

# Project description: On the interaction of information structure and prosody. Evidence from Egyptian Arabic<sup>1</sup>

Dina El Zarka

*Karl-Franzens-Universität Graz, Institut für Sprachwissenschaft*

**Abstract.** This report offers a description of my current research project on the interaction of information structure and prosody in Egyptian Arabic. It gives some preliminary results and offers the first conclusions concerning questions of prosodic representation and the encoding of topics, foci andthetic sentences.

**Keywords.** information structure, topicality, focality, theticity, prosody, Egyptian Arabic.

## 1. Primary goals<sup>2</sup>

The present research project is aimed at investigating the information structure of Egyptian Arabic (EA), which is the spoken language of Egypt and is considered the spoken standard for approximately 80 million people. The project will focus on the interaction of prosody and information structure.

Information structure has attracted growing interest among linguists within the past decades. One thing in particular that has greatly contributed to the understanding of the phenomenon is the increasing number of detailed studies on typologically and genetically different languages and the establishment of comprehensive typological projects investigating information structure (e.g. the SFB 632 "Informationsstruktur" ZAS (Berlin) / Univ. of Potsdam funded by the DFG). Information structure in Arabic is still a relatively under-researched field and it is hoped that the detailed investigation of one variety of Arabic will start filling this gap within the area of Arabic language studies and, at the same time, contribute to the general picture of information structure in the languages of the world from a typological perspective. Thus, this study seeks to conjoin two major research traditions, namely Arabic linguistics and general linguistics.

---

<sup>1</sup> The project was funded by the Austrian Fund for Scientific Research FWF (Elise-Richter Project V 131-G19) from October 2009 until January 2011.

<sup>2</sup> I wish to thank all those who assisted me in the various stages of the project, especially my students, collaborators and informants Heidi Abdel Shahid, Doulagy Hanna, Bassem Asker, Matthias Ulbl and Doris Greiner for their invaluable help in various respects.

Being the first in-depth study of EA information structure, this study will cover – at least cursorily – the entire range of information structural phenomena which have either not yet been described or have been described only in a preliminary way. The small number of in-depth studies that already exist focus on specific phenomena as in Mughazy (2008), Mughazy (2009), Holes (2009) and Owens et al. (2009). Thus, the task of describing the range of information structural phenomena is quite a major task and still needs to be dealt with.

At the same time, the study will contribute to the general linguistic debate by looking at some controversial issues like contrastiveness, embedded categories (topic within focus or focus within topic), multiple foci and the issue of theticity, not only by providing new data from a previously unexplored language, but also by looking at this data from a prosodic point of view and under the assumption that the empirically observable constructions arise from the interaction of different grammatical components. Thus, the main objectives of the research project are:

- to provide the first investigation of information structure in EA,
- to establish the language-specific constructions and explore their discourse functions on the basis of naturally occurring speech,
- to study the interaction of prosody and information structure and thereby
- contribute to the discussion of the thetic/categorical distinction as proposed by Kuroda (1972) and Sasse (1987 and work thereafter) and investigate the various constraints on theticity in EA,
- attempt to differentiate between different kinds of topics, pairing the observable syntactico-prosodic forms with various discourse functions,
- address the question of multiple and embedded categories, such as multiple foci, topic-in-focus or focus-in-topic constructions,
- and, last but not least, provide evidence for the pairing of certain prosodic features with specific informational categories, using a model of intonation developed in El Zarka (2011a).

The main methodological contribution to researching information structure is the combination of experimental/laboratory prosodic research with a corpus-based qualitative examination of naturalistic data. This method, I believe, is apt to gain more insight into the intricate interaction between syntax and prosody and to complete the picture of information structural phenomena in the language investigated.

## 2. Methods and data

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

As far as the study of prosody is concerned, a twofold method is being pursued: 1) data from spontaneous discourse and elicited data constitute the database for the corpus-based investigation and 2) production experiments serve as the basis of a quantitative study of specific information structural categories. Prosodic analyses are carried out on the basis of acoustic analysis using PRAAT (Boersma & Weenink 2007) and auditory interpretation alike. For the acoustic analyses, the standards of recent research methodology in the field of intonational phonology are 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 laboratory phonologist doing acoustic phonetics.

Based on such considerations, the entire 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 et. al. 2006),
- Egyptian narratives (read and told),
- other narratives (narrating task based on a picture book, “The frog story” by Mercer Mayer),
- natural conversations
- Egyptian television talk shows from (*Saba:h il-xe:r ja mas'r* ‘Good morning, Egypt’)
- a play written in Egyptian Arabic by Tarek Eltayeb (*Al-Asanser*) (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

As an associate researcher of the D2 project of the SFB 632, I translated the instructions for the QUIS experiments and transcribed and annotated the entire corpus with the help of student assistants. Additionally, I recorded narratives and conversations and carried out production experiments. A considerable amount of the spontaneous speech data, however, was gathered from audio-visual media and downloaded from the internet. These data were annotated and translated by myself. The transcription was done by student assistants and subsequently checked for their correctness by myself. The issue of data collection, transcription and annotation is considered to be an important aspect of the project, because a considerable amount of different genres of EA were documented and will be made available to the scientific community for further research.

The data have been transcribed and annotated roughly along the lines developed by the SFB 632 (Dipper, Götze & Skopeteas (Hgg.) 2007) and archived as xml-files in the EXMARaLDA database, a software tool developed by the Sonderforschungsbereichs 538 “Mehrsprachigkeit” (Schmidt & Wörner 2005). As EXMARaLDA lacks a search tool that offers the possibility of multiple queries, the data are also transferred to the language documentation software ELAN, developed at the *Max Planck Institute of Psycholinguistics* in Nijmegen, Netherlands. ELAN offers complex search options.

The data of the production experiments were manually annotated by myself and subjected to quantitative analysis, using the sound analysis system PRAAT. The results were subsequently assessed statistically, for which task I relied on the help of a professional statistician.

### 3. Information structure and prosody: a theoretical introduction

By way of introduction to the theoretical framework some theory-related conclusions I have obtained so far will be summarized. This will be done in this early section rather than at the end of the report to complete the theoretical picture that will be referred to in section 4.

In the past, information structure has often been dealt with either in studies that were exclusively dedicated to syntax or in prosodic research. Where syntactic studies incorporated prosodic descriptions, the latter were often based on impressionistic observations. On the other hand, prosodists have mostly integrated syntax into the description of prosody by defining prosodic units (phrases) based on certain syntactic constituents. This study is based on the assumption that syntax and prosody (and also other components of grammar) interact and cooperate to encode information structure and that grammatical structures, whether syntactic or prosodic, can be directly motivated by pragmatic ends.

The theoretical framework for the present study is functional-typological and deals primarily with the correlates of information structure in sentence grammar (Lambrecht 1994; Sasse 1987, 2006). Unfortunately, for reasons of time and space, it will not be possible to properly go into larger discourse units and pragmatics using exclusively authentic spoken discourse; such issues will need to be taken up in subsequent work.

In light of what has been said above, I consider prosody not only as a means of interpreting syntactic structure, but also as a means of encoding information structure on a par with syntactic or morpholexical means. It is thus essential to this approach that prosodic entities are not defined in syntactic terms, although they frequently coincide with syntactic constituents. Prosody is considered to follow its own rules that have rhythmic and melodic motivation. I assumed that the rhythmic/melodic structures can be utilized or altered to encode information (El Zarka 2005). At the same time, the different components of grammar are not viewed as autonomous subsystems, but rather seen to interact with each other to convey information. Following Lambrecht (1994: 12), I assume

a linguistic theory of high explanatory value to be one in which these forces [the different grammatical components, DE] are not only analyzed in isolation but also in their multiple dependence relations to each other. In such a theory the grammatical structures found in particular languages would then be seen as language-specific manifestations of the interplay between the different components of grammar.

### 3.1. Categories of information structure

To give some background of the theoretical basis of the study, I will first introduce the main categories to be examined and review some basic definitions to start with.

