L*).\textsuperscript{31} Turning to EA now, we clearly see that it differs from the European languages discussed in the literature. Given that in EA deaccentuation of content nouns is avoided and that the default accent implies a rising part on the accented syllable of a word, we do not find many cases of a low linking contour in EA. However, the cases identified are mostly associated with continued topics, which is up to expectations. The important fact is that these topics are in the overwhelming majority of

\textsuperscript{31} Hedberg & Sosa also report some difficulty in distinguishing deaccentuation from low accent in some cases, which is again evidence for the gradience of accentuation assumed in the present proposal.
cases expressed by pronouns. Although lexical NPs with a low linking contour do occur, these occurrences are marginal. They are mostly associated with a verbatim repetition from a preceding clause. But again, deaccentuation here is not obligatory. On the other hand, mentioning of the topic in the immediately preceding clause is not required, it seems that the referent of the low topic expression only has to be salient in discourse. On the other hand, continued topics in EA are also frequently associated with less prominent accents than the accents in the FOCUS domain. In such cases, the TOPIC is realized in a compressed pitch range in the left periphery and the first accent in the FOCUS domain shows a marked upstep. This low prominence is not a characteristic of topicality, however, but rather related to its information status and in further consequence of its being taken for granted, thus having the pragmatic property of background.

The great majority of TOPICS/topics in EA are associated with a leading-linking contour or pitch accent. This contour is preferentially associated with temporal frames as well as S-TOPICS which may be continued or shifted, but are fairly accessible. A leading contour is more typical of left dislocated TOPICS, such as D-TOPICS and R-TOPICS. It is also frequently used with clausal TOPICS, whether they are used to anchor and specify an entity topic or as a frame for the whole proposition, as in the case of IF-TOPICS. Finally, closing contours on topic expressions are not uncommon, too. They are associated with ‘new’ topics, i.e. with topic shifts and, just like leading contours, may also signal contrast. Importantly, leading accents that involve a high boundary tone and necessarily an intonation break, which are thus especially prominent, are usually associated with FOCUS constituents and not with TOPICS.

The leading and leading-linking categories are difficult to compare with the accents identified in other languages reported in the individual papers, as already observed above. A further problem for the comparison is that some approaches differentiate between contrastive topics and shifted topics, while others do not. As the question of contrastivity is still a pending issue and such a distinction is not made in the present work, I will treat these two categories together for the present purpose. In doing so, we are in the position to discern a certain pattern that emerges from the descriptions. It has been noted that LH* accents mark topics in English (Steedman 2000, Gundel & Fretheim 2004, Vallduví & Engdahl 1996, Hedberg & Sosa Juan M. 2008) and German (Frascarelli & Hinterhölzl 2007). Likewise, L*H has been found to be associated with topics in Italian (Frascarelli & Hinterhölzl 2007) and German (Uhlmann 1991, Féry 1993, Büring 1997; Jacobs 1997, 2001; Frascarelli & Hinterhölzl 2007). Finally, the H* accent has been found on topics in English (Hedberg & Sosa 2008), in Italian and German (Frascarelli & Hinterhölzl 2007). Some authors have explicitly claimed that LH* marks topics whereas H* marks FOCI (Vallduví & Engdahl 1996, Gundel 1999, Steedman 1991, 2000; Gundel & Fretheim)

32 Of course, this does not do justice to the proposals made by some authors who claim that the prosodic differences between these topic types are categorical (e.g. Frascarelli & Hinterhölzl 2007). The reader is explicitly referred to the individual papers for the details of the proposals.
2004) and new information (Pierrehumbert & Hirschberg 1990). However, Hedberg & Sosa (2008) clearly demonstrate that there are no dedicated topic and FOCUS accents and that the H* accent in English is compatible with every information structural category, and that specifically LH* is used with topics and FOCI alike, provided they are contrastive or shifted. The authors thus conclude that LH* is “a mechanism for emphatically highlighting an element relative to its context” (Hedberg & Sosa 2008: 115). As the above discussion shows, it is problematic to correlate pitch accents directly with specific information structural categories. However, the data from the individual languages suggest that the contours identified for EA are like the contours found in the European languages, if the tonal shape after the pitch accents is also taken into account. To complete the picture, it is therefore necessary to look at the tonal ‘continuation’ of these pitch accents.

It is noteworthy in that respect that some approaches have explicitly claimed that it is whole tunes that are marking topic or focus. Such proposals have been made by Jackendoff (1972) who claims that the B accent (a fall-rise) is used for presupposition and the A accent (a fall) is used for foci. This proposal was adopted by Pierrehumbert (1980) who translated it into AM annotation: topical H*LL% vs. focal H*LL%. Steedman (2000) suggests that “themes” are characterized by an LH* LH% and “rhemes” by an H* LL% tune.33 Let us ignore for a moment that these proposals all involve obligatory boundaries and boundary tones - an assumption that was not adopted in the present model - and only look at their tonal shape. This clearly shows that topical accents are predominantly rising and FOCAL accents predominantly falling. If we now take into account the additional information about boundary tones given in two of the descriptions of topic accents (Frascarelli & Hinterhölzl 2007, Hedberg & Sosa 2008), we find the following picture emerging: The first challenge to the strong correlation of information structural constituents with prosodic categories posed by the fact that Hedberg & Sosa (2008) found that both focus and topic may be associated with H* or LH* may be partly resolved if we resort to the contour after the pitch accent. Looking at their examples, we find that most of the FOCI they cite involve a low boundary. However, a relatively high number of topics showed a low boundary as well. Again, looking at the examples suggests that these topics frequently are contrastive or at least shifted or new (or unratified in Hedberg & Sosa’s terms). Similarly, Frascarelli & Hinterhölzl (2007) note that their examples of Italian shifted and contrastive topics always involve a low boundary tone after the pitch accent. These results are in concordance with the EA data that also exhibits a fair number of closing accents on shifted or contrastive topics. At the same time, Hedberg & Sosa conclude from the results of their study that L* tends to mark focus and L*H tends to mark topic, while upstep seems to mark focus (p. 119), thereby supporting Steedman’s (2000) proposal that L* is a variant of H*L and typically rhematic, which is also in line with the results of the present study, where

33 As already noted in Chapter 2.2.5, a similar proposal has also been made by Brazil (1975) for British English, however Brazil does not refer to information structural categories such as topic and focus, but rather speaks of referring and proclaiming tones, cf. also Gussenhoven's selection and addition (1983, 1984).
final L* is considered to be a variant of downstep. The other occurrences of the rising pitch accents described for the various languages above do not seem to involve a subsequent fall and may thus be taken at face value as corresponding to our leading or leading-linking type. It thus seems that the EA preference for leading and linking TOPICS is shared by the European languages discussed in the literature, with some notable language-specific differences that shall be summarized here:

- The major difference between EA and the other languages concerns the low frequency of deaccentuation.
- This fact implies that the low linking contour (deaccented or L*) in the European languages, especially those of the West-Germanic family, corresponds to a leading-linking contour (i.e. H-, LH) in EA, sometimes realized in a compressed pitch range.
- Consequently, the leading-linking contour or default accent in EA does not give rise to the implicature of contrastiveness and is thus preferentially used for highly accessible topics.
- To mark topic shifts (whether in sentence topic or discourse topic) or to emphasize aboutness, EA thus resorts to phonetically more prominent accents of the leading or the closing type, with the leading type implicating aboutness and the closing type newness or contrastiveness.
- EA is an “H* language”, to use Hedberg & Sosa’s (2008: 119) term, contrary to Spanish or Italian that frequently use rises with the low tone associated with the stressed syllable. Consequently, L*H rises are less frequent in EA.

The results of the detailed investigation of topicality in EA thus supports the hypothesis that the leading(-linking) contour is preferentially used for TOPICS, thereby supporting claims made for other languages in the literature. On the other hand, it has also been shown, that prosodic type and information structural category do not show a one-to-one correspondence. This result is in line with Hedberg & Sosa's general conclusion of their study of English spontaneous speech:

We conclude that while there are systematic correlations between intonation and information structure categories, these correlations are not as straightforward as is suggested in the literature. In particular we deny that there is any prosodic category as distinctive as a ‘topic accent’ as opposed to a ‘focus accent.’ (Hedberg & Sosa 2008: 119).

The existence of a systematic correlation found between TOPIC and the leading-linking contour is due to the universal property of rising or ‘open’ contours marking non-finality (Cruttenden 1986, 2006), which is rooted in the production code (Gussenhoven 2002). The correlation of TOPIC with the leading contour implicating aboutness, expressing a look-ahead strategy is rooted in the frequency code (Ohala 1983, 1984; Gussenhoven 2002). On the other hand, the imperfect correlation between prosody and pragmatic categories is due to the fact that the meaningful contours of intonation serve a number of linguistic and non-linguistic functions that co-occur on the same linguistic material. The fine details of accent-type and contour-type as well as their stronger or weaker conventionalization for certain functions, however, are language-specific and move prosody a little closer to grammar.
6  Theticity

If we should ever need to talk regularly and frequently about independently operated sawmills from which striking workers are locked out on Thursday when the temperature is between 500° and 600° F, we would find a concise way to do it. (Bolinger & Sears 1981: 114).

In Chapter 1 I have given a brief overview on the history of the discourse-pragmatic approach to theticity and its philosophical foundations. Nothing has been said about other treatments of the phenomenon, especially from the semantic perspective. In this chapter I will also be concerned with the semantic aspect of the phenomenon. I will only be referring to individual proposals that are relevant for the arguments proposed in the present study.¹ As already noted in the introduction, Lambrecht (1994, 2000) reinterpreted Sasse’s thetic statement in terms of Sentence Focus, arguing that Sentence Focus is a universal linguistic category on a par with other focus construction types. In Section 6.1 I will argue that theticity is only a special type of an all-FOCUS utterance. It will be argued that the phenomenon of theticity has a semantic basis, whereas a category of Sentence Focus may not be necessary at all. I will rather refer to the relevant cases as all-FOCUS utterances to indicate that the utterance lacks any relational presupposition, thereby avoiding the invocation of a specific ‘category’. Section 6.2 will present some constructions typical for all-FOCUS utterances in EA. In Section 6.3 I investigate the question of whether EA has a construction similar to the well-known cases of English and German subject accentuation. To begin with, let us look at the phenomenon and the way it has been claimed to be realized in a great number of European and non-European languages. The following list of thetic utterances is part of the one given in Sasse (2006: 255f.):

(1) My NECK hurts.
French: J’ai mon COU qui me fait MAL.
German: Mein HALS tut weh/Mir tut der HALS weh.
Hungarian: Fáj a TORKOM.
Italian: Mi fa male il COLLO.
Japanese: KUBI ga ITAI.
Modern Greek: Ponai o LEMOS mu.
Serbo-Croatian: Boli me VRAT/GRLO.
Spanish: Me duele el CUELLO.
Irish: Tá pian i mo SCORNACH/Tá mo SCORNACH nimhneach.

(2) The PHONE’s ringing.
French: Y’a le TELEPHONE qui SONNE.
German: Das TELEFON klingelt.
Hungarian: Csöng a TELEFON.
Italian: Squilla il TELEFONO.

¹ For an overview of the history of the research on theticity the reader is referred to Ulrich Haberland ((1994)), and Matić ((2003))
Japanese: DENWA ga NATTE iru yo.
Modern Greek: Xtipai to TILEFONO.
Serbo-Croatian: Zvoni TELEFON.
Spanish: Suena el TELEFONO.
Irish: Tá an FÓN ag ringáil.

(3) Nominalizations
fit(h) darb (Modern Arabic varieties, Sasse 1987: 553)
EXIST thrashing
'There is thrashing,'
d’etégé-d’uwinna (Boni, Sasse 1987: 546)
ear-my–hurts
'My EAR hurts'

The examples in (1) and (2) not only show a remarkable similarity across languages as far as their formal properties are concerned, they are also said to be used in these languages in the same type of situation, where it is appropriate to present a (new) situation as a whole, thereby exemplifying a specific type of information packaging. As far as the formal similarities are concerned, we may notice that in all languages the subject carries an accent. Furthermore, in most languages this is the only - or rather the primary - accent of the phrase, the predicate being downtoned. This is a marked type of accent pattern. In the unmarked case, i.e. in cases of topic-comment utterances, the predicate is accented. It is only in two languages, French and Japanese, that we find double accentuation. These languages, however, exhibit another strategy that has been claimed to express FOCUS. In Japanese, the subject is marked with the particle ga which stands in opposition to the particle wa that is said to mark topics. In French, the subject is introduced via the presentative construction y’a (il y’a in written French) and the predicate phrase is expressed as a relative clause. The theticity effect of ga marking in Japanese has been pointed out by Kuroda (1972) whose proposal later on became very influential in establishing what may be called the linguistic approach to theticity, most prominently represented by Sasse (1982, 1987, 1995, 2006). Finally, the word order in some languages, most notably in Italian and Spanish, is VS, where it has been assumed that the canonical word order, or at least the order reflecting a topic-comment structure, is SV (cf. Lambrecht 1994, et passim). In (3) thetic utterances from Arabic and Boni, an Eastern Cushitic language of Kenya, are presented. In the Arabic case, the 'thetic' property is quite obvious as the event is presented without mentioning any participant and is thus an instance of the 'event-central' type of thetic utterances (see below). The Boni case is interesting. Sasse (987: 546) notes that the noun and the verb form "a tight phonological unit", i.e. one accent domain or phonological word, with the "verb encliticized to the noun", a construction that can be regarded as a case of incorporation. Thus, in Boni the relationship between subject and predicate is comparable to that in the European languages cited under (1) and (2).

Following Sasse (1987: 526), we may distinguish between two types of thetic utterances: the entity-central and the event-central type, which may be shown to be typologically relevant (p. 527, 563f.). Sasse identifies three types of thetic statements (p. 559):
We may now summarize that the strategies identified by Sasse to express theticity, namely VS order, subject accentuation (henceforth SAcc construction, following Sasse 2006), nominalization, noun incorporation and split structures, all “help diminish the grammatical predicativity of thetic sentences” (Sasse 1987: 519) and successfully obviate the bipartiteness characteristic of topic-comment structures. Intuitively, the formal criterion of a monomial structure is met in all thetic statements discussed above, with the notable exception of French and Japanese, by integrating the entire information under the scope of one accent.

Ever since the suggestion of the thetic/categorical distinction as a categorial linguistic dichotomy, much research has been done on the issue, especially in a study conducted by a group of researchers (Theme Group 1) within the EUROTYPO project (Typology of Languages in Europe) devoted to the functions of VS word order in European languages, the results of which were published in Matras & Sasse (1995). These results have led the authors to cast doubt on the existence of a cross-linguistic category of theticity. Additional evidence comes from a study on VS structures in the languages of the Balkan by Matić (2003). In his recent publication on the topic, Sasse himself abandons the ‘strong theticity hypothesis’, adopting a ‘weaker theticity hypothesis’ that assumes theticity to be a cross-linguistically comparable phenomenon, but not a linguistic category (Sasse 2006: 300). It has become increasingly clear that the prior assumption of a simple thetic-categorical distinction proposed at the outset of the study of the phenomenon cannot be maintained.

As already noted by Sasse, thetic constructions in Arabic are expressed by nominalization and VS order. We will see instances of the latter in the course of the study. In this study, however, I will concentrate on prosody and investigate whether EA has some type of SAcc construction. The information structural property of word order in Arabic vernaculars has been discussed in several studies (Brustad 2000; Dahlgren 1998; Ingham 1994, 2010; Holes 2010; Edwards 2010), in Modern Standard Arabic by Pashova (2003) and in Classical Arabic by Moutaouakil (1989), but a special treatment of thetic utterances is still missing. A thorough investigation of the syntax of thetics, however, must be referred to future research, as it would exceed the scope of the present work.

