Intonational pitch accent distribution in Egyptian Arabic

Samantha Jane Hellmuth

Submitted to the Department of Linguistics in partial fulfilment of the requirements for the degree of

Doctor of Philosophy in Linguistics and Phonetics

School of Oriental and African Studies
University of London
March 2006

Abstract

Egyptian Arabic (EA) is a stress-accent language with postlexical intonational pitch accents. This thesis investigates EA pitch accents within the autosegmental-metrical (AM) framework (Ladd 1996). The goal of the study is to identify the place of EA in the spectrum of cross-linguistic prosodic variation, and to resolve the challenge it presents to existing phonological accounts of pitch accent distribution.

In a corpus of read and (semi-)spontaneous EA speech a pitch accent was found on (almost) every content word, and in the overwhelming majority of cases the same pitch accent type is observed on every word. The typological implications of EA pitch accent distribution are explored in the context of the typology of word-prosodic variation (Hyman 2001) and variation in the domain of pitch accent distribution is proposed as a new parameter of prosodic variation.

A survey of EA prosodic phrasing and of the relative accentuation of function words and content words shows that the correct generalisation for EA is that there is a pitch accent on every Prosodic Word (PWd). A phonological analysis is proposed within Optimality Theory (Prince & Smolensky 1993), formalising the two-way relation between tone and prosodic prominence at all levels of the Prosodic Hierarchy.

An experimental study suggests that alignment of the H peak in EA pitch accents varies with stressed syllable type (cf. Ladd et al 2000), and is analysed as phonological association of the pitch accent to the foot. A final experiment quantifies the prosodic reflexes of information and contrastive focus. Even when post-focal and 'given' EA words still bear a pitch accent, but there are gradient effects of focus in the form of pitch range manipulation. Independence of pitch accent distribution from information structure supports the formal analysis of EA pitch accent distribution within the phonological part of the grammar.

Table of Contents

Transliteration scheme used in the thesis for Egyptian Arabic

Arabic script	IPA symbol	Symbol used here
أ	ʔ	?
ب	b	b
ت	t	t
ث	θ / t / s	θ / t /s
ج	g	g
ح	ħ	H
خ	x	x
د	d	d
ذ	ð	ð
ر	r	r
ز	z	z
س	s	s
ش	ʃ	š
ص	sˤ	S
ض	dˤ	D
ط	tˤ	T
ظ	ðˤ	Z
ع	ʕ	9
غ	ɣ	G
ف	f	f
ق	q	q
ك	k	k
ل	l	l
م	m	m
ن	n	n
ه	h	h
و	w	w
ى	j	y

ا	aː	aa
ي	iː	ii
و	uː	uu
َ	a	a
ِ	i	i
ُ	u	u

Acknowledgements

I would like to thank the members of my PhD committee for their very generous support throughout the last four years - Monik Charette, Bruce Ingham, Andrew Simpson and Justin Watkins. I would particularly like to thank Justin Watkins and Laura Downing for their help in the final stages of the project and for their investment of time in giving comments on drafts of this thesis.

I am grateful to Lisa Selkirk for extending an invitation to me as an academic visitor to the Department of Linguistics, UMass (Amherst) during the Spring of 2004 and for her advice and encouragement during my stay. I am also grateful to the Department of Linguistics at Amherst for providing me with access to their LDC corpus during my stay and for use of recording facilities for one of the pilot studies.

Thanks are due to Bernard Howard and to the staff of the SOAS AV department for help with recording facilities for the pilot studies, and loan of recording equipment for my trip to Egypt, and to Janet Watson for the loan of Sanaani Arabic speech data.

This research would be nothing without the patience and goodwill of the speakers of Egyptian Arabic who assisted me, first in London for the pilot studies, and then in Egypt. Special thanks are due therefore to the staff and students of the English Training Centre, Maadi, Cairo, for welcoming me during July 2004 while I was collecting the experimental data, and especially to the core group of speakers who participated in both experiments, and who were a delight to work with. It was a special privilege to work alongside gifted Egyptian colleagues who were patient and generous in sharing their insights on and incredulity at the style and content of my experimental stimuli.

I am grateful to Professor Khaled Rifaat (University of Alexandria) & Dr Dina El Zarka (University of Graz) for their time and insights as native speaker linguists and for their welcome extended to me as a *xawaaga* working on their language. I have also been very much helped in the last four years by the wider intonation and prosody research community, particularly Sonia Frota, Carlos Gussenhoven and Pilar Prieto.

This work was supported by the UK Arts & Humanities Research Council (Postgraduate Award 59198), who also provided additional funding for my academic visit to UMass (Amherst). I would like to thank my godmother Barbara Pope whose generous financial gift at the start of my PhD bought the computer on which this entire study was recorded, analysed and written.