#### 3.1.1. Topicality

Although the definitions of topic are commonly viewed as rather vague, it cannot be doubted that topic is a cross-linguistically valid, probably universal category of information structure. The notion 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. The notion of (sentence) topic dealt with in the present study is illustrated by the following example (1) from Krifka (2007: 41):

- (1) a. [Aristotle Onassis]<sub>Topic</sub> [married Jacqueline Kennedy]<sub>Comment</sub>  
 b. [Jacqueline Kennedy]<sub>Topic</sub> [married Aristotle Onassis]<sub>Comment</sub>

Definitions of topic vary in the literature. However, the most widely accepted definition seems to include pointing out some *aboutness*-relation that holds between the referent of a topic expression of a specific sentence or clause and the rest of the proposition expressed by the same sentence or clause (Hockett 1958; Gundel 1976; Dik 1978; Reinhart 1982; Lambrecht 1994). In (1a) it is Onassis who is the topic of the sentence and “marrying Jacqueline Kennedy” is being predicated about that topic; in (1b) it is the other way round.

Topic has also been defined as a *scene-setting* expression that sets the frame for a given proposition (Chafe 1976) or as a *filecard* on which new information is stored (Vallduví 1992; Krifka 2007). Topics that fulfill a the sense of Chafe (1976), among which count the so-called *Chinese-style topics* of *topic prominent languages* (cf. Li & Thompson 1976), are considered as a subcategory of a comprehensive category *topic* and referred to as *frames* in the present study (cf. example 2).

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

Accordingly, different constructions have been said to express topicality, such as topicalized (fronted) and clause external (extraposed, detached) constituents that refer to participants of a proposition or delimit the spatial or temporal frames in which a certain predication holds. Lambrecht (1994) considers the prototypical topic expression to be an unaccented (or zero) pronoun.

The present study starts out with an investigation of the different formal types of expressing sentence topics cross-linguistically that have been identified in the

literature. Following Lambrecht (1994) and Maslova & Bernini (2006), a very broad understanding of topic is adopted in this research which includes both *subjects* and *topics* in the sense of Chafe's definition (1976, 1994). It is assumed that these constructions may be used to fulfill a variety of functions, including the base for a predication (aboutness), spatial and temporal frame setting and expressing a reference point (Maslova & Bernini 2006).

### 3.1.2. Focality

The concepts of focus greatly differ among the various approaches with different focus definitions making different predictions about the parts of a proposition considered to be in or 'the' focus. For example, some definitions of focus explicitly or implicitly rely on the salience of constituents as the expression of what is seen to be the most important information. Among these is the definition offered by Dik (1978: 19) who defines focus as the "relatively most important or salient information" and definitions explicitly relying on prosodic highlighting such as those by Halliday (1967, 1968) or Chafe (1976). But all definitions that rely on intuitively appealing, but vague notions like salience, are considered unsatisfactory by Krifka (2007: 28) who suggests defining focus as indicating "the presence of alternatives that are relevant for the interpretation" (p. 18).

Probably the least ambiguous definition is that offered by Lambrecht who takes focus to be the "semantic component of a pragmatically structured proposition whereby the assertion differs from the presupposition" (Lambrecht 1994: 213). Such a definition entails that a topic cannot be in focus, thus Lambrecht distinguishes between *contrastive topic* and *contrastive focus*, while it has been suggested in the literature that a contrastive topic contains a focus (Molnár 1998, Krifka 2007). I will return to that question below. A classical way to determine focus is with the help a question-answer (Q/A) pair, where the new information given in the answer is assumed to be in focus; furthermore it is assumed that focus is somehow indicated by highlighting. The different Q/A pairs in (3) also illustrate how the scope of focus (see below) may vary in accordance with the question (accented constituents (words) are marked in capital letters).

- (3) a. What happened?  
[John stole the COOKIES]<sub>FOC</sub>
- b. What did John do?  
John [stole the COOKIES]<sub>FOC</sub>
- c. What did John steal?  
John stole [the COOKIES]<sub>FOC</sub>.
- d. Who stole the cookies?  
[JOHN]<sub>FOC</sub> stole the cookies..

As the examples in (3) show, it is not the content of a sentence that is altered by the different focus structures, but rather the ‘packaging’ of the content as suggested by Chafe (1976). Krifka (2007: 21) calls this the “pragmatic use of focus” as opposed to semantic uses that affect the truth-conditional content, as in the case of focus-sensitive particles such as *only*, *also* or *even*. In this study, I am concerned with information structure as a pragmatic or packaging phenomenon.

In the literature, it has been suggested that coherent discourse is structured by implicit questions (e.g. Büring 2003). The assumption of such question-answer pairs to identify the focus part of a structured proposition is not unproblematic, though. Examining spontaneous speech soon shows the limitations of such an assumption.

Another notion related to focus is contrastiveness. A much debated question in the literature is whether a contrastive topic is in focus or contains a focus or not (Molnár 1998) and whether *contrast* is a separate category of information structure (Chafe 1976, Molnár 2006). As mentioned above, according to Lambrecht’s definition the definition of topic and focus are mutually exclusive, whereas the highlighting-based and alternative-based definitions have the advantage of allowing a unified account of contrastiveness as contrastive topics can be interpreted as also involving focus (4).

(4) A: What do your siblings do?

B: [My [SISter]<sub>foc</sub>]<sub>TOP</sub> [studies MEDicine]<sub>FOC</sub>  
and [my BROther]<sub>foc</sub>]<sub>TOP</sub> is [working on a FREIGHT ship]<sub>FOC</sub>  
(Krifka 2007: 4)

Based on my EA data, but assuming a more widespread application of this proposal, I assume two different types of focus here, *rhematic* focus, which will be referred to as *FOCUS*,<sup>3</sup> spelled with capital letters, the definition of which is along the lines of the one given by Lambrecht, and another type of focus that is compatible with a definition in terms of alternatives (e.g. Krifka 2007), I call this second type *focus of interest* or simply *focus* (spelled with lower-case letters), following Bolinger’s terminology (Bolinger 1958 and subsequent work). The first type is basically equivalent to the old category of *rheme* assumed by the Prague structuralists. The second type is intimately related to what is informally referred to as highlighting and has a clear prosodic correlate, namely prominence or accent, in many languages (i.e. English, German or Egyptian Arabic).<sup>4</sup> In the linguistic literature of the eighties and nineties, especially in the German tradition (cf. the volumes edited by Jacobs (1992) or Reis (1993)) a distinction between rheme and focus is made. In another linguistic tradition, however, there have been attempts to reduce the number of dichotomies

<sup>3</sup> In the following I will only use the spelling FOCUS when especially referring to this category my own work. I will not, however, attempt to differentiate between notions of focus in the discussion of the literature.

<sup>4</sup> It has to be noted that prosodic prominence cannot to be expected to be a correlate of focus in all languages. It has, for instance, been suggested for some African languages that focus is not marked by prominence (Downing 2003, Zerbian 2006).

(topic-comment, theme-rheme, focus-background, given-new) for the sake of establishing a simpler inventory (Vallduví 1992, Lambrecht (1994), and Büring (1997) among others).

In the present study, I also distinguish between the two types of focus mentioned above which are illustrated in (5). Relating to the Q/A congruence, each of the (a) answers exhibits one large FOCUS domain covering the whole sentence while the (b) answers are bipartite constructions with the FOCUS domain covering the predicate phrase only.

- (5)
- a. A: What happened?  
 B: [John stole the COOKIES]<sub>FOC</sub>
- B': [[JOHN]<sub>foc</sub> stole the [COOKIES]<sub>foc</sub>]<sub>FOC</sub>  
 B'': [[JOHN]<sub>foc</sub> [STOLE]<sub>foc</sub> the [COOKIES]<sub>foc</sub>]<sub>FOC</sub>
- b. A: What did John do?  
 B: [John]<sub>TOP</sub> [stole the [COOKIES]<sub>foc</sub>]<sub>FOC</sub>  
 B': [[JOHN]<sub>foc</sub>]<sub>TOP</sub> [stole the [COOKIES]<sub>foc</sub>]<sub>FOC</sub>

But if we wish to account for the different prosodic realizations of the given answers, we have to account for the presence of the individual accents within the larger domains. 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 1983b; Selkirk 1984; Bolinger 1986). Some theories allow for the embedding of these narrow foci into larger focus domains, such as in Selkirk (1984). Other proposals identify only one focus accent at the end of a domain and consider the additional accents to be secondary; Büring (2003), for instance, calls accents outside the focus domain *ornamental* accents. In the present study, I take on the view, following Bolinger (1986, et passim), that every accent is meaningful itself and that the different answers shown in (5) have a different pragmatic impact. Because limitations of space do not allow me to elaborate more on these distinctions, I only shortly explain the difference between focus of interest contained in a topic and in a rhematic domain. If we compare the (B') examples of (5a) and (5b) we find that the structures are alike in terms of syntax and in terms of accent distribution. They both contain a focus associated with the subject *John* that puts *John* in the foreground. In (5b) however, *John* is clearly the topic of the sentence and, judging on the foregoing answer, and the presupposition that *John* did something is evoked, whereas what he did is the new information and thus the FOCUS of the assertion. As frequently observed in the literature (cf. references in section 3.2.1.2), the tonal shape of the accent in the topic constituent *John* differs from that of the accent in the FOCUS constituent. In (6) these different shapes are illustrated with a schematic contour; the accent on *John* in (6a) is a *closing* accent (see 3.2.1.2) indicating a rhematic FOCUS and in (6b) it is a *leading* accent indicating topicality.