6.1 Theticity and the notion of 'Sentence Focus'

In the view presented here, theticity is a semantico-pragmatic phenomenon closely related to other types of integration that have been already discussed in Section 4.5.5. We will reconsider the
phenomenon of integration shortly. In this section, I will examine Lambrecht's (1994, 2000) reinterpretation of theticity in terms of Sentence Focus. Lambrecht's proposal sheds the remaining semantic basis in Sasse's conceptualization of the phenomenon and establishes it as a genuinely grammatical category that serves a discourse-pragmatic function. Sentence focus\(^2\) in Lambrecht's model is defined as a structure in which no pragmatic presupposition is formally evoked, which means that FOCUS and assertion coincide. The thetic or Sentence Focus utterance, as conceived by Lambrecht, is a "universal information-structure category" (Lambrecht 2000: 611) on a par with his other focus categories (Predicate Focus and Argument Focus). As we have already noted in several places in this study, the assumption of information structure categories is a problematic issue; this is all the more so in the case of a 'Sentence Focus' category, which Lambrecht (2000: 617) defines as a

\[\ldots\] sentence construction formally marked as expressing a pragmatically structured proposition in which both the subject and the predicate are in focus. The focus domain is the sentence, minus any topical non-subject arguments.

Lambrecht's strict pragmaticization of theticity and his claim of its categorial status are flawed on several grounds. First, there are many utterances in which a pragmatic presupposition is not evoked and in which both, the subject and the predicate, are in FOCUS, but that cannot be considered as thetic, at least in the original sense of the concept (see (4) and (5) below). Second, the category is far two constrained semantically to be regarded as a universal category of information structure. Contrary to Predicate Focus and Argument Focus, which may be applied to virtually every subject-predicate combination, Sentence Focus is only possible in a very limited number of cases. Third, some of the alleged Sentence Focus constructions, e.g. split constructions of French and Arabic, simply do not fulfil the conceptual requirement that the proposition expressed should be simple in the sense attributed to it by Kuroda (1972) in his much cited example *Inu ga hasitte iru* 'There is a dog running' as opposed to *Inu wa hasitte iru* 'The/A dog is running.' Kuroda (2005: 1) explicitly denies that his thetic/categorical distinction was meant in terms of information structure, with *wa* marking topics and *ga* marking focus, but that the distinction is a semantic one between different types of propositions. Kuroda explains the meaning of the *ga*-sentence as an event of running taking place, in which a dog is involved as a participant of the event without being singled out conceptually.

To justify the objections made above, I invite the reader to consider the following pair of sentences discussed by Sasse (1987: 521) (5) and Lambrecht (1994: 309) (6) (the latter being cited after Bolinger 1954):

\begin{enumerate}
  \item What's new? - HARRY's coming.
  \item What's going on outside? - HARRY's SINGING.
\end{enumerate}

\(^2\) Note that it would be FOCUS in our terminology. I will, however, stick to the original writing when referring to Lambrecht's concept.
(6) Why didn't she come to work today?
   a. Her HUSBAND is sick.
   b. Her HUSBAND broke his NECK.

While the (a)-examples of (5) and (6) are considered as thetic (in all accounts of theticity I am familiar with), the (b)-examples are not. From a pragmatic perspective, the exemption of the (b)-examples from Sentence Focus is counter-intuitive. First, the subject of the sentences is discourse-new and thus exhibits one important characteristic of a prototypical FOCUS constituent. Second, and more importantly, the (a) and the (b) sentences are uttered in the same type of discourse situation and may both be counted as the \textit{pragmatic} assertion to an open presupposition 'What is happening is X' in (5) and 'The reason why she didn't come to work is X' in (6). On the basis of Lambrecht's Sentence Focus definition alone (without any further requirements he posits as well), the (b)-sentences should qualify as instances of Sentence Focus. And in fact, Lambrecht himself notes that in the statement given in (7) (Lambrecht 1994: 121, ex. 4.2c) the subject NP is "non-topical" and that the answer is not in the first place construed, i.e. interpreted, as information about the children, its pragmatic function being "event-reporting" with the FOCUS covering the whole proposition (p. 124).

(7) (What happened?)
   The CHILDREN went to SCHOOL!

What Lambrecht and Sasse point out is the fact that only in the (a)-examples of (5) and (6) this fact is unambiguously marked linguistically. The (b)-examples, which are in Sasse's proposal instances of the categorical statement and in Lambrecht's model cases of Predicate Focus, are regarded by Lambrecht as allowing different types of focus \textit{construal}, being the unmarked type of focus \textit{construction}. But if this is so, why then should we assume a pragmatically based grammatical category in the first place, when we do not assume it to encode a specific pragmatic function?

Lambrecht and Sasse both take it as a point of departure that the (b)-examples above are more or less homophonous with a topic-comment sentence (e.g. one with a contrastive topic). This assumption obviously stems from the conviction that accent placement is the only relevant prosodic feature in connection with focus, as I have repeatedly criticized in this study. I hope I am not accused of mere speculation when I claim that speakers can make a distinction between an all-FOCUS articulation and a topic-comment articulation, if they wish to in the cases at hand. I have already pointed to that possibility in Chapter 4.5.1 in the discussion of the German example \textit{Die Präsidentin hat den Dekan gesehen}. It is an uncontested fact that context does not fully specify the information structure of an utterance. A speaker does not \textit{have} to answer with an all-FOCUS utterance in the above cases. \textit{Harry} and the \textit{husband} in principle denote accessible referents, i.e. they may well be accommodated as topic.

\textsuperscript{3} In Lambrecht's terminology the term \textit{construal} is used to refer to interpretation, whereas \textit{construction} refers to the linguistic expression.
referents by way of an existential presupposition. I suggest that there are in principle three categorically different possible answers to the questions in (6) and (5b), respectively:

(8) What's going on outside?
   a. HARRY'S SINGING.
   b. HARRY'S SINGING.
   c. Harry's SINGING.
   d. # HARRY'S singing.

(9) Why didn't she come to work today?
   a. Her HUSBAND broke his NECK.
   b. Her HUSBAND broke his NECK.
   c. Her husband broke his NECK.
   d. # Her HUSBAND broke his neck.

The cross-classification shows the preponderance of the leading-linking accent (50% of the cases), followed by the linking type. Among the interesting aspects of the count is the behaviour of pronouns. Although in both categories, the leading-linking accent is the most common one (almost 50%), we also find a fair number of linking tones on pronouns (37% for subjects, 27% for non-subjects). Pronouns are given by definition and low prominence is thus an expected feature. The quantitative results thus support what has been said about pronouns in Section 5.3.1.5, namely that the prosodic shape of the pronoun correlates with its function as a continued topic expression or a contrastive or at least shifted topic.

\[4\] I am only illustrating the difference schematically, ignoring the possibility of a more complex topical accent in English (cf. Brazil 1997; Gussenhoven 1984; Büring 1997) for the sake of simplicity.
While the (a) answers have two closing accents indicating FOCUS, the (b) answers have a leading and the (c) answers a linking contour, which typically indicate TOPIC. The existence of Harry or the husband in (c) is taken for granted and the referents are treated as ratified topics, whereas the topical referents in the (b)-cases are not. Instead they are activated and established as topics. The sequence of a leading and a closing accent, whether in the form of a hat pattern or a roof pattern (cf. 2.2.5.2 and 4.5.5) mapped to the syntactic sequence of subject + predicate, may thus be regarded as a construction that is typically associated with a pragmatically structured proposition of topic + comment or - in accordance with the discussion of TOPIC in Chapter 5 - of TOPIC (i.e. relational presupposition) + FOCUS. If we take the idea of constructions seriously, we must distinguish this type from a similar, though not identical, construction involving a subject that is coded as FOCUS (the (a)-cases), which in turn diminishes its prototypical topic qualities and weakens the topic-comment or TOPIC-FOCUS relation of the constituents. What about the (d)-answers? The (d)-answers are clearly infelicitous in the given context as they only permit a narrow argument FOCUS reading, which is inadequate as an answer to the question of concern. But why does it only permit an interpretation that takes the subject to be the FOCUS? Because integration would not be possible in these cases. The reason for the impossibility of integration has to do with the type of the predicate involved: its information value, its semantic type and its lexical solidarity (Coseriu 1967) with the subject, i.e. its likelihood of forming a collocation with the subject.

So far, we have only looked at theticity from a pragmatic point of view. At this point an excursus about the semantics of the phenomenon is called for. The semantic aspect has been in the foreground of a number of treatments, of which I shall only refer to the explanations offered by Chafe (1974), Fuchs (1980), Bolinger (1954, 1985, 1989) and Jacobs (1993, 1999). Let us look at the way integrated sentences are treated by these authors. Chafe (1974: 115) notes that in (10) and (11) the noun and the verb both convey new information. He takes the difference between them to be explainable in terms of how the speaker perceives the event, whether he conceives it as a "conceptual unity" or not. This explanation is reminiscent of Kuroda's approach to the difference between Japanese wa and ga.

(10) (Chafe's examples 5b and 5c)
   a. My SISTER is DYING.
   b. The BUTTER MELTED.

(11) (Chafe's examples 6b and 6c)
   a. My SISTER is dead.
   b. The BUTTER melted.

Chafe (ibid.) notes that in (10), the sister and her death have not been established as a unity yet, but after her death the event of her dying will have acquired the status of a "single idea" in the speaker's mind. Similarly, the melting of the butter is regarded as a conceptual unity of ‘butter-melting’, an instance of which is expressed in (11) but not in (10). The idea of the whole proposition forming a 'conceptual unity' was criticized by Sasse (1987: 558) who notes that the phenomenon was rather one
of how the event is being "presented", i.e. a matter of what Chafe called 'packaging' and therefore a pragmatic and not a semantic one. Although this is obviously true, there is definitely something intuitively right about Chafe's explanation. Considering the (a)-examples again, note that the verb form in them is different. In the first case (10a), the verb is in the present progressive tense and the event is thus stated as happening right now. Given the semantic content of the utterance, it would perhaps not be impossible, but at least somewhat odd, if the speaker deaccented the verb, thereby playing down the situation denoted by the verb. The accent on the verb and the choice not to use integration here is a matter of attitude and the wish to put some emphasis on what is happening. In (11a), however, the speaker presents the action as a matter-of-fact, in fact he only presents the result of the action by stating it with some emotional distance. This is the Bolingerian theme again: the difficulty to disentangle the attitudinal and the grammatical in intonation.

An important contribution to the discussion of subject-accented patterns was made by Anna Fuchs (1980). Like Chafe, she recognizes the 'all-new'\(^5\) property of multi-accentual patterns and contrasts them with all-new patterns that only carry one accent. These patterns she calls "integrative". Fuchs (1980: 453) notes that while every 'all-new' SAcc sentence may also be uttered with accents on subject and predicate, the reverse is not true. As a result, the SAcc pattern has a more limited distribution and a more specific meaning or discourse function. Fuchs (ibid.) suggests that the limitations are of a discourse pragmatic nature which I will summarize here: (1) Neither the subject, nor the predicate must have referents that "are at the center of attention", which I take to mean that neither must be part of the presupposition of the utterance at hand. Fuchs is clearly aware of the difference between information status and pragmatic relations as she notes that "[w]hat accent signals is not newness of a constituent in discourse, but its relevance to the issue at hand" (p. 454). I thus believe that the first restriction may safely be translated into our terminology as the requirement for the subject referent as well as the denotation of the predicate to be in FOCUS. The second restriction is more interesting. (2) Fuchs notes that the concepts introduced by the subject or predicate must be "centered" (p. 455). The function of her "local centering" and "temporal centering" may be interpreted as the pragmatic requirement of a spatiotemporal anchoring of the utterance, which lies at the heart of the claim made in the literature that every thetic utterance has a stage topic that situates it locally or temporally (Erteschik-Shir 1997). This also excludes generic statements, i.e. the ascription of an individual-level predicate to an entity (Kratzer 1995). Finally, Fuchs' restriction (3) is concerned with the lexical type of predicates allowed or their predictability as Bolinger puts it (Bolinger 1954, 1972, et passim). Lexical restrictions have been most thoroughly discussed by Allerton & Cruttenden (1979) who, however, only offer an enumeration of partly heterogeneous verb types admissible in thetic utterances, such as empty verbs, verbs of appearance, disappearance, and also of misfortune. Fuchs' third restriction tries to account for the exclusion of subject-predicate sequences that are not already

\(^5\) Note that what Fuchs calls 'new' corresponds to our FOCAL.
excluded by the first two restrictions. Restriction (3) is similar to Chafe's *conceptual unity* in that it requires the subject and predicate "to denote some conventional kind of unit" (p. 457). Importantly, Fuchs notes that this unit must be related to the situation at hand, i.e. that the subject-predicate combination must be expectable at this very moment in discourse, thereby referring to the lexical solidarity issue invoked by many scholars and the contextual appropriateness at the same time. As I understand it, the two important contributions made by Fuchs are to have pointed out the pragmatic and semantic factors connected with the issue of theticity and secondly to have identified subject-accented patterns as an instance of *integration*, a concept that she had introduced in Fuchs (1976), to account for the fact that in English and German different types of lexical combinations, whose individual constituents offer new information, may be formally coalesced by putting it under the scope of a single accent and thereby forming "one unit of information" (Fuchs 1980: 449).

The proposal made by Fuchs was taken up by Jacobs (1993) who furnished the concept with the formal definition of a limited number of syntactic and semantic requirements. Although the conditions for integration listed by Jacobs cover many of the issues relevant for thetic sentences that have been discussed in the literature, such as the semantic types of predicates and subjects and monoargumentality, the list is not complete and primarily based on grammatical restrictions. Jacobs’s proposal especially falls short of accounting for the subtle differences in integratability due to lexical-semantic effects. Its main contribution, however, is the attempt at a semantic explication of the phenomenon in terms of cognitive processing. Jacobs defines integration, or *informational nonautonomy* as he calls the phenomenon in a later paper (Jacobs 1999), as the merger of semantic elements within a semantically compact unit to the effect that establishing extralinguistic reference for the latter only requires *one step* of cognitive processing. This means that an integrated construction permits a non-compositional or holistic decoding of its meaning.

The important aspect of Jacobs's proposal is its acknowledgment of the common motivation behind observable accentual patterns of constituents as diverse as compounds, predicate phrases and certain short sentences (12), this common motivation being the establishment of an informationally holistic unit. The integrational approach thus likewise accounts for those cases where integration is not permissible, such as coordination, individual-level predication and the modification of an event or action (13). The following examples are all taken from Jacobs (1993), accent positions have been added.

(12)
  a. FLÜSSIGtreibstoff (liquid fuel)
  b. ROTgrün (a reddish type of green)
  c. Ein GeWITter zieht auf (a thunderstorm is approaching)
  d. eine TÜre öffnen (to open a door)
(13)
   a. Der BUNdeskanzler und der AUßenminister (the chancellor and the foreign minister)
   b. ROT-GRÜN (red-green)
   c. Das Gewitter/GeWITter ist SCHRECKlich (the thunderstorm is terrible)
   d. aus ANGST daVONlauf en (to run away out of fear)

The clearest cases of integration vs. non-integration are compounding vs. coordination. It is intuitively compelling that in the former the semantic autonomy of the individual parts is given up, while in the latter the constituents remain informationally autonomous. Concerning thetic vs. categorical utterances (12c and 13c), Jacobs notes that in (12c) the reference to the event of a thunderstorm approaching is made in one go as opposed to the identification of a specific thunderstorm first and the ascription of a certain property to it afterwards, as in (13c). The door-opening event in (12d) allows the same explanation given for (12c). It is noteworthy that it is not the grammatical constraint of *indefiniteness* that is at stake here, as the event *die TÜRE öffnen* (to open the door) could also be articulated in an integrated manner. It rather seems that the frequency of occurrence and the lexical solidarity between object and verb are the facts that make integration possible here. In (13d) the act of running away is further modified by naming the reason for it, which also requires two steps of semantic processing, as argued by Jacobs.