Thanks to my peers and colleagues, especially Carlo Colella, Birgit Hellwig, Friederike Lüpke, Sabine Zerbian and Ihab Shabana, for invaluable moral support as well as a wealth of practical help. Thanks also to my former teaching colleagues in Sanaa, Yemen, who encouraged me to take up serious study of linguistics. On a personal level I would like to thank the following for their friendship, humour and good company during the last four years: Jo & Jeannie, Gill, Helena, Sarah, Patrick and (latterly) Arthur B. Finally, I could not have carried this project through to completion without the unconditional support and love of my family: Lynn, Nikki, Dad - thank you. What follows is testament to my Mum from whom I learned the value of not giving up on trying to do what you really want to do.

1 Outline summary of the thesis

1.1 Aims of the study

Egyptian Arabic (EA) is the dialect of Arabic spoken in Cairo and by educated speakers throughout Egypt. The sentence phonology of EA is relatively under-researched and thus a descriptive goal of this thesis is to establish the phonological properties of EA words, phrases and sentences, by investigating their intonational properties.

The investigation reveals that in EA there is an intonational pitch movement (a 'pitch accent') associated with almost every content word, and this is argued to generalise to the association of a pitch accent with every Prosodic Word (PWd). The theoretical goal of the thesis is thus to identify the position of EA (and other languages which can be argued to share the property of having a pitch accent on every word) in the spectrum of cross-linguistic variation, with respect to both intonational typology and the typology of word prosodic types.

In pursuit of these goals a series of experimental investigations are reported in which the properties of EA pitch accents above and below the level of the word are explored. These are designed along the lines of 'laboratory phonology' (Ohala & Jaeger 1986) in which phonological theory informs the experimental hypotheses to be tested.

The thesis is couched in the framework of autosegmental-metrical theory (Ladd 1996) in which patterns of behaviour of phonological tones (pitch) are seen as a reflex of their autosegmental association to metrical/prosodic constituents in a hierarchy of prosodic structure. Working within Optimality Theory, a constraint-based theory of phonology (Prince & Smolensky 1993), the thesis argues that in fact all of the properties of EA pitch accents can be formalised within a particular notion of the relation between phonological tone and prosodic prominence (Selkirk 2004b).

1.2 Chapter 2

A key assumption of autosegmental-metrical theory (AM) and of the present work is the autosegmental nature of tone. This was argued for by Goldsmith (1976) in relation to lexical tones, which are said to be autosegmentally associated to elements of the metrical structure of the word (such as the mora or syllable). A key insight of Pierrehumbert (1980) was that the intonational pitch contour, that is, 'postlexical' tones, can also be successfully analysed as a series of pitch targets or tones, autosegmentally associated with the prosodic structure of the utterance.

Chapter 2 sets out the basic properties of AM theory and in particular of the notion of 'the unity of pitch phonology' (Ladd 1996:147ff.). This aspect of the theory means that it can be used to analyse any language, regardless of whether pitch functions lexically and/or postlexically. Some details of the theory of prosodic phonology, which addresses the interface between syntax and phonology, are also discussed in order to establish a working assumption as to which elements of the prosodic hierarchy are relevant for analysis of EA. The theory of violable ranked constraints, Optimality Theory, is briefly introduced, and prior work on the intonation and prosody of EA and of other Arabic dialects is reviewed in detail. The chapter concludes with a description of the corpus of speech data and discussion of issues involved in data collection.

1.3 Chapter 3

Chapter 3 reports the results of auditory transcription of a corpus of EA speech, which reveals that in EA a pitch accent occurs on (almost) every content word. This generalisation holds in a variety of contexts which, in other languages, would be conducive to 'de-accenting', such as in fast speech, in long rhythmic utterances, in non-neutral contexts and in spontaneous speech. The corpus comprises read speech sentences and paragraphs collected for use in chapters 7 & 8, as well as longer read narratives, narratives re-told from memory and spontaneous conversation.

The detail of the auditory transcription reveals that, not only is there a pitch accent on every content word in EA, but that it is in the overwhelming majority of cases, the same *type* of pitch accent (a rising pitch movement). These observations are incorporated into a preliminary working model of EA intonation, which is compared to alternative AM models of EA.

The co-occurrence of rich pitch accent distribution and high frequency of one pitch accent type has however been noted in other languages such as Spanish & Greek (Jun 2005b). The chapter concludes by suggesting that 'accent on every word' languages are a typologically valid category within the range of cross-linguistic intonational variation.

1.4 Chapter 4

Having established in chapter 3 that EA is a language in which every content word bears a pitch accent, Chapter 4 seeks to identify the position of EA in the range of variation of word prosodic typology. The chapter starts with a survey of cross-linguistic word-level prominence marking, in the context of both intonational typology and word-prosodic typology. The detail of prior descriptions and studies of EA word-level prominence is also reviewed, and suggests the uncontroversial hypothesis that EA is an accentual language, with both tonal and non-tonal phonetic correlates of word-level prominence.