- (6)
- a. [JOHN stole the COOKIES]<sub>FOC</sub>
- b. [JOHN]<sub>Top</sub> [stole the COOKIES]<sub>FOC</sub>

Although John is clearly not the answer to the question, the mere accenting evokes alternatives, as the close paraphrase ‘As for John, he stole the cookies’ suggests. The same is true for the contrastive topic in Krifka’s example (4) above where the alternatives are explicitly cited. The sentence in (6b) with the indicated intonation might well be followed by a statement such as ‘but Mary did something worse.’

In EA, these categories have clear formal correlates. Focus of interest is signaled by prominence, i.e. an accent that is not specified for its tonal shape, whereas rhematic FOCUS is signaled by the tonal contour in the first place (but see section 4.4.2 for more details).

Another important issue is the *scope* of focus. Prosodists usually distinguish between *broad* and *narrow* focus while syntacticians and semanticists more accurately define the scope of focus in terms of semantic constituents like arguments, predicates and the like. As focus of interest is by definition limited to one constituent, its scope is always narrow. With other words, the issue of scope does not arise. Scope is only an interesting issue in what I call rhematic FOCUS.

### 3.1.3. Theticity

Theticity as a linguistic phenomenon has been extensively studied by Sasse (1987, 2006). The basic idea is that there is a distinction between *thetic* and *categorical* statements as the linguistic expressions of pragmatic choices a given speaker may make. Categorical statements are characterized by the division of a proposition into two major parts, a *predication base* (Sasse 1987) (i.e. a topic), and a *predication* that ascribes some property or an action to the entity previously established by the topic expression. However, 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 VS order instead of SV 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.”

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 & Sasse (Hgg.) 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.

### 3.2. Prosody

Pragmatic functions have always been a major concern of prosodic research, especially intonation research. Developed on the basis of data from European languages (mostly English), theoretical models (British tradition, Autosegmental-metrical (AM) models) have mostly operated with the concept of the *nucleus* as the most prominent tonal event, typically the last tonal event, in an intonation unit.<sup>5</sup> It is commonly held that the nucleus indicates focus. Another common assumption has been that the given material is *deaccented*. Cruttenden (2006) suggests in his typological sketch that there are languages that favour *deaccenting* and others that favour what he calls *reaccenting* which – according to Cruttenden – is especially common in Romance languages and fairly common in (Tunisian) Arabic. A reluctance to use deaccenting has also been reported for some European languages, such as Italian (cf. Ladd 1996).

As research on pitch accent and tone languages (Gårding 1983; Xu 1999) suggests, focus and background are rather marked by the expansion and compression of pitch range as intonational correlates of focus and background, respectively.

The position defended here is that one correlate of focus (of interest) is prominence which is relational in nature (cf. section 3.2.1.1) and which is only partly brought about by expansion vs. compression of pitch range, other correlates of prominence being intensity and duration. Deaccenting, according to that view, is only a special instance of pitch range compression.

It is further assumed here that another equally iconic correlate of focus, in our terms of rhematic FOCUS, is the type of the tonal contour, namely the falling type. The communicative value of the fundamental opposition between rising and falling gestures has been recognized by different scholars of intonation in the British tradition and by Dwight Bolinger half a century ago (cf. Brazil 1997; Bolinger 1986; Gussenhoven 1983a; Cruttenden 1997, 2006) and will be elaborated below (3.2.1.2).

The investigation of prosody and intonation is carried out using the descriptive labels of autosegmental-metrical (AM) approaches (Pierrehumbert 1980; Gussenhoven 1983a, b, 1984, and work thereafter; Ladd 1996), although my approach cannot be viewed as strictly abiding by the basic tenets of generative models of intonation. The approach taken here is based on identifying whole configurations that have a communicative function and are working in a top-down manner rather than being constructed in a bottom-up fashion via the concatenation of intonational ‘phonemes’ and by looking into the phonetic make-up of such configurations rather

---

<sup>5</sup> In some definitions, most notably that of Pierrehumbert (1980), the nucleus is not necessarily the most prominent, but simply the final accent of an intonation unit.

than defining intonational units on a purely phonetic basis as in a Standard AM-approach. A more detailed description of the prosodic model underlying this study is given in El Zarka (2011a).

### 3.2.1. Egyptian Arabic Prosody

As this study heavily draws on prosodic facts to explain the constructions expressing pragmatic relations, a brief overview of EA prosody and of my own prior work on which the main hypotheses of this investigation are based will be given.

It is a commonly accepted fact that intonation units are characterized by a succession of high-low sequences in EA. This tendency to accent all words was recognized, for example, by Mitchell (1993) and has been verified in instrumental studies such as Rifaat's study of Classical Arabic declaratives (Rifaat 1991), my own dissertation on Egyptian Modern Standard Arabic (MSA) (Rastegar-El Zarka 1997) and finally in Hellmuth's dissertation (2006a) devoted solely to the subject of pitch accent distribution. The studies show that almost every content word in EA carries a *pitch-accent*, as the up-and-down tonal contour has been termed within an AM-framework (cf. Ladd 1996). This is very different from the situation in West-Germanic languages such as English or German, where given information is usually deaccented (cf. above). El Zarka (2005) has suggested that the situation in EA parallels that of Romance languages and pitch accent languages in that respect.

#### 3.2.1.1. The relational nature of intonational peaks

Based on what has been said so far, the basic assumption concerning the investigation of prosody is that it has to go beyond identifying *accented* vs. *deaccented* constituents. As indicated above, accentuation is seen as a means of creating prominence and should not be viewed as a categorical feature operating in an all-or-none manner. It is rather the relation between individual accents that matters (cf. also Dille 2005 for introducing a formal account for tonal relativity to Standard AM-theory). This can be observed in the EA data as well, and likewise in the MSA-data described in Rastegar-El Zarka (1997).

#### 3.2.1.2. The iconic nature of tonal contours

In El Zarka (2011a) I have also argued that, in EA, topical and focal (i.e. rhematic) constituents are encoded with rising and falling contours, respectively. I suggested that, in EA, topics can either be characterized by a flat or much reduced tonal contour, the so-called *linking tone* or by a rising contour which may optionally be accompanied by a phrase break, a contour I call *leading tone*. On the other hand, rhematic FOCUS is characterized by a (rise-)fall or a *closing tone*.

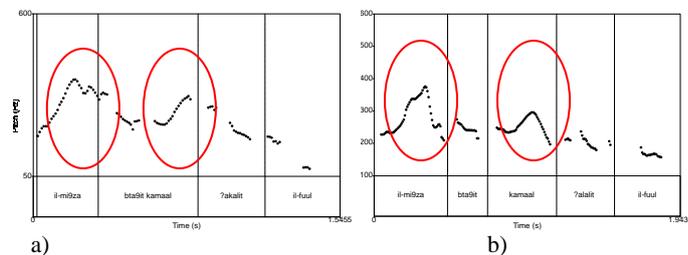
A three-way distinction of tones had also been proposed by Brazil (1975, 1997) who distinguished between *referring* (rising and falling-rising) and *proclaiming* (falling and rising-falling) tones and a neutral tone that is formally characterized by level intonation in British English. Without necessarily adopting Brazil's functional

explanation, I adopt the notion of a level or compressed tonal contour that I take to be used for background constituents, especially if these refer to entities that are sufficiently activated and equally sufficiently established as topics. Contrary to such reduced contours, a fully rising contour may be used to activate a referent or to set a frame within which a predication holds (in the sense of Chafe's definition of topic, Chafe 1976). Brazil's referring tone corresponds to what Gussenhoven (1983a) called *selection*. This notion of *selection* seems especially useful in connection with investigating information structure as it is conceived as selecting "a variable from the background" (Gussenhoven 1984: 201). This is an important notion if we are to understand the actual use of certain contours in discourse.