Obviously, Jacobs's suggestion does not solve the problem that the notion of *semantic compactness* remains rather vague. It remains particularly unclear what the linguistic conditions for integration are. But this problem is most probably unavoidable because of the culture-specific and language-specific property of the notion itself. Take for example the Classical Arabic word *xaʃʃa* and its German translation in Wahr mund's dictionary 'einem Kamel einen hölzernen Nasenring in die Nase legen' ('to put a wooden nose-ring into a camel's nose'); this example shows how complex a single concept can be. The culture-specific property of this complexity is also illustrated by the German or English 'cultural quasi-equivalents' to the Arabic word: *beschlagen* and *stud*. In short, it will obviously never be possible to find a universal rule for the amount of semantic contents and, for that matter, of lexicosyntactic constituents an integrated expression may consist of - as is humorously expressed in the quotation by Bolinger & Sears at the beginning of this chapter. Bolinger (1989: 215ff.) shows many examples where the accentuation of compounds is not even uniform in English itself. Depending on the frequency of a collocation, and therefore the degree of lexicalization of a compound, the accent pattern may vary. Thus, it is Élm Street, but Élm Court and stéam engine, but stéam locomóti ve (p. 222). According to Bolinger, it is possible to say *She has to undergo an órgan transplant* or *She has to undergo an órgan tránsplant*, “given the medical conventionalization of a particular surgical procedure.” We may conclude that integration can be 'more’ or ‘less’ clear. While *órgan transplant* is fully integrated, *órgan tránsplant* is syntactically integrated by compounding, but not prosodically. On the other hand, *the transplantation of an organ* is not integrated at all, neither syntactically, nor prosodically. Conversely, in *héart attack* complete integration has already been lexicalized, both
?héart attáck and definitely *an attack of the heart are not felicitous expressions for the given meaning.

The same observation is true for presumably thetic sentences. Bolinger (21: 101) argues against a structural basis of subject accentuation, showing how the information value of a predicate influences its likeliness of being integrated.

(14)
   a. My MOTHER’S coming.
   b. My MOTHER’S sick.
   c. ? My MOTHER had an accident.
   d. *My MOTHER fell off the ladder.

The pattern shown in (14a) arguably is the highly preferred articulation of an all-new utterance in that case, being on the verge of lexicalization. There are a number of quasi-lexicalized expressions, such as the notorious cases of phone-ringing and door-opening cited over and over again in the literature. (14d) stands on the other end of the scale and is clearly not felicitous. We do not all too often announce that somebody fell off the ladder (14b), the ladder being sufficiently informative to be articulated as a separate point of information. Conversely, announcing that someone is sick is a frequently made announcement. Similarly, the announcement that someone had an accident in (14c) may be regarded as a quite familiar collocation, but two factors make integration less likely here: the information value of the noun accident and the phonological weight of the predicate. As a result, speakers’ intuitions about the possibility of integration in this case vary. It has been claimed that monoargumentality is a major restricting factor obviating the possibility of thetic articulation. But this is dependent on the language and the thetic construction used (cf. Sasse 2006: 277f.). German, for instance, seems to be more tolerant of heavy predicates to occur in a subject-accented construction, cf. the example cited by Fuchs (1980: 451) du, der FERNseher macht so'n komisches Geräusch 'listen, the television's making a funny noise'.

In this excursus on the (lexical-) semantic aspect of theticity I have tried to show that theticity should rather be viewed as a semantico-pragmatic phenomenon based on informational integration intimately tied to other types of integration that more clearly belong to the semantic domain. This is not to say that I believe theticity to be primarily a semantic notion, I rather believe it to be based on a complex interaction between pragmatic motivation and lexico-semantic and generally cognitive restrictions.

Having established the connection between theticity and semantics, we may now return to the original topic of this section, the hypothesized identity of theticity and Sentence Focus as proposed by Lambrecht (1994, 2000) and the question of its categorial status. We have shown that some utterances

---

6 Note, however, that this could also be a case of narrow subject FOCUS, the borderline between the two FOCUS domains is not fuzzy.
such as *Her Husband broke his Neck* or *Her Husband had an Accident* with two closing accents should in fact be counted as Sentence Focus according to Lambrecht's pragmatic definition, but they are definitely not thetic because they do not fulfil the semantic requirements for integration, involving two 'points of information' in the Bolingerian and Hallidayan sense.

My next argument against the equation of theticity and a presumed category of 'Sentence Focus' concerns split constructions, as they typically occur in French and Egyptian Arabic. (15a,b) and (16a-c) illustrate typical examples of the thetic type:

(15)

a. [a noise coming from the hallway. A says to S.:]
   
   *Ah oui, j'oubliais. Il y a votre FRÈRE qui est ARRIVÉ.*
   
   'Oh, I forgot. Your brother has arrived.'
   
   (Film by J. Rivette, Secret defense, 1999; after {Wehr 2000, ex.82)

b. *Y'a le TÉLÉPHONE qui SONNE.*
   
   'The PHONE is ringing.'
   
   (after Lambrecht 1988a: 137)

c. *Qu'est qu'il y a? - C'est MAMAN qui me BAT.*
   
   'What's the matter? MUM is hitting me.'
   
   (Wehr 1984; cited after Sasse 1987: 538)

d. *Maman, c'est AUGUSTINE qui met ses MAINS dans mon ASSIETTE!*
   
   'Mum, Augustine is putting her hands in my plate'
   
   (Zola; cited after Wehr 2000, ex. 85)

(16)

a. fi: *TILIFU:N ʕaʃA:N-AK*
   
   EXIST phone for-2SG.M
   
   'Someone's on the phone for you'
   
   (constructed example)

   
   EXIST 40 robber rob.PTCP.PL-1PL
   
   'What are you doing here? - They said: 40 ROBbers kidnapped us'
   
   (Abul-Fadl 1961: 97; cited after Sasse 1987, ex. 77; transcription and glossing DE)

c. fi: *BINT isma-ha RAPINZIL itTWALAD-IT*
   
   EXIST girl name-3SG.F Rapunzel be.born-3SG.F
   
   'A girl named Rapunzel was born.'
   
   (Rap_F0_04)

d. fi: *KALb bi-J-ʕUDʔDʔ RA:GIL*
   
   EXIST dog IND-3SG.M-bite man
   
   'There's a dog biting a man.'
   
   (SFB 632-D2, translation task)
On the Sasse-Lambrecht view, these examples are all instances of thetic or sentence-focus constructions. The arguments put forward by the authors are the detopicalization of the subject (Sasse 1987: 540; Lambrecht 1988a: 135 et passim), and the formal subordination of the relative clause (Sasse 1987: 561). Lambrecht (1988a: 160ff.) shows that the avoir-cleft construction in French (as in 15a,b) may have an event-reporting function. The same function may also be fulfilled by a ce-cleft as (15c,d) illustrate. Event-reporting, in turn, is one of the functions typically fulfilled by thetic utterances. In the following, I will argue that not every event-reporting split structure is necessarily an instance of theticity. If this were the case then French and Arabic thetics were different from thetics in other languages in that they do not have any semantic restrictions. The first problem for the theticity interpretation that strikes us is the fact that the split structures in both languages exhibit accents on subjects and predicates. Let us now first deal with French. The Sasse-Lambrecht interpretation was rejected by Wehr (1984, 2000) who claims that the subject of a split structure is the topic of a 'bipartite' construction (zweigliedrige Struktur). Wehr notes the intonational difference between the 'all-new' clefts and the subject FOCUS clefts, the former lacking the 'tonal break' after the subject and the following low tone of the latter (Wehr 2000: 242) as in C'est PIERRE qui le bat with a presupposed predicate in contradistinction to (15c) where bat is accented. According to Wehr (2000: 275), the cleft construction is only a strategy to put the subject in a clause-final position (i.e. at the end of a mot phonétique) to make it eligible for receiving primary stress, which is necessary to activate the referent of the subject expression. On this assumption, Wehr analyses ce-qui and (il) y a-qui as discontinuous morphemes that mark topics. She thus is able to interpret the whole presentational construction as an ordinary sentence topic marked for its being new. The topic-comment interpretation is especially conspicuous in example (17) given by Lambrecht (1988a, ex. 9):

(17) A l'heure actuelle j'm 'plains pas, y'a un camarade d'usine qui m'ramène en voiture jusqu'aux Quatre Routes pour prendre l'autobus

'right now I'm not complaining, there's a friend of mine from the factory who drives me back to Quatre Routes to take the bus'

That (17) is not a thetic sentence is intuitively clear. In Lambrecht's (1988a) analysis - which has a lot in common with Wehr's interpretation - a new referent is introduced into the discourse by a cleft construction, thereby making it eligible as a topic. But what is the difference between (17) and the examples in (15)? As it seems, expressing the predicate in a subordinate sentences, as suggested by Sasse, is obviously not enough to diminish its predicative power. It is true that the whole construction crucially differs from a topic-comment sentence. If the colleague had been salient in the discourse, he could have been readily coded as a topic in something like mon camarade d'usine, il m'ramène...., but this lack of familiarity obviously does not make the whole construction thetic. The bipartiteness of the construction is felt clearly. This fact is also accounted for by Lambrecht who distinguishes between two types of cleft constructions, the presentational cleft construction, such as in (17), and the event-
reporting cleft construction as in (15), the latter of which he takes to be a thetic statement. He suggests that in sentences such as (17) the referents are "pragmatically salient discourse participants" which are likely to become topics in subsequent discourse (1988a: 160) while the phone in (15b), for example, is not likely to become a discourse topic. Obviously this cannot be the decisive factor for a differentiation, as it is not true that the participant introduced by the cleft is not going to be topical in subsequent discourse (cf. for instance 14a), and, what is more, event-reporting sentences may have semantic characteristics that are normally barred from thetic utterances, as (18) (from Lambrecht 1988a, ex. 23) suggests:

(18)
ben ouais ah ben dimanche quand on arrivait i'y'a une voiture qui a passé sur les pattes du chien de Gomez (Jeanjean 1979)

'well yeah oh well Sunday when we were coming home a car ran over the legs of Gomez' dog'

In (18) and similarly in (15d) above, it is definitely the case that an event is being reported, without the subject being coded as topic, i.e. it is clearly not a pragmatic predication about 'a car' (une voiture), but rather about the moment of arriving on Sunday or probably about Gomez' dog. Although Augustine is somehow a functional topic in (15d), the primary topic is the situation of a child complaining. The utterance would be a felicitous answer to a mother who, upon hearing her children quarrel in the kitchen, is scolding them and asking something like 'What's that fuss?' But does this imply that the answer is necessarily thetic? I believe not. If sentences such as (18) and (15d) are counted as thetic, this would mean that French cleft constructions allow for agentive subjects and transitive verbs with lexically coded objects, thus violating the well-known constraints against theticty in other languages (Sasse 2006, Jacobs 1993, 1999). Although Lambrecht (1988a: 161) invokes Chafe's concept of a conceptual unity for these sentences, it is difficult to see how a car running over a specific person's dog should form a 'conceptual unit' in the same way as the ringing of a phone or the arrival of a train etc. At the outset of the discussion, we have noted that one problem with the French split structures is the fact that both, the subject and the predicate, are accented. We have already seen that identificational clefts that code a narrow FOCUS subject differ intonationally from event-reporting clefts. Let us now take a closer look at this difference. Wehr (2000: 242) mentioned the 'tonal break' after the subject and the following low tone in narrow FOCUS clefts. Her analysis is partly corroborated by an investigation of different cleft constructions in spoken French by Clech-Darbon, Rebuschi & Rialland (1999) who also present pitch tracks along with their analysis7 which translates into our model as earlier alignment of an L tone after the H on the 'subject' in narrow FOCUS cases and late alignment in the all-FOCUS cases. Importantly, the tone associated with the verb is low in

---

7 Clech-Darbon et al. analyse the intonation of a narrow FOCUS declarative cleft structure as HL%L% and of an all-FOCUS structure as H%L%. 384
both cases, or at least downstepped in the case of an all-FOCUS sentence. (19) presents a schematic representation of the two cleft types:

(19)

a. Who is gone?
C'est PIERRE qui est sorti.
L- L H L%

b. What's the matter?
C'est PIERRE qui est SORTI.
L- L H( ) L%= (L%)

Even though the authors do not say anything about the prominence of the specific lexical items (or the mots phonétique), it can well be assumed that in (18a) sorti is deaccented while in (18b) it is not. But as we have already argued at length in Chapter 4.1.3, complete deaccenting is not necessary for an accent to be subordinated; it is enough if there is a prominence difference between the individual accents. In Section 6.3 I will show how downstep may be an intermediate step between full prominence and complete deaccentuation. Contrary to (19b), the examples in (16) and (17) clearly contain too much of new and interesting information within the 'relative' clause to be downtoned and prosodically subordinated. I assume that additional points of information coded by prominent accents mitigate against the full integration of subject and predicate.

If this is correct, we can assume that the criterion of prosodic integration is necessary for the thetic interpretation of an utterance. Under this assumption we will find that the well-known lexico-semantic restrictions do also hold for French, i.e. long utterances containing lexical objects - that will almost inevitably be associated with a (strong) prominence - do not count as thetic.

Now, we can eventually turn our attention to the EA examples (16a-d). Note that they contain accents throughout. Yet, some accents are transcribed with small caps, indicating their lower prominence. In agreement with my analysis of the French cleft constructions, I regard the sentences in (16a-c) as thetic, but not the one in (16d). Let us now look at the pragmatic, semantic and formal difference between them. Pragmatically, all EA utterances inform the hearer about something new. The utterances are all felicitous out-of-the-blue statements and do not refer to an already salient discourse topic. (16a), of course, is a clear case of entity-central thetic, and thus there is no need for argumentation. In the example from Abul Fadl (16b), the verbal part is not a finite verb, but a participle with a resultative meaning. The 'subject' NP ʕarbaʔi:n hara:mi 'fourty robbers', however, is not a prototypical topic. It is new to the discourse and even unidentifiable, and therefore it is introduced by the existential marker fi:. Moreover, the participle is not a prototypical predicate being a

---

8 Abul Fadl's data are from another dialect of EA spoken in the rural areas of the Šarqiyya province, but this does not affect the argument here.
non-finite verb form that does not ascribe the action of kidnapping to the 40 robbers, but rather indicates the state-of-affairs that is the result of the kidnapping, which is the main assertion to be made. A formally close translation to English would yield something like 'there are 40 robbers having kidnapped us'. Thus, the utterance is a clear incidence of an event-reporting thetic serving as an explanation why they are where they are, which may be interpreted a comment on a 'stage topic'.

The same is true for (16c), the beginning of a fairy tale, telling us how the wicked enchantress came to know that 'a girl named Rapunzel was born'. Again, as is also illustrated by the English translation, this is a clear case of a thetic utterance. It has been noted that a subject may be heavy and contain additional constituents, such as appositive relative clauses, and still be integrated with the predicate, while a predicate may not (Fuchs 1980, Jacobs 1993). In the EA example (16c) the subject bint 'girl' is further described by a relative clause, which in fact contains a topic-comment construction isma-ha rapinzil 'her name is Rapunzel', but yet the whole construction is perceived as thetic, as articulated in one go and the verb downstepped and downtoned (Figure 1).

Contrary to these, (16d) involves a presentational construction introducing the 'dog', but the dog here is an agentive subject and the predicate consists of a transitive verb 'bite' and its direct object 'man'. The sentence is a translation from English There is a dog biting a man. Just like its English counterpart, this is not a thetic sentence. Firstly, it does neither qualify semantically as a thetic utterance, nor does it qualify prosodically, as every constituent is equally prominent, which adds up to three (focus) accents, indicating three points of information (Figure 2).
Fig. 2: All-FOCUS utterance with three accents, the last one being partially downstepped, but with no accompanying loss in prominence.

Summarizing the discussion of the EA examples, we may note that (16a-c) and (16d) differ in (almost) all semantic, syntactic and prosodic aspects that have been claimed to be characteristic of thetic vs. categorical utterances. Although (16b) involves an agentive subject and a transitive verb, this effect is mitigated by the use of the participle denoting a resultative meaning, thus referring to a state that is the consequence of the action rather than to the action itself. As far as the prosodic realization is concerned, it is highly probable that sarginna 'having kidnapped us' would be pronounced as downtoned. More importantly, the clause may not be split in two intonation phrases, the occurrence of a break between hara:mi 'robber' and sarginna 'having kidnapped us' is highly unlikely. Similarly, there cannot be an intonation boundary between rapinzil 'Rapunzel' and itwaladit 'she was born'. A break and a strong accent on the verb would mean something like 'There is a girl named Rapunzel - and she was born' - which would be rather meaningless. A comparison of the syntactic structures in the French and the EA examples shows that they are parallel, no matter whether we analyse them as split structures involving a presentative relative clause, or if we adopt Wehr's analysis and interpret ce-quí, y'a-qui\(^9\) and fi: as presentational morphemes. EA (just as other Arabic) relative clauses do not have an overt relative marker when the relative clause refers to an indefinite noun. EA also does not require overt subjects. Consequently, the person marking in the verbal paradigm may be interpreted as a suffixed pronoun or as an agreement morpheme in a SVX structure. As a result, fi: kalb (Ø) bi-j-ʕudˤdˤ ra:gil is syntactically ambiguous between (i) 'there is a dog bites a man', (ii) 'there is a dog who bites a man' and (iii) 'there is a dog he bites a man'. We can argue that the above articulation of this specific example without a break (indicated by the epenthetic Schwa-vowels) favours an interpretation as in (i). There are, however, many occurrences of the construction with a clear boundary after the subject and a leading tone on the subject that would favour (ii) or (iii). We will be looking at the different constructions more closely in Section 6.2.

\(^9\) Lambrecht (1988a: 157) notes that il y'a is not to be analysed as a "phrasal idiom" as it can occur in other tenses, such as il y'a eu.
Wehr's analysis also permits another comparison. An interpretation in the sense of (i) makes the structure equivalent to an SVO sentence without a presentational particle, on the condition that the subject carries a closing accent - as in the English example (9a) *her HUSBAND broke his NECK* - that indicates the subject's being in FOCUS. We may argue that the presentational particle foi: is an optional morpho-lexical device that can be dispensed with if FOCUS is marked intonationally. I will not pursue the issue any further for the time being, but we will return to such cases in the next section. Note, however, that I do not agree with Wehr's interpretation that the subject in the split constructions is not a FOCUS, but I agree with her in that it could be a topic (and TOPIC) at the same time (cf. 4.5.1 for discussion). Given the equivalence of the split structure and the 'normal' SVX structure with a closing accent on the subject we could also assume that they have the same information structure and that they should either both be counted as instances of a Sentence Focus, or not. Let me illustrate that point using an example from Austrian German and its French translation. The situation is the following: In a family meeting, one of the grown-up children is not present. The father is being asked why his daughter has not come to the meeting and he answers:

(20) overheard example
a. Why is Laura not here?
   *Najo, weil die Eltern vom Johannes* / *die zwei ins Kino eingladn ham*
   LH   LHL*- L- LHL*-  
   well because the parents of Johannes / the two to-the cinema invited have
   'Well because Johannes' parents have taken them to the cinema.'

b. *Eh bien, il y a les parents de Johannes qui les ont invités au cinéma.*
   (French translation)

Figure 3 presents the pitch track of the German utterance. It shows three accentual peaks, the first one on the introductory *najo 'well' the second peak on *Eltern 'parents' and the third one on *Kino 'cinema'. In accordance with the German prosodic system, every lexical noun is associated with a fully prominent pitch accent, the predicate phrase *ins Kino eingladn ham 'have invited to the cinema' is integrated by deaccentuation of the verbal part and the proper noun *Johannes* is deaccented because the name refers to *Laura's boyfriend who is therefore accessible and may be easily inferred from *Laura who is the main topic of discourse.

![Pitch track diagram](image)

Fig. 3: Explanative utterance from Austrian German, showing closing accents on the main information points.
The interesting point, however, is the shape of the accents and their respective prominences. Both accents are of the closing type that are typical for FOCUS constituents, i.e. Johannes’s parents, although being readily identifiable to all interlocutors and coded as a definite NP, are prosodically realized with a 'FOCUS' accent, mitigating against their status as topics. And indeed, they do not fulfill the role of the pragmatic topic here, being no matter of immediate concern. Their relevance in this situation is limited to their participation in the event that Laura and Johannes have gone to the cinema instead of appearing at the family meeting. In that sense they are part of the reason why Laura and Johannes are not present and therefore they are FOCAL. Considering the sentence without its context, the parents are, of course, the subject on which a predication is made, and thus perhaps some kind of topic, but this interpretation is not realized prosodically in this case. The reason why I adduce this example here is to illustrate the parallelism between German intonation and French and Arabic clefting. Considering (16d), the EA example is even prosodically parallel - with the sole exception that EA verbs are not readily integrated with their objects. I therefore argue that French and EA clefts on the one hand and German utterances as in (20) and English utterances as in (8a) and (9a) on the other hand should be treated the same. But this does not necessarily mean that they are thetic.

Interestingly, Lambrecht (2000: 772, fn. 6) points to this solution himself, noting that he assumes the prosodic structure of the topic-comment clause and the all-FOCUS clause to be the same, but that this may turn out to be a false assumption. Any *systematic* difference would mean that "English has two prosodically marked SF [Sentence Focus, DE] constructions, one for intransitive, one for transitive, sentences. I have already discussed the problem of the systematic difference between TOPIC and FOCUS in 4.5.1 and suggested a solution that assumes that the different realizations point the choice of the speaker which of the pragmatic relations to make explicit – the TOPIC-FOCUS construction as in (8b) and (9b) or the all-FOCUS construction as in (8a) and (9a).

The conclusions to be drawn from the discussion so far are the following: Within the group of *all-FOCUS* utterances, we may distinguish a *thetic* type from a *non-thetic* one on the basis of information structure - i.e. whether an utterance qualifies as thetic at all on discourse-pragmatic and semantic grounds - and, crucially, on the condition that there be some formal strategy of integration, i.e. of removing bipartiteness. An intuitively very clear strategy is offered by prosodic integration. In Chapter 4.5.5 I have suggested that EA prosody offers two types of integration, *integration I* (the combination of a leading and a falling contour without a break between them) and *integration II* (the subordination of one (or several?) accent(s) under the scope of another accent). In the languages discussed here - or at least in the examples discussed so far - it seems, that only *integration II* is correlated with the discourse functions typical for thetic utterances. In EA, for instance, *integration I* and *integration II* are both to be found in nominal compounds, but *integration I* is associated with a bipartite topic-comment structure and is therefore not suited for thetic utterances. It is, however, conceivable that other languages exploit other means of prosodic integration, depending on their prosodic system as a
whole. Furthermore, there are other means of integration mentioned in the literature, such as noun incorporation (Sasse 1987) or discontinuous noun phrases (Schultze-Berndt & Simar 2012). For EA, German and English I take it to be a hard criterion that a thetic utterance be prosodically integrated by integration II. While in German and English this is necessarily related to the use of only one pitch accent, in EA, the lexical/rhythmic accent is not completely deleted in accordance with the prosodic requirements of the language (for a detailed discussion cf. 4.3 and 4.5.5).

If such a distinction is accepted, what then about the rest of the cases that Sasse and Lambrecht subsume under the category of thetic or Sentence Focus? To decide this issue, let us first review the arguments for the existence of a Sentence Focus category put forward in the literature. Lambrecht's (2000) characterisation of Sentence Focus is primarily in opposition to its Predicate Focus counterpart, crucially requiring the absence of a topicality presupposition which he defines as a speaker's assumption "that the hearer considers it a center of current interest in the discourse and hence a potential locus of predication" (Lambrecht 2000: 613). If this assumption on the side of the speaker is formally evoked, the utterance is a case of Predicate Focus and not of Sentence Focus in Lambrecht's terms. In utterances that involve a subject and a predicate, it is thus necessary to code the subject as a non-topic (Lambrecht's Principle of Detopicalization) in order to avoid a topic-comment interpretation. Detopicalization is formally expressed by coding the subject with typical object features (Lambrecht's Principle of Subject-Object Neutralization). A second requirement to assume a 'grammatical' category of Sentence Focus is that its realization be categorically different from the category of Predicate Focus.

Let us now consider the left-over cases that could be members of a putative category of Sentence Focus without being thetic. The first requirement, that the subject should not be coded as a topic, seems to be fulfilled in all examples that have been excluded from the thetic type discussed so far. In all split constructions, the subject is introduced by a marker that diminishes its subject/topic properties. In the EA case, the existential marker turns the subject into a predicate noun, and the same could be argued to hold for the French examples. The same is also true for those cases where the subject exhibits a closing accent indicating FOCUS - the double-peak utterances of English, for example. The second requirement is also fulfilled, as the split constructions as a whole differ from the topic-comment clauses. Only in the case of the double-peak utterances, Lambrecht's criterion that the predicate must not contain an accent, is not met. The best argument for the existence of an independent sentence FOCUS category are cases such as (16d) which crucially differs from a prototypical topic-comment sentence il-kalb / bi-j-ṣudʾdˤa raːgil, 'the dog is biting a man', both morpho-lexically and intonationally, and similar examples of American English substandard varieties, cited by Lambrecht (1988b), such as there was a farmer had a dog. Formally, they are clearly not bipartite and this can be taken as an argument that the FOCUS domain stretches out over the whole sentence. But whether this is enough to suppose an independent category of sentence FOCUS is at least doubtful.
To further investigate that issue, let us now take a closer look at EA split constructions. In (16d), every single constituent is introduced as informationally autonomous. First the dog is introduced as being new to the discourse and its existence is asserted by the existential marker in combination with a closing accent that expressly underlines the assertion. The fi-clause is introducing an indefinite noun into the discourse and is clearly presentationational here. Then the verb and the object are both presented as independent information points. We may take the whole construction to mean something like: What I see is: a man - he is performing the action of biting - it is a man whom the dog is biting. On the other hand, this interpretation is thwarted by the uncontroversial fact that all constituents are uttered in one single IP, thereby indicating that the sentence should be interpreted as one proposition. Furthermore, a topic-comment relation is not indicated due to the lack of a rising contour on the subject. This type of articulation, however, is by no means obligatory. The EA corpus shows many cases of a 'new' subject introduced by the existential marker in an independent proposition (phrased in an independent IP), followed by a comment in a second proposition (another IP), as in the following typical introductory utterance of stories and fairy tales, which is usually preceded by a frame, e.g. ka:n ja-ma ka:n, the equivalent of English 'once upon a time.'

(21) Mido_01_F0_01
be.3SG.M EXIST boy beautiful very very name-3SG.M M
L H LHL L^H^L °H °H- °!HL !HL
'There was a very beautiful boy, and his name was Mido.'

Another frequent possibility is to use one integrated proposition as in (22).

(22) Mido_01_M1_01
kan fi: walad ism-u mi:du
be.3SG.M EXIST boy name-3SG.M M
L H ^H !HL
'There was a boy named Mido.'

The utterances do not differ syntactically, but, as it seems, they differ in terms of their internal information structure. Similarly, a bipartite rendition of (16d) as fi: kalb / bi-j-ʔudˤdˤ ra:gil 'there is a dog, he is biting a man’ would be perfectly natural. It seems that what can be done employing morpho-lexical and morpho-syntactic means in English is done by prosodic means. For instance, the difference between (21) on one side and (22) or (15c) on the other side, is similar to the difference between There was a boy whose name was Mido and There was a boy named Mido. The existence of different constructions used in one and the same discourse situation is not in itself an argument against the existence of different categories. However, I think the proliferation of categories should be avoided if the constructions can be explained otherwise (cf. the concluding remarks at the end of this chapter). In the case of all-FOCUS utterances, I believe that we get by with the assumption of FOCUS and TOPIC as pragmatic relations and simply state that certain discourse situations optionally trigger the occurrence of an utterance that is all-FOCUS, being the comment on an implicit or explicit pragmatic
stage topic (cf. Matić 2003 for a similar view). Depending on the semantic and phonological
complexity of the utterance it may be uttered holistically by way of integration, thereby producing a
thetic statement, or else, the all-FOCUS utterance will be split up into multiple FOCUS domains, with
the optional variant of an internal structuring into TOPIC and FOCUS (cf. 6.2 for more examples). We
have also noted above that integration may be very clear as in the case of univerbation, such as subject
incorporation as in Sasse's example from Boni, to less clear cases such as some of the split
constructions in EA and French. Another area that shows the fuzzy boundary between what is a case
of integration and what is not yet, was suggested to be the variable accentuation of compounds by
Bolinger. Under the assumption that syntactic compounding is a strategy of integration, I suggested
that a compound with two accents - Bolinger's intermediate compounds (Bolinger (1989: 245) - could
be regarded as less integrated (because less lexicalized) than a compound with only one accent. On
that view, we could say that cases like the EA utterance in (16d) are more integrated than the one in
(21). It is not, however, thetic because it does not fulfil the cognitive-semantic criteria for theticity.

Although I am sympathetic to the idea that a construction as a whole may fulfil a different function
from that of its individual components, it seems to me that the occurrence of all-FOCUS utterances
may be interpreted compositionally in most cases. As we have seen above all utterances unequivocally
fulfil the requirement of detopicalisation of the subject. We may take this as evidence for the FOCAL
marking of the subject. In spoken French this is necessarily realized by a presentational construction
with ce or (il) y a, due to the relatively inflexible syntax and relatively inflexible prosody of the
language which mitigate against the occurrence of VS structures (the Italian strategy) and the
occurrence of SAcc structures (the English strategy). I.e. the use of a relative clause is a requirement
of the syntax. It is the only way to express a grammatical predication of a 'subject' that has been
moved out of its canonical subject position. The case of EA is a little different, as EA clauses of that
kind are asyndetic. In EA, it is the intonation of the utterance that decides whether a verb plus its
complement(s) is to be interpreted as a relative clause or a main clause or perhaps only as a predicate
phrase. In any case, it seems that the coding of information as all-FOCUS - as is frequently the case in
out-of-the-blue utterances - strongly depends on the kind of information to be conveyed. If this
information is a simple assertion (or denial) such as in the Arabic example in (3), fi: d‘arb 'there is
thrashing', it is inherently thetic. If the information is complex, containing agents and patients and the
like, it is most likely to contain several assertions expressed in independent - in Jacobs' terms
'informationally autonomous' - FOCUS domains. In the comparably rare case that the information is of
a sort to be integratable, a thetic utterance will arise. On this view, theticity is not a matter of
information structure.

To summarize the foregoing argumentation, a thetic utterance arises in certain discourse situations
where an all-FOCUS utterance is strongly preferred, but theticity is not a pragmatic category. To cut a
long story short: a thetic statement is an utterance whose function is pragmatic, whose definition is semantic and whose grammatical form is integration.

6.2 All-FOCUS utterances in EA

This section will offer a description of those constructions found in the corpus that are typically found in situations evoking an all-FOCUS utterances. The constructions will be introduced and their pragmatic functions will be identified. Based on the theoretical premises laid out in Section 6.1, we will be concerned with the occurrence of FOCI, TOPICS and thetic utterances.

6.2.1 Existential and presentational constructions

The first group of constructions to start with is the family of fi:-constructions that have already been introduced in 6.1. The existential marker fi: is used in a genuine 'existential' function asserting and denying the existence of entities or events. Likewise it is the most frequent device in EA to introduce new referents and new concepts in general, thereby fulfilling a presentative function.

6.2.1.1 The existential fi:

With fi:, a grammaticalized form of the preposition fi 'in' plus a third person masculine suffix (expressed by the lengthening of the vowel), the existence of an entity may either be affirmed or, in combination with the discontinuous negative marker ma:-fi, it may be denied. This entity may be animate (23) or inanimate (24), abstract or concrete (25). Such utterances are thetic by definition.