A tonal characterization of topics and focal elements has also been suggested for German, e.g. by Uhmann (1991), Féry (1993) and Grice and her colleagues (cf. Grice et al. 2005) on the part of prosody research, but the issue has also been discussed by scholars of information structure under different headings such as *I-Topikalisierung* (intonational topicalization) (Jacobs 1997) or *bridge accent* (Büring 1997) for a basically rising-falling contour that expresses a certain type of topic-comment structure.

The claim that EA topics are frequently encoded as rises and rhematic FOCUS constituents as falls constitutes a central theme throughout the study, and cases where the intonation tune is the major factor or even the only one distinguishing between different informational meanings are identified. Figure 1 illustrates the intonation contour of a topic-comment sentence (panel a) and a narrowly focused subject (panel b) in which the sole difference is the intonation contour. The contour also shows that both, the topical argument in panel (a) and the rhematic argument in panel (b) contain foci of interest associated with the individual nouns.

Fig. 1: *il-miʕza bita:ʕit kama:l ʔakalit il-fu:l* 'Kamal's goat ate the beans' with K's goat as a topic with two rising accents in panel (a) and as a narrowly focused subject with two falling accents in panel (b) (from the QUIS corpus)



## 4. Preliminary results of the corpus study

The study has two main objectives (see Section 1): (1) a theoretical one which is to contribute to the discussion on the interface of prosody and information structure and (2) a descriptive one, namely to give an overview of TOPIC and FOCUS constructions in EA. In order to conduct an in-depth examination of the phenomena

it is, however, not possible to treat all aspects of information structure in EA equally thoroughly. Instead, the detailed study will be limited to two aspects of information structure: topics and theticity.

The picture that emerges from prior research (Moutaouakil 1989; Ouhalla 1993; Brustad 2000) 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. Rhematic FOCUS is also encoded syntactically, either by fronting a focused 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.

Sections 4.1 to 4.3 offer a brief overview of syntactic topic and focus constructions occurring in EA, with a slightly more comprehensive listing of the thetic constructions found so far. This survey makes no claim to completeness, nor do I give any information about the frequency, context dependence or pragmatic functions of the identified patterns, a task that will be done in the next phase of the project. Thus, the first three subsections have to be regarded as an outlook rather than a report. Section 4.4 will again deal with the main theme of the study, which is the interaction of information structure and prosody that has already been a subject of section 3. This last section is based on the prosodic study that has already been done and will summarize the major results obtained so far.

## 4.1. Topics in EA

Neo-Arabic varieties in general and EA in particular show features of a topic-prominent-language according to the classification by Li & Thompson (1976), as suggested by Brustad (2000). This finding comes as no surprise, as spoken varieties in general are characterized by the extensive use of clause-external topic constructions (Chafe's (1976) *Chinese-style topics* or Maslova & Bernini's (2006) *hanging topics*).

### 4.1.1. Left dislocated topics

As in other Arabic varieties left-dislocated topics abound in EA. Such topics are referred to by resumptive pronouns within the main clause (7)

- (7) *ir-ra:gil*    *itraddit*                      *fi:-h*            *ir-ro:h*  
 the-man    was.brought.back.she    in-him            the-soul  
 'The man was brought back to life.'                      (Woidich 2006:178)

In the above example the topic expression *ir-ra:gil* 'the man' is referred to by a pronominal suffix attached to the preposition *fi:* 'in' in the main clause. Syntactically, the topic thus fills the position of a prepositional object in the main clause. In example (7) the predication is expressed by a passive verb but the contexts in which passive constructions are used in EA are limited to agentless propositions. A sentence such as 'The man was hit by a car' has no structural equivalent in EA. The

corresponding construction is one with a preposed topic and an independent clause taking up the topic as an object of the verb ‘hit’:

- (8) *ir-ra:gil dʿarabit-u ʕarabijja*  
 the-man hit.she-him car.S.F  
 \**ir-ra:gil itdʿarab min ʕarabijja*  
 \*the-man was.hit.he by car

The limited use of passive constructions has been claimed to be one of the diagnostics of topic-prominent languages by Li & Thompson (1976). The above use of left-detachment points to the fact that in Arabic, unlike English or German, this type of topic-comment structure is not only a device for pragmatic information structuring, but that it is grammaticalized to the extent that it has become the only grammatical way to express the patient as a topic if the agentive participant is also mentioned. Thus Eisele (1999: 112) claims that such constructions may have a “purely syntactic function” in EA.

Brustad (2000) further argues for the primacy of the VS word order and that all SV constructions, not only those with dislocated topics, should be viewed as topic-prominent. Although the question of ‘basic’ or dominant word order is not at the center of my study, careful investigation of the corpus will certainly contribute to the understanding of the distribution of VS and SV constructions and their discourse functions. Even in Classical/Standard Arabic, which is commonly described as a verb-initial language, the choice of VS or SV order is functionally motivated, with verb-initial sentences being used for narrative discourse types and noun-initial sentences for descriptive passages and expository genres. Holes (1995) referred to that property of word order as event-oriented (VS) and entity-oriented (SV), an analysis that gains even more support from the fact that VS order tends to co-occur with perfective verbs and SV order with imperfective verbs (Benmamoun 2000).

#### 4.1.2. Deictic and anaphoric expressions

In EA sentences containing a verb, fully active referents are generally not referred to anaphorically by a pronoun if they are in subject position. Personal subject pronouns are commonly said to be used to mark contrast or disambiguate referents or indicate a subject switch (Cowell 1964; Eid 1983, 1991; Jelinek 2002) while a zero anaphor is the appropriate linguistic expression for a continuous topic.

Generally, EA pronouns, whether demonstrative or personal, exhibit a variety of different forms and cover a wide range of functions, among which pragmatic functions are the most noticeable. The use of the pronouns will also be thoroughly investigated in the present study.

While the personal pronouns *huwwa*, *hijja*, *humma* (‘he’, ‘she’, ‘they’) – which will subsequently be referred to as the *h-set* – have received great attention among linguists, the demonstrative pronouns in pronominal use *da*, *di*, *do:l* (‘this’ m., ‘this’ f., ‘these’), which will be referred to as the *d-set*, have not been subject to investigation. Both sets may be used anaphorically or deictically. It seems that in EA

deictic and anaphoric expressions bear some similarity to those of many other languages, including German, Dutch or Modern Hebrew (Matras & Bolkestein 2006). As in Hebrew, the *d-set* seems to be used as a device for immediate reference or extratextual/situational deixis whereas the *h-set* is, among other things, specialized for contextual use. In addition, the *d-set* is purely colloquial while the *h-pronouns* have to be used in formal speech or writing (cf. Matras & Bolkestein 2006).

The corpus also contains a number of cases that suggest that the form *da* of the *d-set* has been grammaticalized and functions as a discourse particle that I have argued to be associated with FOCUS (9) (for a detailed argumentation cf. El Zarka (2012)).

- (9) *da ka:n-it kida bi-ti-ʕmil so:ʔ*  
 PRT was.she so she.is.making voice  
 ‘It made a kind of voice’ (corpus, conversation)

#### 4.1.3. Hanging topics, multiple topics and reference point

A wealth of other topic constructions have partly been described by Brustad (2000), relying on Chafe’s definition of topic. The example may show that EA allows multiple topics (including individual, spatial and temporal frames) that are left-attached to the main clause (slashes indicate where phrase breaks are likely to occur; glosses and segmentation DE).

- (10) *fi jo:m / su:su / s-sana lli fa:tit*  
 in day / Susu / the-year REL has.passed /
- ʔablə- ma tsa:fir/ yimkin ma-jgi:f fahr /*  
 before-PRT she.departs / maybe NEG-it.comes-NEG month
- ka:n xamastʕa:far ʔaw ʕafar tijja:m /*  
 was.it fifteen or ten days
- ʔa:mit taʕba:na sʕ- sʕubħ ʔawi*  
 got.up.she feeling.poorly the-morning very

‘One day Susu, last year before she left maybe less than a month, it was fifteen or ten days, she woke up in the morning feeling very poorly.’  
 (Brustad 2000: 388f.)