(23) HRuh_F0_02_01

<table>
<thead>
<tr>
<th>fi</th>
<th>jo:m</th>
<th>min</th>
<th>zat-l-ijja:m</th>
<th>/</th>
<th>kan</th>
<th>fi:</th>
<th>/</th>
</tr>
</thead>
<tbody>
<tr>
<td>in day from self-DEF-day.PL</td>
<td>be.3SG.M</td>
<td>EXIST</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

'Once upon a time there was... (lit.: in a day from the same days')

baker-and fisherman-and shoemaker-and Hosna
H'H'L H L L^HL 'H'L%
a baker and a fisherman and a shomaker and Hosna'

(24) Knee_MG1_17

<table>
<thead>
<tr>
<th>huwa</th>
<th>g-gira:ha</th>
<th>fi:</th>
<th>?awwalam</th>
<th>[giqaari:a:t]</th>
<th>[wuqa?ijj-a]_roc</th>
<th>roc</th>
</tr>
</thead>
<tbody>
<tr>
<td>PRT</td>
<td>DEF-surgery</td>
<td>EXIST</td>
<td>first</td>
<td>surgery-PL.F</td>
<td>preventive-SG.F</td>
<td></td>
</tr>
<tr>
<td>L</td>
<td>H'LL</td>
<td>H</td>
<td>LHL</td>
<td>L^H</td>
<td>!HL^H^L-</td>
<td></td>
</tr>
</tbody>
</table>

'Concerning surgery, there are first of all preventive operations.'

393
In (24), the entity 'preventive operations' whose existence is asserted is only partly new, the topic being 'surgery' in general, coded as a TOPIC at the beginning of the utterance *huwa g-giraha:ha* 'concerning surgery' with a leading accent (LH*L), followed by the marker *fi:* and the FOCUS *giraha:t wuqa?:ijja* 'preventive surgery', with a focus on the adjective *wuqa?:ijja* 'preventive' coded by a double-peak contour (cf. 4.5.4), evoking alternatives for the type of operations only, not for operations as a therapeutical method in general.

By the same token, an event may be asserted or negated, as has already been pointed out by Sasse (1987: 553) and presented in (3) above. Another example quoted by Sasse (1987: 554) is *fi: be:taza:kir* 'tickets sold (lit. there is selling of tickets'. EA, as well as other Arabic varieties, offer the opportunity to express an event in a nominalized form by the so-called *mas?dar* or verbal noun (*d?arb* 'thrashing, *be:t*’selling'). This use is, however, slightly formal, and mostly found in educated speech - note the use of the /q/ in (25) indicating a *wust?a*-form. The example is taken from a TV-discussion between journalists and university professors about the miserable situation of the Egyptian educational system. Contrary to the other TV-programmes in the corpus, which are predominantly informal interviews, this discussion is held in a rather formal register, much closer to MSA than the highly informal interviews of the youth programme *Shabab Beek* or the interviews in the popular morning show *Sabab el kheyr ya Masr* 'Good morning, Egypt', from which (26) and (27) are taken. These examples lack the formal flavour of (25), as the verbal noun *tat?awwur* 'developing' is already lexicalized as an abstractum meaning 'development' and the same is true for *sihr* 'magic' (cf. also *giraha:ha* 'surgery; operation' above). Also note that the word for 'real' is the EA variant *ha?i:qi* in (27), not the MSA *haqi:qi* as in (25).

(25) Mehwar_01_02_03
* ma-fi:*f
  NEG-EXIST
* iltiza:m
  adherence.VN
* bi-qijam
  PREP-value.PL
* haqi:ji:ja
  real-SG.F

'Nobody adheres to true values'

(26) Sihr_MG2_11
* ma-fi:*f
  NEG-EXIST
* tat?awwur
  development

'LH 'H*L%
'There is no development.'

(27) Sihr/MM3_16
* fi:
  EXIST
* sihr
  magic
* ha?i:qi
  real

'Is there real magic?'

It has been claimed that the existential marker *fi:* is only felicitously used with indefinites, on the assumption that the referent introduced by *fi:* must be hearer-new or at least discourse-new. Although this is a frequent use of the marker, there are many examples where *fi:* introduces a given referent as shown by Mughazy (2010). In (24) and (27), the referents introduced are only partly new to the
discourse. In the first case (24), it is surgery that is being talked about and the new referent is a specific subtype of it. The same is true of (27), where sihr 'magic' has been the main topic of discourse (the interviewee being a magician) and sihr haʔi:ʔi 'real magic' is again a subtype of this category. That is, both instances represent a part-of-a-whole relation between the new referent and the old one.

The following dialogue between a pharmacist and a client, cited after Mughazy (2010: 107f.), is a similar example, only that the relation here is an equivalence relation between old and new.

(28)

A: ʕandu-ku  id-dawa da ?
at-2PL DEF-medicine DEM.M.SG
'Do you have this medicine?'

B: la: wa-lʔa:hi xulus⁵ bass fi: il-badi:l il-mistawrad
NEG by-god sold.out but EXIST DEF-alternative DEF-imported
'I am afraid it is sold out, but there is the imported alternative.'

In (28) the introduced referent is even definite, indicating the presumed familiarity of the hearer with the item at hand. As Mughazy (2010) convincingly argues, what is new here is not necessarily the referent itself, but the information, which is equivalent of saying that fi: is necessarily associated with a FOCUS as that part of the proposition by which the assertion differs from the presupposition. Thus, in (28) the open presupposition would be 'there is X', and X is il-badi:l il-mistawrad 'the imported alternative'. Consider now example (24) again: Here the new information was expressed by an indefinite plural noun giraha:t wuqaʔijja 'preventive operations' denoting a concrete countable entity. But the answer could have equally well been fi: ig-gira:ha l-wuqaʔijja (the-surgery the preventive) 'there is preventive surgery', using the abstract noun in a generic sense instead. In sum, it is clear that neither information status, nor the grammatical feature of indefiniteness, are prerequisites for the felicitous use of fi: in its function as an existential quantifier.

My EA corpus also contains a number of such instances, for example the following from a frog story, when after a long story of searching, the frog is finally found.

(29) Mido_F6_01_03

bi-j-busʔsʔ wara j-fagara / laʔa fi: dʔ-d‘ufdʔaʔa
find.3SG.M EXIST DEF-frog
LH- LH⁷ HL H LH^H%
'he looked behind the tree... and there was the frog' (lit.: he found there is the frog)

(29) also illustrates an optional, but by no means infrequent, companion of fi:, the verb laʔa 'find' (see also 30). In the frogstory examples, the frequent use of the verb 'find' is obviously triggered by the activity of searching, but it could also be used in other contexts. Other verbs that may occur in this position are verbs of perception such as simiʕ 'hear' or fa:f 'see'. All of these verbs can be used preceding fi: or also instead of it. Like the existential marker they serve as a strategy to bestow object-
features to the subject, thereby changing it from a prototypical topic to a prototypical FOCUS constituent.

(30) Mido_F6_01_03

laʔa kida fi: hufra f-ilʔardˤ
find.3SG.M so EXIST pit in-DEF-ground

'and there was a hole in the ground'

6.2.1.2 Locative inversion

One construction that was already a matter of much debate in the Arabic grammatical tradition is **locative inversion**. The ancient Arab grammarians called it *xabar muqaddam* 'preposed predicate' and regarded it as a genuine 'detopicalization' strategy. Examples of this kind are legion in CA/MSA (30).

In EA, however, the *fi:*-construction in combination with the locative seems to be used instead, although (32a) is not ungrammatical in EA, (32b) or (32c) are the constructions in use.

(31) CA/MSA locative inversion

*fi*-l-hadi:qat-i kalb-un
in-DEF-garden-GEN dog-NOM.INDF

'In the garden is a dog.'

(32) EA translation

a. *fi*-l-gine:na kalb
in-DEF-garden dog

'In the garden is a dog.'

b. *fi:* *fi*-l-gine:na kalb
EXIST in-DEF-garden dog

'There's a dog in the the garden.'

c. *fi:* kalb *fi*-l-gine:na
EXIST dog in-DEF-garden

'There's a dog in the the garden.'

Note that (32b) and (32c) also differ prosodically. As only the existence of a dog is the new information to be conveyed and the garden is obviously given - an explicitly mentioned stage-topic so-to-say, *kalb* 'dog' carries the strongest accent in both utterances. For (32c) that means that *fi*-l-gine:na 'in the garden' is necessarily downtoned and thus integrated with the noun (see 6.3). If the strongest accent were on the 'garden' in this case, the garden would immediately be interpreted as a *focus*, i.e. contrastively - in the garden as opposed to the house, for example.
An example of fi: + locative from the corpus is the following description of a working place. A girl, working as an entertainer for children, is describing the apartment where she and her colleagues prepare the events (birthday parties etc.) for the children.

(33) Work_F4_01_02
EXIST on-the-left / EXIST drawing.room
'There's on the left ... there's a drawing room.'

The example might lead to the conclusion that the existential marker and the entity it introduces have to be adjacent, hence the double use of fi: - but this is not the case. On the contrary, we frequently find spatio-temporal stage-topics to be inserted in the existential phrase (34, 35).

(34) Work_F4_01_01
fi: ʔawwil ha:ga matˈbax
EXIST first thing kitchen
'There's first a kitchen.

(35) Taxi_F0_01_01
ma-kan-fɔ fi: fi-l-be:t wa-la malliːm
NEG-be.3SG.M-NEG EXIST in-DEF-house and-NEG penny
'There wasn't a penny in the house.'

A special instance of locative inversion is the have-construction of EA. In EA, as in other varieties of Arabic, the possessive construction is not expressed by a verb, but by a locative phrase with the preposition ʕand 'at' or maʕ 'with'. (36) illustrates the use of this preposition in a possessive construction (35a) and in a prepositional predicate phrase indicating a location (36b).

(36)
a. EA locative inversion indicating possession
ʕand-i kalb
at-1SG dog
'I have a dog.'

b. EA prepositional phrase as a locative complement
il-kalb ʕand-i
DEF-dog at-1SG
'The dog is with me.'

The use of the possessive construction can also be metaphorically extended to assume a presentational function. It is not always easy to distinguish between the existential and the presentative, as the following example illustrates. (37) is a sequence from an interview, the well-known actor Hussein Fahmy conducted with the young female journalist and university teacher Marwa Rakha who is famous for her very liberal ideas, the great majority of traditional Egyptians do not approve of.

Concerning marriage, Hussein Fahmy had suggested that Marwa should rather stick to the traditions and marry an 'oriental' man (raːgil faːr̪iː). Then she asks what Fahmy meant by an 'oriental man'. Of
course, the question is not a neutral one, it rather gives rise to the implicature: Why should I do that? What is so special about oriental men? And what does your suggestion imply?

(37) Mar_HF_MM5_FG3_01

MR: huwwa jasni ?e: ra:gil farqi ?
What's that supposed to be, an oriental man?

Q at-1PL tradition.PL and-at-1PL custom.PL and-at-1PL
Well, we have our traditions, we have our customs and we have...

quju:d fi-t-ta'amul-a:t / wi-Sand-ina... /
restriction.PL in-DEF-interaction-PL and-at-1PL
restrictions in our interactions, and we have...

makani it-mar?a / wi-makani ir-ragul...
position DEF-woman and-position DEF-man
'the social position of the woman, and the social position of the man...'

In the above example, the original possessive meaning might be present in the first clauses, e.g. we have (our) traditions, but less so in the final clause 'we have the position of the woman and the position of the man.' The extended use of the locative PP will be further illustrated in the next section when it is used as a mere presentational device (ex. 40).

6.2.1.3 The presentative function of 'fi:' and locative inversion

We have already seen in 6.1 how EA makes use of the existential marker as a presentational device. In this section, we will see some more examples and have a closer look at the different intonation patterns associated with the split constructions. Example (16d) above is a typical case of a brand-new referent that is introduced by the existential marker: fi: kalb 'there is a dog'. So far the function of fi: does not seem to differ from its existential function. At the same time, fi: fulfils a presentative function in introducing an NP to become the topic of the following sentence bi-j-?ud'da ra:gil '(who) is biting a man'. The presentative function is especially clear with unspecific NPs, for instance in the frequent collocation fi: na:s 'there are people', where it is not the function of fi: to assert the existence of people but rather to introduce them as topics (38). Unspecific referents are by definition unidentifiable and therefore do not readily qualify as topics according to the topic acceptability scale (5.8).

(38) Makarona_F0_01
fi: na:s bi-t-hiba t-hut'ti fwaqj malh=um filfil...
EXIST people IND-3SG.F-love 3SG.F-put a little salt-and pepper
'There are people who like to put some salt and pepper...'

(39) Asanser_F0_03_02
fi: na:s kiti:r bi-t-sa: fir il-xa:rig
EXIST people much/many 3SG.F-travel DEF-abroad
'There are many people who travel abroad'

As (38) and (39) suggest, fi: plus an unspecific noun, sometimes modified by a quantifying adjective, is used with the meaning of other quantifiers, such as 'some' and 'many'. The literal translation 'there are people who...' in (37) is synonymous with 'some people'. Thus we could translate (38) as 'some people like to put some salt and pepper' and (39) as 'Many people travel abroad'. There is no real equivalent in EA for the fi:-construction to express 'some X', while 'many X' may be expressed without fi: as in (40). This suggests that the ban against unspecific and indefinite NPs sentence-initially (i.e. in TOPIC position) is less strong in EA than in CA/MSA.

(40) Mar_Shab_MM6_01

na:s kiti:r-a / mayru:r-a / wi-na:s kiti:r-a /
people much/many-SG.F conceited-SG.F and-people many
L- L-0L L-0HH% L- LHL

'Many people are conceited, and many people...

ja-ʕe:n-i / maksu:r-a
VOC-eye-1SG broken-SG.F
LHL LHL%

...are broken, poor things'

In the following example na:s 'people' is specific, marked by the definite marker il-, but unidentifiable to the audience. To be used as a topic the referent has therefore to be introduced in a presentational construction, this time a locative PP is used for that purpose.

(41) Mehwar_FG2_01_06_04

jašni nta šand-ak in-na:s illi ka:n-it hat... taʔri:ban
DM 2SG.M at-2SG.M DEF-people REL be-3SG.F [...] approximately

'There are those who were...like..

bi-ta-dˤu ʔilə harb be:n masr wi-ggazaʔir
IND-3SG.F-call PREP war between Egypt and-Algeria
calling for a war between Egypt and Algeria'

The presentative function is also very clear in the next example, where fi: introduces a definite subject as a new topic. The discussion is about self-confidence and the way boys and girls choose their partners. Up to this moment the three hosts and Marwa were talking about a situation when a man chooses a wife at a very young age and then life changes him, he may start a career and wish another type of partner for himself than the one who was suitable when his personality had not yet been fully developed. In the present sequence the topic changes to the girls, introduced here as contrastive FOCUS, indicated by a strong closing focus accent. 'The girls' (il-bana:t) is introduced by the existential marker fi:. It goes without saying that fi: cannot have an existential function here.

---

10 The literal translation baʔdi: in-na:s 'part (of) the people' is clearly a wusṭa-form that is not used in colloquial speech.
The examples shown in this section illustrate the use of the existential marker as a presentative particle as well as some cases commonly referred to as locative inversion whose sole function seems to be to present a referent and introduce it to the discourse. In Chapter 5.8 I have argued that quantificational expressions may even be used as topics of their propositions. In the next section, we will investigate this suggestion, looking into the prosody of such examples.

6.2.2 The prosody of existentials and presentatives

Now let us consider the prosodic make-up of the various constructions, prosodic annotations have been added in some of the examples. In the existential use, the entity (or event) introduced is necessarily the FOCUS of the utterance and therefore mostly associated with a closing contour. In (23) the list of persons in FOCUS are phrased together in one IP partially linked by default accents, the prefinal one being higher than the ones preceding it. This has the effect of enlarging the fall to the end integrating the final downstepped accent on Hosna into the fall, thereby enhancing finality and assertion. Similarly, a downstep to low is used in (26) with an early peak on the FOCAL item ‘development’. In 4.5.1 it has been shown that FOCUS domains may also have high rising terminals, mostly with a high boundary tone after the H tone on the stressed syllable, if the element at the end of the IP is not the final one in FOCUS. In (29) dafḍaša ‘frog’ is the first item of a list, followed by ‘and his parents... and his siblings’ and thus shows a rising focus accent. In (24) the final accent on wuqaʔiija ‘preventive’ is a truncated fall indicating continuation. In this case the focus is indicated by double accentuation.

Now, what about the intonation of fi: itself? We find that the existential marker is almost always associated with a leading contour. Sometimes it is integrated in the leading part of an accent on the
following element, as in our familiar example (16d) fi: kalbə bi-j-ʕud'də ra:gil, but in most of the occurrences, fi: is associated with a high tone itself. It is not uncommon that fi: is associated with a typical TOPIC contour, i.e. a leading accent with an intonation break as in (23). This intonation is especially common with lists due to a rhythmical preference of splitting up longer utterances into shorter rhythmic chunks. These cases can be interpreted as 'there is/was the following: X (and Y and Z). Consequently I regard fi: as not belonging to the FOCUS domain itself, but rather leading up to the FOCUS as it is typical for genuine TOPICS. The existential marker may be directly followed by the FOCUS as in (23, 24-30, 32c and 33), or it may be followed by another TOPICAL element, mostly a spatio-temporal frame as in (24, 32b, 34-35, 42).

The following example (43) presents a split construction, where fi: introduces an unspecific plural, followed by a predicative relative clause that contains new information. Judging from the context, the only relevant information to be coded as a FOCUS is the NP mafs'al bitaʕ ir-rukba 'the knee joint', whereas the rest of the information is retrievable from the context before. The concept of xufu:na 'arthrosis', particularly 'gonarthrosis' xufu:ni: ir-rukba, is the topic of the interview, and the 'many cases' refer back to a referent that has been introduced in the sentence before, namely the mummies of the ancient days of Old Egypt. The speaker tells us that it has been found that many of them show gonarthrosis, i.e. arthrosis in the knee joint. The contextually determined interpretation is supported by the intonation that shows a leading contour up to the word xufu:na 'arthrosis' and a FOCAL fall on the final NP with higher prominence on the word mafs'al 'joint', the only really new information here. We may thus consider the whole relational presupposition that there is arthrosis in X to be a TOPIC and X to be the FOCUS. Syntactically, the 'subject NP' is introduced by fi: with a presumptive pronoun referring back to it in the relative clause, as it is common with left-detached topics (ch. 5.3.1).

(43) Knee_MG1_04
fi: hal-a:t kit:r-a fi: ha xufu:na
EXIST case-PL much/many-SG.F in-3SG.F arthrosis
H LH LH- aH aH
There are many cases of arthrosis...

fi-l-mafs'al bitaʕ ir-rukba
in-DEF-joint POSS DEF-knee
L in the knee joint.

In Chapter 5 (5.7.1 and 5.8) it was suggested that such quantificational expressions may indeed be interpreted as topics, allowing the paraphrase 'As far as such cases are concerned, many of them exhibit arthrosis in the knee joint.'

We have seen that the NP introduced by a presentative construction functionally may be topic of the relative clause that serves as a comment on that topic. The prosodic realization also supports this analysis - the whole presentative construction is mostly associated with a leading contour and
frequently followed by a break or even the TOPIC marker baʔa as in (44). This analysis is thus in line with Wehr's analysis of French clefts. In a limited number of cases, however, the antecedent NP of the relative clause is not coded as a TOPIC. Among these are the cases of integration identified in 6.1. In these cases, the contents of the relative clause is not construed as a comment on a certain topic, but rather as an event that is the comment on a situational stage topic in which the subject referent is involved. In the following two examples the FOCUS of the utterance is the whole event of Rapunzel’s birth (43, a repetition of 16c) and the passing by of a prince (45), thus being instances of entity-cum-event thetics.

(44) Rap_F0_04
   fi:  bint isma-ha rapinzil itwalad-it
    EXIST girl name-3SG.F Rapunzel be.born-3SG.F
    'A girl named Rapunzel was born.'

(45) Rap_F0_10
   fi: ʔami:r hilu ʔawi miʕaddi [mʕad:i]
    EXIST prince beautiful very pass.by.PTCP.SG.M
    'There was a very beautiful prince passing by.'

In (44) and (45) the utterances are uttered in one IP, which is why in (45) phrasal syncope (the elision of the high vowel in miʕaddi) applies. The tonal coding also suggests one FOCAL domain extending over 'subject' and 'predicate', one of the highest prominences being on the name rapinzil 'Rapunzel' in (44) and on the adverb ʔawi 'very' in (45), falling down to low with a downtoned 'predicate' (relative clause) that is associated with a low contour (or at least a partial downstep in (45)) and lower prominence. The pitch track of (44) was presented in Figure 1 above; the intonation contour of (45) is illustrated in Figure 4.

![Fig. 4: Pitch track and intensity tracking for the integrated FOCUS: fi: ʔami:r hilu ʔawi mʕaddi 'There's a very beautiful prince passing by.'](image)

6.2.3. **Experimental data**

In this section, I will only shortly point to a number of cases that have been discussed by Hellmuth (2010) as they are interesting for their syntactic behaviour. The data investigated by Hellmuth was 402
collected within the SFB 632/D2 project, part of which is also included in the corpus used for the present investigation (see 4.2). Hellmuth analysed 128 potential thetic utterances to be elicited in an all-new context (10 Event Cards). She found three strategies that were employed to produce all-new statements: i) existential-constructions (with fi:), ii) deictic presentational constructions (using da 'this, M', di 'this, F' and do:l 'these') and iii) "monoclau sal (indefinite subject is clause-initial)" (Hellmuth 2010a: 263). As Hellmuth (ibid.) notes, only 6% of the felicitous tokens were produced with the existential marker which, according to Sasse (1987), would be the expected construction in an event reading. This outcome is not at all surprising, given the experimental design that was used. The participants in the experiment were shown individual pictures that illustrated a particular scene. When looking at the picture, they were instructed to respond spontaneously to the question "What's happening?" As the results suggest, however, speakers simply described what they saw. They did not imagine themselves being in the situation and reporting an event in a holistic manner as a would-be thetic requires. What they did is describe a picture, they saw a person that was fulfilling an action and so in most of the responses, first the person was identified and then the action ascribed to that individual this person - which yielded a regular topic-comment sentence. This also explains the high number of deictic presentationalss as in example (46).

(46)

a. 27ARZ-A03ANT-01M0900
   da walad ?a:ʕid ?udda:m il-kumbju:tir
   DEM.M boy sit.PTCP.SG.M in.front.of DEF-computer
   'This is a boy sitting in front of a computer.'

b. 27ARZ-A01ANT-01M0900
   da be:t / il-be:t bi-jii-thiri?
   DEM.SG.M house DEF-house IND-3SG.M-burn.down
   'This is a house. The house is burning down.'

The strategy employed is particularly clear in (45b), where the introduction of the new referent and the predication about him is actually done in two completely independent clauses with the FOCUS of the first clause repeated as a ratified topic in the second clause. But also (46a) is not a thetic utterance as I have argued in Section 6.2.1. Obviously the attempt to elicit an equivalent to the well-known English thetic 'The HOUSE is on fire' was doomed to failure. What we get here, is the typical split construction illustrated with the presentative particle fi: and with the locative PP ʕand 'at' + personal pronoun. In (46) and many other of these cases it is the demonstrative instead, consistent with the task of picture description. This is to say that the utterances elicited by the experiment are not supposed to be thetic at all. However, the referent of the predication has to be introduced first. As a result, the utterances using a definite NP are comparatively rare.

---

11 Hellmuth's study comprised 128 utterances. The data used here is only a subset of that data.
The interesting cases I would like to point to here are the ones classified as “monoclausal” by Hellmuth. As Hellmuth (p. 263) reports, this is the largest group of responses, making up for more than 50% of the total number of utterances. The syntactic structure of these sentences is the same as in the cases with a presentative particle, the only difference being the lack of such a particle. Consider, for example, the utterances in (47a) and (47b); the latter is an equally felicitous response in the context at hand, and no strong pragmatic difference can be discerned between the two. That means that the syntactic analysis is again ambiguous between mono-clausal and bi-clausal split construction, in the same way as already discussed in connection with example (16d) above. We can therefore also interpret such utterances as responses to the implicit question 'What do you see?' which could be answered with a deictic presentative (to use Hellmuth's term) or in other words, with an explicitly expressed ratified topic da X 'this is a X' or without it, simply as X 'an X'. In English we would rather expect the shorter ('elliptic') response. And interestingly, this also seems to be the case in EA. As already noted above, this fact shows that the ban against clause-initial indefinite NPs is not active in EA.

(47) 27ARZ-A06ANT-02F1000

a. 
\[a::: \text{walad } \text{bi-ji-nzil} \quad \text{fi-l-pisi:n}\]

HESIT boy IND-3G.M-go.down in-DEF-pool

'... ah, a boy (who is) stepping into the pool'

'... ah, a boy is stepping into the pool'

b. 
\[\text{da} \quad \text{walad } \text{bi-ji-nzil} \quad \text{fi-l-pisi:n}\]

DEM.SG.M boy IND-3G.M-go.down in-DEF-pool

'this is a boy (who is) stepping into the pool'

The decision concerning the internal information structure in this type of construction can, as I have suggested above in the discussion of examples (21) and (22), be taken with the help of prosodic coding. For that purpose, we will consider the utterances produced in another experiment from the QUIS corpus. In experiment (13 Animal Game), the participants also had to describe cards, but in this case the referents were either new or given and likewise were the properties associated with them, i.e. the pictures, for instance, showed a red horse bending down, followed by a blue horse bending down. The aim was to see whether givenness of referents and properties would be coded in the language at hand. Here we will only consider those utterances that contain all-new material. In (48), the speaker first introduces the new referent, the bird (ʕasˤfu:ra) using an indefinite NP and a focus accent (very high rise with a boundary tone), because she is going to continue her utterance with the property of flying tˤajra that is predicated on the subject referent and which again carries a leading focus accent, followed by an ordinary topic-comment clause lon-ha ?azra? 'her colour (is) blue. Notice the high rise of the FOCAL accent and the moderate rise of the TOPIC lonha 'her colour', the final FOCUS - the colour blue - is then associated with the final low fall, indicating assertion and finality.
Fig. 5: An all-new utterance in response to a picture prop showing a big blue bird. (The high frequency at the end of the utterance is an artefact caused by the glottal stop. The final L is in fact very low.)

The event card game also triggered similar utterances such as the one in (48), illustrated in Figure 6. Again, we see that the utterance is divided into several IPs (breaks are indicated in the third tier in Figure 6). Every IP closes with a focus accent indicating the relevance of the referent. The first referent introduced is wahda bint, ‘a girl’ and the predicate is rakba hus'a:n ‘riding on a horse’, all constituents are phrased as single IPs and all of them are high rising indicating continuation, and involving focus accents indicating their high information value, while the final IP is associated with a closing contour. What is particularly noteworthy here is the high prominence associated with the final accents in the individual IPs. The medial IPs rakba hs'a:n, bitim'fi fi:, and bitim'fi bi: not only show high intensity values, but more crucially scaling differences between the accents. In final position, however, where the fall is obligatory, the lower scaling is counterbalanced by very high intensity on the final item. Compare the intensity of the final two accents in Figure 6 with the final two accents of the thetic utterances in Figure 4 or in Figure 1 above. The difference is clear evidence that intensity values play a major role in the perception of prominence, contrary to a widely held assumption that it is only pitch that matters. The whole utterance is very rhythmic containing two accents per phrase with the second one being more prominent and is a typical example of a staged descriptive event, when the speaker codes every individual information point on its own. Such utterances are quite unlikely to appear in naturally occurring speech.

(49) 27ARZ-A09ANT-03M1100

wahda bint / rakhi-ra / hus'a:n / bitim'fi / FL:-Ø / one-F girl embark.PTCP-SG.F horse IND-3SG.F-go in-3SG.M

'A girl, sitting on a horse, riding IN it, ...
IND-3SG.F-go  PREP-3SG.M  middle  DEF-woods
riding ON it, in the middle of the woods.‘
(The fourth IP is a correction of the third one).

Fig. 6: An all-new utterances as a response to a picture prop showing a girl riding on a horse, and the question ‘What’s happening?’

We have seen in this section that EA noun-predicate constructions with an indefinite noun are syntactically ambiguous between an NP followed by a non-restrictive relative clause, a monoclausal subject-predicate interpretation and a multicausal interpretation consisting of individual clauses with only one constituent each as in (49). In such cases the first clause only contains the argument to be introduced, which is the implicit topic of the second clause which in turn contains the predication on that topic. Evidence for this analysis comes from prosody, (1) from phrasing and (2) from type of accents involved. As illustrated in Figures 5 and 6 all individual FOCUS domains exhibit a focus accent, the prefinal ones a high-rising leading accent and the final one a low-falling closing accent. The high rise has been identified as a typical focus accent in FOCUS domains in Chapter 4.5.

6.2.4 All-FOCUS utterances with the discourse particle ‘da’

In 6.2.3 we have encountered the deictic presentational construction involving the masculine singular demonstrative *da* with its feminine counterpart *di* and the plural *do:l*. In the cited examples the demonstrative is a real pronoun with deictic reference and therefore has derived forms inflected for gender and number. I have argued elsewhere (El Zarka 2012) that *da* has been grammaticalized to a discourse marker conveying a certain modality and at the same time indicates that the utterance as a whole is in FOCUS. In that sense, the discourse particle *da* presents an utterance as a whole, on a par with the existential marker which presents a referent. Such an analysis has been offered by Ouhalla (1997) for the CA particles *ʔinna* and *qad* that ‘reinforce/confirm’ the propositional content of a given sentence’ (Ouhalla 1997: 21). Similarly, *da* enhances the whole predication, but diminishes the
topicality of the subject. It is usually only employed in narrative, not in qualitative predications that ascribe a property to a topical referent. In (50) the particle *da* is used to express the immediacy of the event in the situation at hand. The speaker is imagining an egg that has been boiled and describes the situation before the egg was boiled and after that. She describes the raw egg imagining that she was holding it in her hand. If she had simply ascribed the property of making a noise to a generically used topic 'the egg', the use of *da* would not be felicitous. The meaning of *da* is difficult to pin down, but it clearly has an assertive function, also conveying some sense of unexpectedness. Depending on the context, it can be roughly paraphrased as 'guess what...', 'what a surprise', 'that's what happened' or 'believe me' or maybe 'actually'.

(50) Marwa_SV_FM4_14

\[\begin{array}{llllll}
\text{da} & \text{ka:n-it} & \text{kida} & \text{bi-ti-smil} & s'\alpha:t' \\
\text{DM} & \text{be-3SG.F} & \text{so} & \text{IND-3SG.F-make voice} \\
\end{array}\]

'It made such a kind of voice.'

It is important to note that the particle *da* is not primarily a device to indicate information structure, it rather is a modal discourse particle that conveys a certain pragmatic nuance, and incidentally produces a certain information structural effect. This is in line with Ouhalla’s observation that “given that both focus and modality convey extra-propositional information, in the sense that they are both affective operators, it is not surprising that they interact with each other in the way they seem to do [...]” (1997: 21).