Looking into EA conversational data permits the observation of an interesting phenomenon abundant in natural conversation. It is typical of EA (and other varieties of spoken Arabic) that speakers tend to begin their conversational turns with first person pronouns *ʔana/ʔihna* (11a) preposed to the first utterance. But also the use of the second person pronouns *ʔinta* (m.Sg.)/*ʔinti* (f.Sg.)/*ʔintu* or the honorific expression *ħad ʕritak* (m)/*ħad ʕritik* (f) (11b) as a respectful form of address

are very common in questions. The indication of the relation between the extraposed topic and the participant within the main clause may be explicitly indicated or not:

- (11) a. *?ana l-fa:r bijilʕab fi ʕibb-i*  
 I the-mouse he.is.playing in garment.my  
 'I am worried.' (own corpus: play)
- b. *?alit-lu ?inta hadʕritak tiʕraf-ni*  
 said.she.to.him you you.POL you.know-me  
 'She said, "Do you know me, sir?"' (Brustad 2000: 393)

Such constructions are typical of Li & Thompson's (1976) topic-prominent languages. As Maslova & Bernini (2006: 88) point out, this type of construction, which they call *hanging topic*, mostly entails a 'stable' relation between the topic referent and participant of the proposition encoded by the comment that can be encoded as grammatical possession. Following Langacker (1993), they view this type of topic construction on a par with grammatical possession as a "manifestation of the same *reference point*" (p. 90; emphasis DE).

Although such reference point topics exhibit the same syntactic form as aboutness-topics or temporal and spatial frames, their prosody obviously differs; they are not normally set apart from the rest of a clause by forming a separate prosodic phrase nor are they uttered with a leading tonal contour.

In EA, multiple topics are quite common and the relationship between the different topics can be very diverse as example (10) above shows. Example (12) contains three topic expressions whose referents seem to function as reference points anchoring each of the topics to the preceding one, thus building a reference chain from the innermost sentence topic to the outermost referent which is the speaker himself, thereby anchoring the aboutness-topic of the main clause in the immediate discourse situation.

- (12) *?iḥna hadʕarit-na ʕumr-aha sabaʕtala:f sana*  
 we civilization-our age-her 7000 year  
 'Our civilization, its age is 7000 years.' Woidich (2006: 178)

The interesting questions that arise are (i) what exactly constitutes the difference between the various topic constructions that are attested in the language and (ii) in what contexts are they used.

## 4.2. FOCUS in EA

As already indicated EA exhibits relatively free word order. Brustad (2000: 342) states that "almost any basic constituent may begin an Arabic sentence". It has been shown that fronting of rhematic FOCUS phrases is a common way to mark contrastive focus in Classical Arabic (Moutaouakil 1989, Ouhalla 1993). Thus, for example, direct object fronting is a viable way to mark an object as being contrastively focused in Classical

Arabic, as the case system allows for the disambiguation of agent and patient in NVN or VNN structures.

Modern Arabic vernaculars might not be as flexible due to the lack of case endings, but Brustad (2000: 343) also claims focus fronting to be true for modern vernaculars. In verb-initial typology, she reports the contrastive object to occupy the position after the verb, thus yielding a VOS order instead of the unmarked VSO order. The initial impression, which has yet to be systematically tested on the corpus data, is that EA is more restrictive in that respect than other Arabic vernaculars. Brustad claims that if a syntactic object is fronted without being referred to by a pronominal suffix it is to be interpreted as a focus rather than a topic. The one example she adduces containing an object followed by the subject that unequivocally denotes contrastive focus is from Moroccan Arabic: *hadik byit ?ana* ‘this one I want’, a structure that would not be felicitous in EA, where instead a cleft might be used, something like (*hijja di lli ana ?ayizha* ‘this is the one I want.’

Clefts are extensively used in EA and are the unmarked way to express contrastive subject focus, but object focus is more likely to be realized *in situ* as example (13) from the QUIS corpus shows:

- (13) Q: Who hit Karim, Nelly or Nermin?  
 A: *nelli hijja lli d'arabit kari:m*  
 Nelly she REL hit.she Karim  
 ‘Nelly is the one who hit Karim.’

Q: Do you prefer his shirt or his trousers?

- A: *?ahibb ?ami:s'u ?aktar*  
 I.like shirt.his more  
 ‘I prefer his shirt.’

(own corpus: QUIS data)

It seems that object focus is mostly realized *in situ* which frequently coincides with the main prominence at the end of an intonation phrase.

### 4.3. Thetics in EA

Cross-linguistically, certain structures have been identified as typicalthetic expressions. Among these are VS word order (as opposed to SV order for topic-comment structures), split constructions (i.e. subject + relative clause), verb nominalization and subject accentuation. Sasse cites some examples from EA in support of his claims but, apart from that, theticity in EA has not been dealt with in the literature. Three of the four above mentioned constructions have been said to occur in EA; the only construction that has not been identified in EA is subject accentuation. Here are some typical examples ofthetic utterances:

- (14) a. *wi-nizil guha l-qahira*  
 and-went.down Goha Cairo  
 ‘and Goha went to Cairo’ VS order initiating an episode in a narrative  
 (own corpus, narrative)



Holes (1995) thus seems to be the best point of departure in an attempt to capture the facts about word order in EA.

In the same way thatthetic utterances are multifunctional, their occurrence cannot be definitely predicted from the context. The frequent occurrence of SV structures, with a topical subject, especially with introductive function, where one would expect athetic statement has been reported to be quite frequent in a cross-linguistic study of word order by Sasse (cf. Sasse 2006: 284). Similarly, the elicited data from the questionnaire on information structure (QUIS, SFB 632/ D2) frequently show topic-comment structures wherethetic utterances might be expected. Besides frequent SV clauses, there are also many instances of the existential construction mentioned above. The striking fact about these is that they exhibit the prosody of topic-comment structures (cf. below).

The constructions referred to above will be investigated in some detail, the main focus of the study, however, is on the prosodic encoding ofthetic in EA. Contrary to what has been assumed so far in the literature (Hellmuth 2010) I argue that there is a certain prosodic construction that is typical for *sentence focus* and thatthetic sentences display a prosodic contour that significantly differs from a syntactically identical topic-comment structure (cf. section 4.4.4).

#### **4.4. The role of prosody in expressing information structure in EA**

This last section is devoted to the crude results of the prosodic study that have been partly made public in talks and papers (El Zarka 2011a, 2011b) so far. In the next phase of the project, these results have to be refined in a more detailed investigation of the data which is also expected to furnish more statistical evidence for the basic claims. The presentation will include a brief overview of some proposals made in the literature and relate them to the results of the present study.

Based on observations in Romance languages, Vallduví (1992) suggested a principled distinction between *plastic accent languages* like English and *non-plastic accent languages* like Catalán. Vallduví argues that languages that do not implement prosody (i.e. accentuation) for information-structural purposes counter-balance this by exploiting syntactical means to a greater extent. Although it certainly is true that EA, like Romance languages, makes use of syntactic constructions to a very high degree - narrow subject focus, for example, is mostly expressed by cleft sentences and left-dislocated or hanging topics (Maslova & Bernini 2006) are commonplace - Vallduví's claim seems to be too strong. Rather prosody is always an essential part of spoken language and to encode information structure in a very iconic way. Thus, it will be shown how prosody accompanies the different syntactic constructions, and in certain cases may also be used to disambiguate syntactic structures.

##### **4.4.1. Topics and the mental representation of referents**

Topics are frequently used to refer to participants of a proposition. Referents may be more or less within the focus of consciousness (of the hearer) at the time they are

mentioned. They may be anything from fully activated to unused or new, but to express a brand-new referent, i.e. a referent that is not identifiable to the hearer, as the topic of a sentence is highly unlikely. As to formal devices, it is a well-established fact across languages that the overwhelming majority of topic referents are fully active referents that are mostly referred to by unaccented personal pronouns or even zero pronouns. However, lexical nouns or noun phrases are used to refer to referents that are not fully activated in the consciousness of the hearer (Prince 1981, Chafe 1994, Lambrecht 1994), a fact that Chafe (1994) called the *light subject constraint*. Yet topics can of course also be expressed by full NPs whether they serve as subjects in an SV-structure or whether they are left-dislocated topics with a resumptive pronoun referring back to them within the main clause.