6.3. The intonational origin of Prosodic Inversion

In this section, I will present a hypothesis about the possible origin of the thetic SAcc construction, which I believe to be rooted in two cross-linguistically widespread intonational features: an utterance-initial exclamatory accent and downstep. I will particularly argue that the Principle of Paradigmatic Contrast advocated by Lambrecht (2000), of which Prosodic Inversion is a subcategory, is not the motivation for the SAcc construction to arise. Lambrecht's argumentation for an independent Sentence Focus category (see 6.1) is partly based on the observation that what he calls Predicate Focus has an accent on the predicate and what he calls Sentence Focus is accented on the subject. This dichotomous prosodic surface, observable in English or German, for example, is paralleled by 'syntactic inversion', for example in Romance languages. Both dichotomies are assumed to be manifestations of Predicate vs. Sentence Focus structure, as illustrated in Lambrecht (1994: 319; ex. 5.79):

(51) Predicate Focus     Sentence Focus

| Italian | S\(P\) | P\(S\) |
| English | S\(P\) | \(SP\) |
Lambrecht's *Principle of Paradigmatic Contrast* as a functional motivation and the term inversion itself suggest a deviation from the 'basic', 'neutral' or 'canonical' structure for a specific functional purpose, i.e. an arbitrary grammatical choice. Although the account is appealing, my objection is primarily to the idea of *motivation*. As already noted above, I am not concerned with word order here, so my arguments are exclusively based on prosodic facts. Based on data from EA, I will try to show in the following paragraphs that prosodic integration as observed in thetic utterances cited in 6.1 and 6.2 is related to the choice of a certain intonation pattern. The fact that a paradigmatic contrast incidentally arises may help with the grammaticalization of such a construction, which may actually be on its way in West-Germanic languages. It will have become clear by now that in EA there cannot be genuine inversion based on accent placement due to the prosodic properties of the language. As I have shown in Section 6.1, the predicate is not deaccented in EA thetics, but only downtoned.

To collect data on thetics is not a trivial task. The EA corpus collected here is too small to contain enough thetic utterances to make any quantitative statements. In the spontaneous and read data, a total of 126 utterances were labelled as thetic, around half of which were of the *fi:*-type, the other half contained different constructions, mostly VS structures, locative inversions and simply one-term utterances. The case of an argument-cum-predicate thetic is, however, a rare incidence. We have also seen that the QUIS questionnaire was not successful in eliciting thetic utterances. What speakers did was simply to describe the pictures they saw instead of imagining the situation and relating an event. In addition to the handful of entity-cum-event thetics in the corpus, such utterances were elicited in a small experiment with staged dialogues. 14 native speakers of EA (10 female and 4 male), all aged between 20 and 40, participated in the small staged dialogue. This procedure was chosen in the hope that the utterances would be to some degree natural. Therefore the speakers were only instructed about the contents of the dialogue, but were not given a scripted version of what they should say. The intention was to yield thetic statements with an explanatory function, following Sasse (1987), such as 'My CAR broke down' or 'The BRAKES are broken'. This relative freedom unfortunately resulted in a variety of different utterances, many of which were of the categorical type. Nevertheless, some of the utterances show the expected pattern of 'inversion' in the sense that the prominence relation between the subject and the predicate in an SP structure is opposite to that of the typical topic-comment clause. The important difference, however, is tonal, as predicted by the results of the TOPIC-FOCUS study presented in Chapters 4 and 5. The dialogues were digitally recorded and submitted to qualitative analysis with PRAAT. The two utterances that appeared with some frequency in the conversations are presented in (52). To test the potential difference between a thetic and a categorical statement, the participants were interviewed after the recording of the dialogue and the arguments that were supposed to have appeared in a thetic utterance in the dialogue were the topics of the questions in the interview. The results of the experiment were described in El Zarka (2011).

(52)
a. potential thetic: My CAR broke down.
ili-ṣarbijja ṣat'la-n-a
DEF-car out.of.order-SG.F
'The car is not working.'

b. potential thetic: The BRAKES are broken.
il-fara:miš bājz'-a
DEF-brakes broken-SG.F
'The breaks are broken.'

(52a) was a response to the question ‘Why can’t you come to Alexandria with us next weekend?’ and (52b) was an answer to the question ‘What is wrong with your car?’ Some of the car-statements showed the stronger prominence on the subject, but the predicate was also accented, albeit less than in a comparable categorical statement. The perceptual impression is supported by the intensity tracking that show the drop in intensity from the first to the second accent (Figure 7). In the example illustrated in Figure 7, the pitch contour shows an excursion on ṣarbijja and a flat structure, i.e. downstep on the word ṣat'la:na. One explanation for the noticeably audible prominence of the predicate might be the phonological weight of the three-syllable word ṣat'la:na with a stressed syllable containing a long /a/. This makes it rather hard to deaccent the word in EA, so there is still a fair amount of prominence perceivable. Figure 8 shows a typical categorical statement for comparison. The example contains three successive pitch accents with wide excursions, the last one of which is perceived as the main prominence in the utterance, thereby exemplifying typical default topic-comment prosody.

![Figure 7: F0 and intensity tracks of the thetic utterance il-ṣarbijja ṣat'la:na (in response to ‘Why can’t you come with us to Alexandria next weekend?’ with the heavy line representing pitch, and the thin line intensity.](image)

![Figure 8: F0 and intensity tracking of the thetic utterance il-ṣarbijja btašīk bājz'-a! (an echo-utterance upon the other interlocutor’s ‘My car is out of order’ with the heavy line representing pitch, and the thin line, intensity.](image)
The *fara:mil*-utterances are clearer, they show a sharp decline in prominence from subject to predicate. The examples illustrated in Figure 9a and 9c both have a marked accent on *fara:mil* and a downstepped predicate *bajzˤa*. Note that the presence of downstep alone cannot be held responsible for the difference in prominence, as downstep also occurs in topic-comment clauses (9d). The fall of the closing accent illustrated in Figure 9c covers more than 7 semitones. Although a large excursion can also be observed in non-FOCAL positions, this pertains to the rising part of the accent and not to the falling one. The instances of the topic-comment counterpart investigated in this experiment exhibited only a drop of 2-3 semitones. What is important here is not the excursion itself, but the steepness and abruptness of the fall after the peak. Note that the subject accent in (9b) is predominantly rising followed by a slight downstep of the predicate accent. In (9d) the downstep is much stronger. Pitch, however, is not the only phonetic feature that serves as a cue in these examples. Again there is a difference in intensity which is particularly clear in (9c) as opposed to (9d).
A third acoustic cue associated with prominence is duration. In all examples there was a difference in duration between the thetic and the categorical types. As already noted, examples were not enough to obtain quantitative results. Therefore I can only give absolute values for illustration. The greatest difference in duration between subject and predicate in a thetic fara:mil-statement within the corpus was 622 ms (subject) vs. 276 ms (predicate), i.e. a ratio of approximately 2.25:1. Whereas a typical categorical statement showed a relation of 562 ms vs. 493 ms which constitutes a ratio of only 1.14:1. Although these results would have to be verified by a large-scale experiment - i.e. if a reliable methodology can be found to elicit thetics - that produces enough instances of thetic vs. topic-comment utterances to be amenable to statistical analysis, the measurements made in this handful of examples correspond to the features identified in the FOCUS experiment reported in Chapter 4.4.1.

In addition to the prosodic cues, there are also segmental cues involved. As has often been pointed out in phonetic studies, there is more accuracy in the realization of segmental material in FOCUS constituents. When under FOCUS, the determiner il- in initial position is articulated with glottal stop epenthesis as [ʔil] whereas determiners in non-FOCAL constituents are sometimes deleted completely or at least articulated in a reduced manner as [l] or as a slight lengthening of the initial consonant, depending on the morphonological environment.
One issue remains to be addressed here: There is a possible alternative analysis of the *fara:mi*l-sentences as subject FOCUS utterances. The predicate *bajz*:*a* can be viewed as given information, as it has been noted that the car is 'broken' or 'out of order' before, which turns the denotation of *bajz*:*a* into given information. In any case, we have seen that a FOCUS domain may also contain given information. And thetic utterances always bear a strong similarity to subject FOCUS utterances. Thus, the question *il-*ará*bijja fi:*ha ?e:*h? ‘What is wrong with your car?’ could also be understood as a query to identify the ‘wrong’ or broken part of the car. But the utterance is, at the very least, ambiguous between the two readings, and the possibility of a thetic interpretation.

We have also seen some examples of thetics in the spontaneous data and a few more will be presented here. (53) is a spontaneous utterance from a conversation in an Egyptian home. The host had asked if anybody wanted to have tea, and when all denied, she made the comment cited in (54) and illustrated in Figure 10. (53) is from a translation task of the SFB 632-D2 corpus. The utterance is similar to the one cited in (16c) above and thus strengthens the evidence that this is the typical way such an utterance would be articulated (see Figure 11).

(53) conversation data: speaker F0  
*fi:* hadd *ša:wiz fa:*j   
EXIST someone want.PTCP.SG.M tea  
'Would anyone like to have tea?'

*t'ab kwajjis *šafan *il-*bojlar *ša'la:n  
ok good because DEF-boiler broken  
'Well that's good because the water kettle is broken'.

![Fig. 10: thetic utterance with a closing accent on the subject and a low tone predicate.](image-url)
stimulus
What happened? - A child was born.

translation
fi: ʕajjil itwalad
EXIST child be.born.3SG.M
'A CHILD was born.'

Fig. 11: Thetic utterance with a closing accent on the subject and downstep on the predicate.

The final pair of examples I will present here is from an elicitation session with two female Egyptian speakers (F0 and F1). The task was to read the following short dialogue (55) and to answer the questions in (56). The dialogue and the questions were embedded in a dialogic reading task with other unrelated dialogues where they occurred twice. After a short pause, the reading task was repeated with the two speakers changing roles. The results were not homogeneous. The referent 'mum', being existentially presupposed, was coded as a topic three times and only once as a FOCUS. But the intonational difference between this thetic utterance and the one with a ratified topic (56a) and another one with a contrastive narrow subject FOCUS (56b) from another dialogue is very clear (see Figure 12).

(55) elicited dialogue
A: ʔe: raʔik niru:h is-sinima mnharda?
   'How about going to the movies today?'
B: la  mif  ha-ji-nfaʕ  ma:ma  gajj-a
   no NEG FUT-3SG.M-be.useful mum come.PTCP-SG.F
   'No, this won’t be possible – Mum’s coming.'

(56) elicited question-answer pairs

a.
A:  ma:ma  fe:n?
   'Where is mum?'
B:  ma:ma  GAJJIA
   'Mum's COMING.'

b.
A:  nell  gajja?
   'Is Nelly coming?'
B:  laʔ,  MA:MA  gajja.
No, MUM'S coming'

![Fig. 12: Three pitch contours of a thetic clause (panel a), a topic-comment clause (panel b) and a contrastive subject FOCUS (panel c)](image)

The data I have presented here is only a first step in the investigation of thetic utterances in EA. Due to the comparatively low number of thetic utterances found, the results are not as robust as the ones on TOPICS and FOCI reported in the preceding chapters. Nevertheless, it is fair to say that prosodic correlates for thetic utterances could be found - these can be summarized as follows:

- The theticity effect arises when in the appropriate context and in a sufficiently short utterance the second accent is downtoned making integration possible.
- The acoustic correlate of the prominence relation involves several interacting phonetic features which may not all be present to the same extent in one single utterance:
  - wide pitch excursion in the first accent on the subject with a steep fall on the accented syllable
  - (partial or total) downstep of second accent
  - frequently early alignment of the H-tone within the accented syllable of the subject or before
  - longer duration of the whole subject constituent and more articulate pronunciation of the segmental material
  - drop in intensity from first to second accent
  - integrated contour with the 'head' of the contour to the left

Thus, a FOCAL accent on the subject in combination with a weaker accentuation of the predicate enhances the monomial character of the thetic statement in EA in the same way it has been shown for other languages. The resulting prosodic construction runs counter to articulations with two strong accents that are prone to separation by a phrase boundary. It is in particular different from the typical topic-comment utterance in which the accent on the predicate is the more prominent one and whose subject accent is either linking or leading. It is, however, not at all clear whether EA can be said to have a clear strategy to mark thetic utterances by some version of SAcc prosody. I rather suspect that
this is not the case. At least, it is not the preferred strategy to express theticity in the language.
According to Sasse (2006: 268), the SAcc construction is only marginal in some languages, among them Italian, Modern Greek, Rumanian, Russian and Serbo-Croatian. Two remarks are called for here. First, all languages mentioned - including EA - have the possibility of VS word order. Second, I have already pointed to the similarity of Italian and EA as far as the deaccenting is concerned. As the investigation of my EA corpus and the additional experiments and elicitation attempts suggest, the prosodic marking of theticity can at best be considered a marginal strategy, if it can be called a strategy at all. I strongly suspect that only languages that do resort to deaccenting within a FOCUS domain, such as German and English, have the prosodic basis to grammaticalize the SAcc construction. Recall that in EA, there is strong downtoning, if not deaccenting, of presupposed material outside the FOCUS domain, but given items that are in FOCUS are not normally downtoned. In this respect, the marginal incidences of thetic utterances shown in this chapter are an exception.

But if thetic prosody is not the result of a willingly employed strategy, why do we come across such examples at all? I believe that there are reasons, all of which have to do with certain intonational choices that are ultimately and jointly responsible for the resulting thetic intonation. First, the subject in such cases is in FOCUS and therefore is likely to have a closing accent associated with it. In a typical topic-comment sentence the subject accent is linking or leading and the FOCAL closing accent is the stronger one of the two, especially if the topic is not contrastive and does not have to be activated. If there is a closing accent on the subject, signalling its being part of the FOCUS, the predicate may as well have a prominent closing accent - perhaps with a high peak. Such cases can be found in the specific contexts that are assumed to trigger thetic utterances. However, the predicate may also be downstepped. Let us now consider the specific cases of thetic utterances found in the corpus. Take first examples (53) and (16c). The predicate is downtoned with respect to the subject, thereby giving rise to integration and the establishment of a semantic unity between the two items involved. As already noted above, had the speaker in these cases chosen a second strong closing accent - as in the dog-biting-a-man example in (16d) - probably with a pointed hat, the semantic implication would be rather odd, as it would mean something like 'there is a baby and it was born' or 'there is a girl named Rapunzel and she was born.' This is why the speaker chose a downstep accent here and further reduced its prominence.

Now let us turn to example (53): 'The water kettle broke down / is broken.' Although making tea involves the necessity of a water kettle, which in turn suggests that the water kettle can be regarded as concept given in the present situation, the speaker chose to present it as 'new' information employing a closing accent on the subject only and integrating the verb with the subject under the scope of that accent. This is possible because the conceptual unity of a broken kettle is an option. Thus the sentence has the open proposition '[It would have been difficult to make tea anyway] because of X', with X
being the broken tea kettle. The choice of downstep here has a humoristic effect, it plays down the information as something not worth mentioning, putting it as an aside.

Having identified the formal properties and the semantic constraints in the our examples, we will finally take a look at their pragmatic function of thetic utterances as identified by Sasse (1995), the main five of which I reproduce here in Table 1 together with the semantic areas that are commonly associated with them:

<table>
<thead>
<tr>
<th>DISCOURSE FUNCTIONS</th>
<th>ASSOCIATED SEMANTIC AREAS</th>
</tr>
</thead>
<tbody>
<tr>
<td>introductive</td>
<td>existentials + indefinite animate subjects</td>
</tr>
<tr>
<td>annuntiative</td>
<td>appearance and disappearance, beginning, ending; mishaps, gleeful news</td>
</tr>
<tr>
<td>interruptive</td>
<td>sudden events (phone ringing, door opening), appearance</td>
</tr>
<tr>
<td>descriptive</td>
<td>meteorological expressions, existentials with natural phenomena as subjects</td>
</tr>
<tr>
<td>explanatory/elaborative</td>
<td>(open)</td>
</tr>
</tbody>
</table>

Table 1: Discourse functions of thetic utterances (Sasse 1995: 24)

The first function is to introduce new referents into discourse, as a result it does not involve the case of event reporting that is of interest for us here and may thus be neglected in the following considerations. As we are only concerned with the prosodic shape of such utterances, let us now consider what attitudinal and emotional characteristics are associated with the individual discourse functions, on the assumption that, following Bolinger, the attitudinal and the grammatical aspects of intonation go hand in hand. In this respect, the functions can be divided into two major types: (i) The first group involves a decisive attitude and a matter-of-fact connotation which particularly pertain to the explanatory, but also partly to the descriptive function. (ii) The other group comprises the annunciative and the interruptive functions. These may also be expressed as matter-of-fact statements, but it may likewise be exclamative, expressing surprise and suddenness. A statement such as 'The SUN is shining' may thus be uttered as a barren statement about the weather conditions or as a joyful exclamation. The descriptive function is the most elusive of the functions as it only involves a very limited semantic area (weather expressions and natural phenomena) and many utterances of that kind may be classified as annunciative as well, whereas others are not even regarded as thetic by Lambrecht (2000), e.g. German es regnet or its English counterpart it is raining.