It has also been observed that accentuation serves to activate inactive topics (Givón 1983; Chafe 1974, 1994) or to mark a topic as contrastive (Lambrecht 1994; Büring 1997; Krifka 2007). Lambrecht (1994: 323-333) convincingly argues, however, that the mental representation of referents and the pragmatic role of a referent have to be kept separate. According to Lambrecht, an accent on a topic expression may be used to activate a referent, but it can also be used simply to establish a topic and to strengthen the aboutness-relation between that topic and the predication to follow.

We have seen that in EA content words are almost always accented, nevertheless *topic selection* is marked prosodically in EA by the use of what was called a *leading* contour by El Zarka (2011a).

As mentioned above, the lexico-syntactic form of topic expressions may be viewed as covering an explicitness scale ranging from non-explicit or implicit topics (zero pronouns) to fully explicit constructions (extraposed full NPs) at the other end of the scale. In a similar way, the prosodic form may cover an equivalent explicitness scale from a totally flat tonal contour with accompanying deaccentuation of a lexical item over a reduced accentual gesture to a dominantly rising gesture and finally, as the most explicit expression of topicality, to a fully rising gesture with accompanying phrase break after the topical element.

These facts point to the gradience of prosodic phenomena and raise the question of if and how they can be interpreted categorically? Based on the prosodic form of the utterances, a distinction is made here between *leading topic* (rising tone or tune) and *linking topic* (flat or almost flat tone or tune) and the aims of the more detailed corpus investigation will be to find out if such a categorical distinction is justifiable at all and how these constructions differ functionally, thereby relating the findings to the different functional explanations and definitions that have been offered for the notion of topic in the literature.

A further task will be to identify the contexts for the different constructions and the way prosodic, lexical and syntactic strategies interact.

#### 4.4.2. FOCUS and contrastiveness

In the light of what has been said in section (4.2), the question of whether focus (i.e. FOCUS) can be indicated solely by prosody arises. Prior accounts of information structure in Arabic either omit this question altogether, as in Brustad (2000), or they

only incidentally mention accentuation as a means of focus marking (Moutaouakil 1989; Ouhalla 1993). Among prosodic accounts of focus, there seem to be two different positions concerning the availability of prosodic focus marking in EA. It was suggested by earlier studies such as those by Gary & Gamal Eldin (1982) and Mitchell (1993) that EA intonation exhibits properties known from English, namely that the most important information is prosodically highlighted. Norlin (1989), in a small-scale experimental study with one speaker, suggests that focus in EA is expressed by expanding the pitch range of the focused element and compressing all subsequent materials without deaccenting it. Based on Norlin's claim, Hellmuth (2006a, 2009), drawing on experimental data, suggests that *information focus* shows no prosodic reflex in EA and that pitch range manipulation may be used to express contrastive focus.

My own experimental work and corpus data suggest that focus/FOCUS in EA can indeed be indicated by prosody alone. Keeping in mind our distinction between focus of interest and rhematic FOCUS, it can be shown that focus of interest is always prominent, i.e. associated with an accent. While the tonal shape of a focus of interest is underspecified, rhematic FOCUS has as its most important correlate a basically falling tonal contour as suggested in 3.2.1 above. In addition, prominence relations brought about by a mixture of different prosodic means such as pitch range compression after FOCUS (and possibly wider pitch excursion on the focused item), prosodic phrasing, intensity and duration manipulation, or a trade-off between these factors may be involved to render the rhematic FOCUS domain more prominent than the focus of interest contained in a topic expression. It has to be noted, however, that these prominence relations do not obligatorily arise. Thus it is not in the first place the relative prominence that is used to mark rhematic FOCUS, but rather the directionality of the  $f_0$ -movements (El Zarka 2011a). Accordingly, a principled distinction between rising  $f_0$  for topics (containing a focus) and falling  $f_0$  for rhematic FOCUS is assumed.

Another important finding is that deaccenting given elements is not necessary for achieving meaningful prominence relations (El Zarka 2011b). The question whether these meaningful prominence relations are confined to instances of contrastive focus or whether they may also be characteristic of information focus will be dealt with in the next phase of the project.

Evidence for the prosodic reflexes of information focus in EA also comes from the fact that *in situ* focus without accompanying changes in the syntactic or morphological structure is allowed. Unlike Standard Arabic or other Arabic vernaculars, EA even permits question words *in situ*.

(16) Q: *fuft* **MI:N** *fi manzil lamlu:m?*  
 saw.you who in house Lamlum  
 'Who did you see in Lamlum's house?'

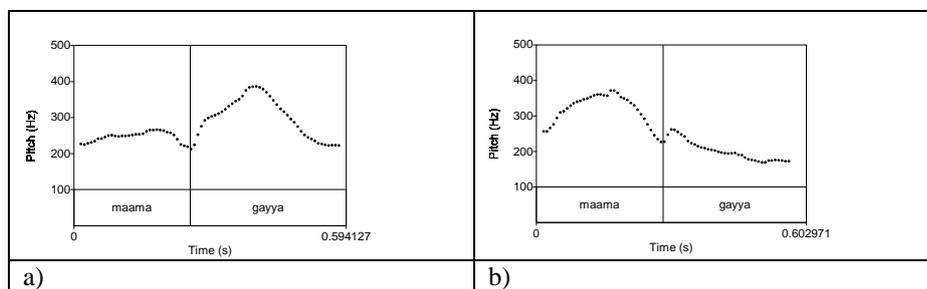
*fuft* **MARJAM** *fi manzil lamlu:m*  
 saw.I Maryam in house Lamlum  
 'I saw Maryam in Lamlum's house.'

(own corpus, production experiment)



larger corpus of spontaneous speech, a task the present study was designed for. It however turned out that the number of thetic utterances encountered in the corpus is not very high and should be amended by more data and, if possible, by experimental data - that is, if a convincing design for such an experiment can be developed. So far, we may say that even if thetic utterances do not inevitably exhibit a special prosody, we may venture the hypothesis that thetic utterances strongly prefer a certain type of intonation contour which may be called integrating, following Bolinger's distinction between tonal integration and separation (Bolinger 1986). This is equal to the claim that thetic utterances resist typical topic-comment prosody, especially the selection of a leading contour for the subject NP. An integrating contour is shown in a thetic utterance from the investigation corpus in panel (a) of Figure 2, as contrasted to a (contrastive) predicate focus contour in panel (b).

Fig.2: Panel a) shows the answer *ma:ma gajja* 'That's not possible. MUM is coming.' to 'How about going to the cinema tonight?' and panel b) shows the answer *ma:ma gajja* '(No,) mom's COMING' to the question 'Is mom leaving?'



In section (4.3) it has been mentioned that some existential constructions, however, exhibit topic-comment prosody. These typically contain two separate propositions that are also rendered as such. Thus the first part of sentence in (18), the existential construction, establishes a discourse referent and the second part predicates something to that referent.

- (18) *fi:h ra:gil t'awi:l/ bijjifrab saga:jir gamb in-nahr*  
 EXIST man tall / he.is.drinking cigarettes beside the-river  
 'There is a tall man smoking beside the river.'

(own corpus, QUIS data)

Contrary to the above quoted example (13c), this sentence cannot be viewed as a monolithic event description, as it clearly involves a predication about the tall man, whereas the 'coming to Mona' in example (13c) can hardly be regarded as something predicated of the suitor. Examples like (18) raise the issue of constraints on the occurrence of thetic clauses, such as the number of arguments and the influence of lexical semantics. Sentences consisting of two or more propositions have to be viewed as complex constructions as far as their information structure is concerned. Consequently, I do not consider this type of sentence a thetic construction, as it does

not fulfill the main formal criterion for the definition of a thetic statement, i.e. *integration* (Fuchs 1976, 1980; Jacobs 1993). The combination of a thetic presentational construction that introduces a participant who serves as the topic of a topic-comment sentence nicely illustrates what Chafe (1994) called the *One New Idea Constraint* or Lambrechts's (1994) *Principle of the Separation of Reference and Role*.

#### Abbreviations:

EA	Egyptian Arabic	SFB	Sonderforschungsbereich
		632	Informationsstruktur (information structure)
EXIST	existential marker		
F	feminine	D2	D2 (Typology of Information Structure), subproject of the SFB
M	masculine		632
MSA	Modern Standard Arabic	QUIS	questionnaire developed by the D2 project of SFB 632
NEG	negation		
P	plural		
POL	polite expression		
POSS	possessive marker		
PRT	particle		
REL	relative particle		
S	singular		
SV	subject-verb order		
VS	verb-subject order		

## 5. References

- Al-Khamissi, Khaled (2006). *taksi. ḥawadi:t al-mafawir*. Cairo: Dar El Shorouk.
- Benmamoun, Elabbas (2000). *The feature structure of functional categories: A comparative study of Arabic dialects*. New York, etc.: Oxford University Press.
- Boersma, Paul & David Weenink (2007). *Praat: doing phonetics by computer* (Version 4.5.21) [Computer program]. Retrieved 27-April-2007, from <http://www.praat.org/>.
- Bolinger, Dwight (1958). A Theory of Pitch Accent in English. *Word*, 14, 109-149.
- Bolinger, Dwight (1986). *Intonation and its parts: Melody in spoken English*. London, etc.: Arnold.
- Brazil, David (1975). *Discourse intonation: Discourse analysis monographs no. 1*. Birmingham: University of Birmingham.
- Brazil, David (1997). *The communicative value of intonation in English*. Cambridge, etc.: Cambridge Univ. Press.
- Brustad, Kristen (2000). *The syntax of spoken Arabic: A comparative study of Moroccan, Egyptian, Syrian, and Kuwaiti dialects*. Washington, DC: Georgetown University Press.

- Büring, Daniel (1997). *The meaning of topic and focus. The 59th street bridge accent*. London, New York: Routledge.
- Büring, Daniel (2003). On D-Trees, Beans, and B-Accents, *Linguistics & Philosophy*, 26/5, 511-545.
- Chafe, Wallace L. (1974). Language and consciousness. *Language*, 50, 111–133.
- Chafe, Wallace L. (1976). Givenness, contrastiveness, definiteness, subjects, topics and point of view. In Charles Li (Hg.), *Subject and topic*. New York: Academic Press, pp. 25–56.
- Chafe, Wallace L. (1994). *Discourse, consciousness and time: The flow and displacement of conscious experience in speaking and writing*. Chicago, Ill., etc.: Univ. of Chicago Press.
- Cowell, Mark W. (1964). *A reference grammar of Syrian Arabic: (based on the dialect of Damascus)*. Washington, DC: Georgetown Univ. Press.
- Cruttenden, Alan (1997). *Intonation*. Cambridge, etc.: Cambridge Univ. Press.
- Cruttenden, Alan (2006). The de-accenting of given information: A cognitive universal? In Giuliano Bernini & Marcia L. Schwartz (Hgg.), *Pragmatic organization of discourse in the languages of Europe: EUROTYP*. Berlin, etc.: de Gruyter, pp. 311–355.
- Dik, Simon C. (1978). *Functional Grammar*. Amsterdam: North-Holland.
- Dilley, Laura C. (2005). *The phonetics and phonology of tonal systems*. Ph.D thesis, Harvard University: MIT Division of Health Sciences and Technology.
- Dipper, Stefanie, Michael Götz & Stavros Skopeteas (Hgg.) (2007). *Information Structure in Cross-Linguistic Corpora: Annotation Guidelines for Phonology, Morphology, Syntax, Semantics, and Information Structure*. Potsdam: Universitätsverlag.
- Downing, Laura J. (2003). Stress, Tone and Focus in Chichewa and Xhosa. *Frankfurter Afrikanistische Blätter*, 15, 59–81.
- Edwards, Malcolm (2006). Pronouns, agreement and focus in Egyptian Arabic. *SOAS Working Papers in Linguistics*, 14, 51–62.
- Eid, Mushira (1983). The copula function of pronouns. *Lingua*, 59, 197–207.
- Eid, Mushira (1991). The Copula Pronoun in Arabic and Hebrew. In Bernard Comrie & Mushira Eid (Hgg.), *Perspectives on Arabic Linguistics III. Papers from the Third Annual Symposium on Arabic Linguistics, Salt Lake City, Utah 1989*. Amsterdam, Philadelphia: John Benjamins, pp. 31–61.
- Eisele, John C. (1999). *Arabic verbs in time: Tense and aspect in Cairene Arabic*. Wiesbaden: Harrassowitz.
- Eltayeb, Tarek (1992). *Al-ʿasane:r. Masrahiyya bi-l-ʿaammiyya l-misʿriyya*. [Al-qaahira (Cairo)].
- El Zarka, Dina (2005). *Rhythmische und tonale Strukturen im Ägyptischen Arabisch: eine Sprache im Viervierteltakt?* Paper held at the *Österreichische Linguistiktagung (ÖLT 2005)*, University of Graz.
- El Zarka, Dina. (2011a). Leading, linking, and closing tones and tunes in Egyptian Arabic – what a simple intonation system tells us about the nature of intonation. In Ellen Broselow & Hamid Ouali (Hgg.) *Perspectives on Arabic Linguistics: Papers from the annual symposia on Arabic Linguistics*. Volume XXII-XXIII:

- College Park, Maryland, 2008 and Milwaukee, Wisconsin, 2009, Amsterdam, Philadelphia: John Benjamins, pp. 57–74.
- El Zarka, Dina. (2011b). Prosodic encoding of thethetic/categorical distinction in Egyptian Arabic: a preliminary investigation. *Grazer Linguistische Studien* 76, 91–111.
- El Zarka, Dina (2012). One *ra* - many meanings. Syntax, semantics and prosody of the Moroccan modal particle *ra* and its Egyptian Arabic counterparts. STUF, Akademie Verlag 65/4: 410–429.
- Féry, Caroline (1993). *German intonational patterns*. Tübingen: Niemeyer.
- Fuchs, Anna (1976). 'Normaler' und 'kontrastiver' Akzent. *Lingua* 38, 293–312.
- Fuchs, Anna (1980). Accented subjects in 'all-new' utterances. In Gunter Brettschneider & Christian Lehmann (Hgg.): *Wege zur Universalienforschung. Sprachwissenschaftliche Beiträge zum 60. Geburtstag von Hansjakob Seiler*. Tübingen: Narr (Tübinger Beiträge zur Linguistik, 145), pp. 449–461.
- Gårding, Eva (1983). A generative model of intonation. In Anne Cutler & D. Robert Ladd (Hgg.), *Prosody: Models and Measurements, Prosody: Models and Measurements*. Berlin, Heidelberg, New York, Tokyo: Springer, pp. 11–26.
- Gary, Judith O. & Saad M. Gamal-Eldin (1982). *Cairene Egyptian colloquial Arabic*. Amsterdam: North-Holland.
- Givón, Talmy (1983). *Topic continuity in discourse: A quantitative cross-language study*. Amsterdam, etc.: Benjamins.
- Grice, Martine, Stefan Baumann & Ralf Benz Müller (2005). German Intonation in Autosegmental-metrical Phonology. In Sun-Ah Jun (Hg.), *Prosodic typology. The phonology of intonation and phrasing*. Oxford, New York: Oxford University Press, pp. 55–83.
- Gundel, Jeannette (1976). *The role of topic and comment in linguistic theory*. Bloomington, Indiana: Indiana University Linguistics Club.
- Gussenhoven, Carlos (1983a). *A semantic analysis of the nuclear tones of English. Distributed by Indiana University Linguistics Club (IULC)*. Bloomington, Indiana (Published as ch. 6 in Gussenhoven (1984)).
- Gussenhoven, Carlos (1983b) Focus, mode and the nucleus. *Journal of Linguistics*, 19, 377–417.
- Gussenhoven, Carlos (1984). *On the grammar and semantics of sentence accents*. Dordrecht, etc.: Foris Publications.
- Halliday, Michael A. K. (1967). Notes on transitivity and theme in English, Parts 1 and 2. *Journal of Linguistics*, 3, 37–81, 199–244.
- Halliday, Michael A. K. (1968). Notes on transitivity and theme in English, Part 3. *Journal of Linguistics*, 4, 179–215.
- Hellmuth, Sam (2006a). *Pitch accent distribution in Egyptian Arabic*. Ph.D. dissertation, SOAS.
- Hellmuth, Sam (2009). The (absence of) prosodic reflexes of given/new information status in Egyptian Arabic. In Jonathan Owens & Alaa Elgibali (Hgg.), *Information Structure in Spoken Arabic, Information Structure in Spoken Arabic*. London: Routledge, pp. 165–188.