I will specifically argue in the following paragraphs that it is exactly the attitudinal aspects of intonation that are responsible for the emergence of the SAcc construction as an expression of
theticity. If this is correct, it comes as no surprise that SAcc constructions predominantly occur in three discourse situations: interruption, announcement and explanation as summarized by Sasse (2006: 293) where either one or both attitudes are to be expected.

<table>
<thead>
<tr>
<th></th>
<th>(X)VS</th>
<th>SAcc</th>
<th>Split</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introductive</td>
<td>+</td>
<td>−</td>
<td>−</td>
</tr>
<tr>
<td>Descriptive</td>
<td>+</td>
<td>(+)</td>
<td>−</td>
</tr>
<tr>
<td>Interruptive</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Annuntiative</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Explanative</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>

Table 2: Range of functions of alternative constructions (after Sasse 2006: 293)

Let us now look at the prosody of exclamation and matter-of-fact. The intonation of exclamatory declaratives is usually associated with a high fall in (American) English (Bolinger 1998: 51). The high fall will also frequently be associated with the beginning of the utterances as in the examples below. It has already been noted by Jacobs (1988: 115) that the emphatic accent in examples such as (57b) have nothing to do with focus marking, but are typical exclamatory accents.

(57) English
   a. What a WONderful DAY!
   b. How NICE he is!

(58) German
   a. Was für ein WUNderbarer TAG!
      what for a wonderful day
   b. Bist DU aber schmutzig!
      are you dirty
      How DIRTY you are!

(59) Egyptian Arabic
   a.
      \paper{AMma} hittit FAS\textsuperscript{a}L/fas\textsuperscript{a}l
      PRT piece trick
      That’s far out!
A look at the prosody of these examples is all it takes to note the striking similarity of these patterns and the SAcc pattern. The strong accent early on in the utterance is a high fall followed by a downtoned second accent or even by deaccenting. This is even true for EA. Now consider typical eventive (all-FOCUS) utterances cited in the literature. They all make perfect exclamations.

(60)

a. The HOUSE is on fire!
b. Your EYES are red!
c. Your EYES are blue! (Adam to Eve upon first seeing her) (Gussenhoven 1983: 395)
d. The PHONE's ringing!

A less compassionate speaker could also say these utterances in a more detached manner as a simple matter-of-fact statement. The intonational difference between those two is gradient, the exclamatory statement has a higher peak and thus a wider pitch range than the simple declarative. But both have a closing accent on the subject. The detached matter-of-fact quality of downstep can be observed in another frequently cited example from Schmerling (1976) that initiated the whole debate of eventive 'news' sentences and their prosodic shape (61a).

(61)

a. JOHNSON died. (routine announcement in the news)
b. ?LYNDON died. (fictional announcement made by Johnson's wife upon his death)

Bolinger (1989: 243) makes the interesting remark that Johnson's wife could hardly have announced the death of her husband in the same detached manner as suggested in (61b). What Bolinger points out is the the connection between downstep and a detached matter-of-fact rendition of an event associated with it. The choice of downstep in this case makes the event of dying a routine issue, which is why we would perceive a thetic statement coming from the wife of someone who has just died as rather cold and inadequate.

The upshot of this digression into the world of paralinguistics is to show the connection between the intonational features of downstep and exclamatory accents and the attitudinal aspects of thetic eventive statements. In terms of acoustic phonetics, downstep with some prominence preserved may be regarded as a step towards total deaccenting, the tonal configuration of total downstep and deaccenting being essentially the same. Thus in EA, we may observe the same motivation, the need to mark the subject as FOCAL and to express detachment or surprise or suddenness. The resulting pattern then has the effect of a thetic utterance, albeit maybe not as clearly as in the well-known German and English examples where paradigmatic contrast is perceived stronger as the prominence relations have become
a matter of accent placement. It is in such cases that the issue of paradigmatic contrast becomes important. If the SAcc construction in English and German is indeed a grammaticalized linguistic pattern or on its way to become one, then the path of grammaticalization could have been as schematically presented in Figure 13, from tonal downstep (on the left) to deaccenting (on the right), which in turn leads to a paradigmatic contrast of left- and right-headed accent domains, facilitating the linguistic exploitation of this difference.

Further evidence against the assumption of Prosodic Inversion as a motivation comes from the syntactic variety that may be encountered in 'thetic' utterances across languages. I only give two examples here from German and EA to illustrate that point. Notice that the function of the sentence in (61) is interruptive and announciative at the same time. The utterance occurred in a lecture on the history of the Palenqueros, a creole-speaking people in Columbia of African origin, when the speaker Armin Schwegler pointed to a major turning point in their history.

(62) Armin Schwegler, (Swiss) German (15-04-2013)

Und Dann kam der große Wechsel.
and THEN came the big change

The utterance is all-FOCUS with no topic selected and it is clearly eventive. But it does not exhibit one of the typical identified constructions, but rather a hybrid form consisting of a (X)VS clause with a SAcc intonation pattern. Thus it cannot be paradigmatic contrast that produces that pattern, but rather the exclamatory intonation pattern described above.

Similarly, the EA example in (62) does not involve an accented subject and deaccented predicate, but rather an accented predicate and deaccented object, which suggests that the primacy here is the intonation pattern, not the inversion. The FOCUS domain again is the whole sentence, although there is not a single 'new' entity or event denoted. We have already discussed this example in Section 4.5.5 in connection with integration. I will repeat part of the example here as (63) together with the pitch contour (Figure 14).
(63) Rap_F0_22

ja rapinzil / ja rapinzil / nazzil-i fašr-ik
VOC R VOC R let.down.IMP-SG.F hair-2SG.F
o Rapunzel, o Rapunzel, let down your hair

rah-it MI-NAZZIL-A-L-HA fašr-aha /
go-3SG.F PTCP-let.down-SG.F-to-3SG.F hair-3SG.F
so she LET down her hair for her.’

Fig. 14: Thetic utterance indicating verum focus with a strong prominence of the accent on the verb and downtoning of the accent on the object.

Here the wicked enchantress demands of Rapunzel to let down her hair, and then Rapunzel does as she was told. To express this situation the narrator uses the same lexical items again, this time making use of the active participle to stress the action. The main prominence lies on the verbal part and the object is downtoned. The implication of this pattern is a so-called verum focus (Höhle 1988). It is the factuality of the action that is in FOCUS, but the domain is the whole utterance. At the same time, the utterance is clearly eventive involving a verb and its object which are integrated in a thetic manner, indicating that Rapunzel did the letting-down-the-hair procedure. Any other prominence relation would suggest a different type of information structure. Letting down something could also be TOPICAL here, being the presupposition in another sentence, such as rahit minazzila-l-ha habl ‘And then she let down a rope for her’, where habl ‘rope’ is the only new, contrastive information given and would exhibit the stronger prominence and most probably a pointed hat to indicate contrast. If the whole action was totally unexpected, say, the enchantress had not asked Rapunzel to let down her hair, neither this time, nor ever before, we would expect equal prominences of the verb and the object due to the high information value of both concepts.

6.4. Summary

In this chapter, the issue of theticity has been discussed in the light of new evidence from Egyptian Arabic. The constructions identified by Sasse (1987) for EA, namely the existential construction, the split construction, and to a lesser extent nominalization, could be shown to be of major importance in EA (see Section 6.2).

420
In Section 6.1, I argued against Lambrecht's radical pragmaticization of theticity and the status of the identified constructions as a manifestation of a category of Sentence Focus. I have proposed that thetic utterances only arise in very restricted conditions. Utterances that are inherently thetic are those in which either the existence of an entity or of an event is asserted. The interesting cases, however, are exemplified by the contested type of thetics I called *entity-cum-event* thetics. From a pragmatic point of view, such a statement has to be of the *eventive* type, which means that the ascription of generally valid properties to an entity is excluded. In addition, various semantic and syntactic properties restrict the theticity of an utterance - which may be language-specific to a certain extent.

I specifically argued that eventive split constructions in French and EA are not necessarily thetic, but that there rather are thetic and non-thetic splits, all of which, however, are instances of all-FOCUS sentences. It was shown that some all-new utterances exhibit an internal information structure that identifies a new topic which in a presentational construction, using the presentative particle *fi:* or the locative construction *sand + pronoun* or the deictic demonstrative *da* (*di* and *do:l*), followed by a presentative relative clause, thus forming a 'presentational cleft'. This strategy was compared to the French equivalents, the *ce*-cleft and the *(il) y’a*-cleft that have been thoroughly discussed by Lambrecht (1988a). It was furthermore shown that the syntactic strategy is accompanied by prosodic coding, mostly involving a leading TOPICAL contour on the new referent, which due to its newness to discourse frequently involves a focus accent manifested as a high rise. In the present study, I have argued that the split construction itself is not a construction for the expression of a thetic utterance as proposed by Sasse (1987), but that the use of a relative clause is syntactically conditioned by the subject-diminishing properties of the presentational construction that changes the would-be subject into a predicate noun which in turn may only be referred to in a syntactic relative clause, if a cohesive construction is chosen.

I have further argued that some of the split constructions, however, are of the thetic type, but that their theticity is crucially dependent on the semantics of the constituents involved. A thetic utterance can only be formed if semantic and hence phonological integration (Fuchs 1976, 1980; Jacobs 1993, 1999) is possible. Thus, theticity here is not regarded as a pragmatic category, but rather as a semantically conditioned cognitive phenomenon that can be exploited for pragmatic purposes, i.e. the monolithic presentation of an event involving an entity. Given the semantic basis of theticity, it is related to other phenomena of integration, such as compounding. This much about the signifié-aspect of theticity. As far as the signifiant-aspect is concerned, phonological integration is the hallmark of a thetic sentence. On the formal side, it is exactly this feature that differentiates thetic splits from non-thetic ones. This means that the prosodic shape of a thetic statement is always integrated. It follows from this fact that the SAcc construction with an accented subject and a downtoned predicate is only one type of a prosodically coded thetic statement.
EA was not supposed to exhibit the prosodic SAcc construction for the expression of thetic statements due to its characteristic non-deaccenting prosody. Contrary to expectations, it was however possible to identify a small number of thetic SAcc utterances in the corpus. The conclusions to be drawn from the investigation are that the SAcc construction has only a marginal status in the language. This groups Egyptian Arabic together with other languages that reportedly do not readily allow deaccenting and show a rich pitch accent distribution, but offer the opportunity of word order variation instead. However, the investigation also showed occurrences that suggest prosodic marking of theticity. We may therefore assume that the strategy found in the West-Germanic languages is related to the one that responsible for the thetic SAcc sentences in EA. It was however suggested that proper deaccenting in the West-Germanic languages may have paved the way for the SAcc construction to be grammaticalized, whereas in EA this specific intonation may be traced back to other factors such as the semantic contents, i.e. the information value of the constituents and the attitudinal aspects of the communicative event.

In Section 6.3 the attitudinal aspects of intonation were considered to be the origin of the emerging of the prosodic SAcc construction in other languages, too. It was argued that the main functions that have been associated with the SAcc construction cross-linguistically are related to two attitudes: a decisive matter-of-fact attitude associated with explanatory, descriptive and annuntiative statements and an attitudinal/emotional factor that may result in exclamations, expressing surprise, joy, fright, or alarm. Such exclamations are frequently associated with annuntiative and interruptive utterances. Together with the FOCAL coding of the subject in SAcc utterances, the intonational features associated with the two attitudes, downstep as matter-of-fact and an early high falling accent in exclamations, were deemed responsible for the emergence of the characteristic SAcc intonation. I suggested that the phonetic difference between downstep with reduced prominence, as in EA, to downstep with complete deaccentuation as in English or German is gradual and that both patterns have a common intonational origin and convey the same meanings. The difference in the prosodic systems of these languages, however, may be the reason why the actual use of a SAcc construction remains marginal in EA, while it seems to be assuming or to have already assumed the status of an independent construction in the deaccenting languages. The possibility of deaccenting allows accentuation to be used in an all-or-none manner, which in turn may lead to the employment of accent placement in a categorical manner. Thus, a true paradigmatic contrast arises between subject-accented and predicate-accented utterances. This is supposed to facilitate the grammaticalization of the prosodic SAcc construction. It has, however, been argued that Prosodic Inversion as an instance of a supposed Principle of Paradigmatic Contrast as suggested by Lambrecht (2000) cannot be the motivation for the construction to arise, but only promote a possible grammaticalization. The primacy of intonation is supported by the different grammatical constructions expressing event-reporting statements that appear with the same integrating prosody. The different motivations involved in the ‘Prosodic Inversion’ pattern and the different
functions it may fulfil – from presenting a referent as new, i.e. FOCAL to the expression of attitude, factuality and the like - rather suggest that the pragmatic effect of theticity arises via inference rather than being grammatical category encoded by a specific construction. It shares this property with other specific FOCUS types proposed in the literature, such as a putative exhaustive/identificational-/contrastive category, as recently discussed by Matić & Wedgwood (2013). In an interesting paper on the meanings of focus, where they regard focus in general as a primarily heuristic tool and an interpretive, albeit not grammatical category, these authors point out that in certain instances of a putative ‘corrective FOCUS’ in English in which a certain syntactic construction with a specific pitch accent position are involved, the “obligatorily corrective meaning plainly arises as a result of the interaction of several independent facts about English prosody and syntax” (Matić & Wedgwood 2013: 150). The same has been shown to be true for the thetic utterances in EA above, not only for the prosodic pattern, but also to the use of the modal particle da presented in 6.2.4 whose primary meaning was said to be assertive with certain additional connotations, thereby producing the effect of FOCUS over the whole sentence domain. By way of conclusion I would like to point to an important caveat expressed by Matić & Wedgwood concerning the common practice of subsuming all different kinds of phenomena under the common heading of focus (or, as I should add, information structure) in linguistic analyses (including my own), which could impede descriptive and comparative work. They note that “[a]n incorrectly established category thus turns into a Procrustean bed for the variety of linguistic categories found in the phenomenological field of focus across languages”, especially if the “lack of differentiation between causes and effects” leads to the “reification of the latter” (p.159).
References


Sciences, San Francisco, 1753—1756.


Dilley, Laura C. 2005. The phonetics and phonology of tonal systems: Harvard University; MIT Division of Health Sciences and Technology Ph.D thesis.


Féry, Caroline & Shinichiro Ishihara. 2010. How focus and givenness shape prosody. In Malte


Grønnum, Nina. 1991. Prosodic parameters in a variety of regional Danish standard languages, with a view towards Swedish and German. *Phonetica* 47. 188–214.


Hartmann, Katharina & Malte Zimmermann. 2009. Morphological focus marking in Gu’ru’ntu’m (West Chad). Lingua 119. 1340–1365.


Hellmuth, Sam. 2006b. Pitch accent distribution in Egyptian Arabic. Ph.D. dissertation, SOAS.


163.


439


Skopeteas, Stavros et al. 2006. Questionnaire on Information Structure (QUIS). *Interdisciplinary* 441
Studies on Information Structure (ISIS)(4).