- Hellmuth, Sam (2010). Functional complementarity is only skin deep: evidence from Egyptian Arabic for the autonomy of syntax and phonology in the expression of information structure. In Nomi Erteschik-Shir & Lisa Rothman (Hgg.), *The Sound Patterns of Syntax*. Oxford etc.: Oxford Univ. Press, pp. 247-270.
- Hockett, Charles F. (1958). *A course in modern linguistics*. New York: Macmillan.
- Holes, Clive (1995). *Modern Arabic: Structures, functions and varieties*. London, etc.: Longman.
- Holes, Clive (2009). Word order and textual function in Gulf Arabic. In Jonathan Owens & Alaa Elgibali (Hgg.), *Information Structure in Spoken Arabic*. London: Routledge, pp. 61-74.
- Jacobs, Joachim (1992). *Informationsstruktur und Grammatik*. Opladen: Westdeutscher Verlag.
- Jacobs, Joachim (1993). Integration. In: Marga Reis (Hg.): *Wortstellung und Informationsstruktur*. Tübingen: Niemeyer (Linguistische Arbeiten, 306), pp. 63-116.
- Jacobs, Joachim (1997). I-Topikalisierung. *Linguistische Berichte*, 168, 91-133.
- Jelinek, Eloise (2002). Agreement, clitics and focus in Egyptian Arabic. In Jamal Ouhalla and Ur Shlonsky (Hgg.), *Themes in Arabic and Hebrew syntax: Studies in natural language and linguistic theory*, v. 53. Dordrecht, Boston, MA: Kluwer Academic Publishers, pp. 71-105.
- Krifka, M. (2007). Basic notions of information structure. *Interdisciplinary Studies on Information Structure (ISIS)*, 6, 13-55.
- Kuroda, S.-Y. (1972). The categorial and the thetic judgment reconsidered: evidence from Japanese syntax. *Foundations of Language*, 9, 153-185.
- Ladd, Robert D. (1996). *Intonational phonology* (1. publ.). Cambridge, etc.: Cambridge Univ. Press.
- Lambrecht, Knud (1994). *Information structure and sentence form: Topic, focus and the mental representations of discourse referents*. Cambridge, etc.: Cambridge Univ. Press.
- Langacker, Ronald W. (1993). Reference-point constructions. *Cognitive Linguistics*, 4(1), 1-38.
- Li, Charles & Sandra Thompson (1976). Subject and topic: a new typology of language. In Charles Li (Hg.), *Subject and topic, Subject and topic*. New York: Academic Press, pp. 457-490.
- Maslova, Elena & Giuliano Bernini (2006). Sentence topics in the languages of Europe and beyond. In Giuliano Bernini & Marcia L. Schwartz (Hgg.), *Pragmatic organization of discourse in the languages of Europe: EURO TYP*. Berlin, etc.: de Gruyter, pp. 67-120.
- Matras, Yaron & Machtelt Bolkestein (2006). Deixis and anaphora: Some case studies. In Giuliano Bernini & Marcia L. Schwartz (Hgg.), *Pragmatic organization of discourse in the languages of Europe: EURO TYP*. Berlin, etc.: de Gruyter, pp. 215-253.
- Matras, Yaron & Hans-Jürgen Sasse (Hgg.) (1995). *Verb-Subject Order and Theticity in European Languages*. Berlin: Akademie Verlag.
- Mitchell, Terrence F. (1993). *Pronouncing Arabic*. Oxford: Clarendon Press.
- Molnár, Valéria. (1998). Topic in focus. In *Acta Linguistica Hungarica*, 45, 89-166.

- Molnár, Valéria. (2006). On different kinds of contrast. In Valéria Molnár & Susanne Winkler (Hgg.), *The Architecture of Focus. Studies in Generative Grammar* 82, Berlin, New York: Mouton de Gruyter, pp.197-233.
- Moutaouakil, Ahmed (1989). *Pragmatic functions in a functional grammar of Arabic*. Dordrecht, Holland, Providence, RI, U.S.A.: Foris Publications.
- Mughazy, Mustafa (2008). The pragmatics of denial: An information structure analysis of so-called “emphatic negation” in Egyptian Arabic. In Dilworth B. Parkinson (Hg.), *Perspectives on Arabic linguistics XXI. Papers from the Twenty-first Annual Symposium on Arabic Linguistics, Provo, Utah, March 2007*. Amsterdam, Philadelphia: John Benjamins, 69–81.
- Mughazy, Mustafa (2009). The information structure of existential sentences in Egyptian Arabic. In J. Owens & A. Elgibali (Hgg.), *Information Structure in Spoken Arabic, Information Structure in Spoken Arabic*. London: Routledge, pp. 107-120.
- Norlin, Kjell (1989). A preliminary description of Cairo Arabic intonation of statements and questions. *Quarterly Progress and Status Report (QPSR)*, 47–49.
- Ouhalla, Jamal (1993). Focus in Standard Arabic. *Linguistics in Potsdam*, 1, 65–92.
- Owens, Jonathan et al. (2009). Explaining Ø and Overt Subjects in Spoken Arabic. In Jonathan Owens & Alaa Elgibali (Hgg.), *Information Structure in Spoken Arabic*. London: Routledge, pp. 20-60.
- Pierrehumbert, Janet (1980). *The phonology and phonetics of English intonation*. PhD thesis, Massachusetts Institute of Technology.
- Prince, Ellen (1981). Towards a taxonomy of given-new information. In Peter Cole (Hg.), *Radical Pragmatics*. New York: Academic Press, pp. 223-255.
- Rastegar-El Zarka, Dina (1997). *Prosodische Phonologie des Arabischen*: Ph.D. thesis, University of Graz.
- Reis, Marga (Hg.) (1993) *Wortstellung und Informationsstruktur*. (Linguistische Arbeiten, 306). Tübingen: Niemeyer.
- Reinhart, Tanya (1982). *Pragmatics and linguistics: an analysis of sentence topics*. Bloomington, Indiana: Indiana University Linguistics Club.
- Rifaat, Khaled (1991). *The Intonation of Arabic: An Experimental Study*. Ph.D. thesis, University of Alexandria.
- Sasse, Hans-Jürgen (1987). Thethetic/categorical distinction revisited. *Linguistics*, 25, 511–580.
- Sasse, Hans-Jürgen (1995). “Theticity” and VS order: a case study. In Yaron Matras & Hans-Jürgen Sasse (Hgg.), *Verb-Subject Order and Theticity in European Languages. (Sprachtypologie und Universalienforschung* 48 [1/2]). Berlin: Akademie Verlag, pp. 3–31.
- Sasse, Hans-Jürgen (2006). Theticity. In Giuliano Bernini & Marcia L. Schwartz (Hgg.), *Pragmatic organization of discourse in the languages of Europe: EUROTYP*, Berlin, etc.: de Gruyter, pp. 255–308.
- Schmidt, Thomas & Kai Wörner (2005) Erstellen und Analysieren von Gesprächskorpora mit EXMARaLDA. *Gesprächsforschung*, 6, 171-195. ([www.exmaralda.org](http://www.exmaralda.org))

- 
- Selkirk, Elisabeth (1984). *Phonology and Syntax*. Cambridge, Massachusetts: MIT Press.
- Skopeteas, Stavros et al. (2006). Questionnaire on Information Structure (QUIS). *Interdisciplinary Studies on Information Structure (ISIS)*, (4).
- Uhmann, Susanne (1991). *Fokusphonologie: Eine Analyse deutscher Intonationskonturen im Rahmen der nicht-linearen Phonologie*. Tübingen: Niemeyer.
- Vallduví, Enric (1992). *The informational component*. New York: Garland.
- Woidich, Manfred (2006). *Das Kairenisch-Arabische: Eine Grammatik*. Wiesbaden: Harrassowitz.
- Xu, Yi (1999). Effects of tone and focus on the formation and alignment of F0 contours. *Journal of Phonetics*, 27, 55–105.
- Zerbian, Sabine (2006). *Expression of Information Structure in the Bantu Language Northern Sotho*, PhD thesis, Humboldt University, Berlin.