An experimental study of the alignment of individual low (L) and high (H) pitch targets within EA rising pitch accents demonstrates that these pitch movements mark the accentual head of the word (the main stressed syllable), rather than, say, the edges of the word (which might suggest non-accentual pitch as observed in Korean).

A small post-hoc study of the phonetic correlates of word-level prominence, on a small subset of data from the same experimental corpus, suggests that pitch is not the only correlate of accent in EA: duration and intensity also consistently mark stressed syllables.

These studies confirm the hypothesis that EA is a stress-accent language (in the terms of Beckman 1986). EA does not however fit neatly into any existing word-prosodic categories, since it has consistently more richly populated pitch accent distribution than ‘archetypal’ stress accent languages such as English. The chapter concludes by arguing that EA forces us to propose density of pitch accent distribution as an additional parameter of prosodic variation, and the predictions of such a proposal are explored.

1.5 Chapter 5

Chapter 5 starts by reviewing the main competing theoretical explanations for pitch accent distribution. A particular area of debate within AM theory concerns the relationship between focus and prosodic prominence. In Germanic languages focus affects pitch accent distribution (words which are not focussed are ‘de-accented’) and some authors have therefore argued that the position of pitch accents directly reflects the position of focus (as determined by 'F(ocus)-marking' in syntactic structure, for example, in Selkirk 1984). Contrasting data from other languages has led other authors to propose a 'structure-based' view (Ladd 1996), in which the distribution of pitch accents arises strictly from the distribution of prosodic constituents (focus may affect prosodic phrasing, and thus lead to de-accenting, but only indirectly). It is argued here that the absence of de-accenting in EA provides new evidence in favour of the structure-based view of pitch accent distribution.

Within the structure-based view, various constituent levels in the Prosodic hierarchy have been suggested to function as the domain of pitch accent distribution, in analyses of different languages. These analyses are reviewed, and then the main body of the chapter explores empirical evidence from EA in detail, in order to decide which constituent level of the Prosodic Hierarchy might be the correct domain of pitch accent distribution in EA: Major Phrase, Minor Phrase or Prosodic Word.

Evidence from prosodic phrasing in EA, from a pilot study and from materials in the thesis corpus, indicates that in EA MaPs are generally very large, containing up to 8 PWds, all of which routinely bear a pitch accent. The MaP thus cannot be the domain of pitch accent distribution in EA.

Analysis of the phrasing data in terms of violable constraints on prosodic well-formedness (rhythmic constraints) and on the interface between syntax and phonology, supports a rhythmically-based working definition of the MiP in EA as minimally branching ('BINMIP'), and thus composed of at least two PWds. Both such PWds again routinely bear a pitch accent, suggesting that the MiP is not the domain of pitch accent distribution in EA either.

Finally, a thorough survey of the accentuation of function words in the corpus is presented. In comparison with analyses of the treatment of function words in English (Selkirk 1996) and Standard Serbian (Zec 2002), evidence is provided to demonstrate that function words procliticise within the MiP to a following PWd, mapped from a lexical content word. The prosodic constituent which acts as the domain of pitch accent distribution in EA is thus the Prosodic Word (PWd).

The chapter concludes with a brief discussion of why, on functional grounds, the phonology of EA might choose to use pitch to mark the PWd level, rather than, as other languages do, some higher level in the Prosodic Hierarchy.

1.6 Chapter 6

Having established in chapter 5 that the domain of pitch accent distribution in EA is the PWd, chapter 6 presents a formal OT analysis of the facts of EA. To do so the chapter starts by reviewing formalisms that have been proposed in the literature to explain variation in the choice of target for the association of phonological tone, with choices ranging across the constituents of the Prosodic Hierarchy.

A formalism of the relation between phonological tone and prosodic prominence is proposed within Optimality Theory using a pair of inherently ranked fixed hierarchies of positive markedness constraints, which regulate the association of tone to prosodic prominence (T→P), and of prosodic prominence to tone (P→T), respectively (following an idea suggested by Selkirk 2004b).

Specifically it is proposed that in EA the presence of pitch accents on every PWd arises because a P→T constraint, requiring the head of every PWd to be associated with tone, is ranked above the faithfulness constraint militating against the insertion of (postlexical) tone: PWD→T >> DEP_{TONE}.

Interaction with interface constraints on the mapping of lexical words to PWds explains why, in general, only content words bear a pitch accent in EA: PWD→T, LEXWD:PWD >> DEP_{TONE}. Interaction with rhythmic well-formedness constraints is argued to account for the 'promotion' of a small number of function words to PWd status, with the result that they bear a pitch accent: NOLAPSE >> LEXWD:PWD. An analysis is also proposed to capture proclisis, rather than enclisis, of function words in EA.

The chapter concludes by arguing that an advantage of using markedness constraints is the ability to directly encode the "unity of pitch phonology" (Ladd 1996:147ff.): tones of any origin, lexical or postlexical, will be subject to the same constraints. In addition, two specific predictions of the proposal are spelled out: firstly that there will be effects of not only of P→T constraints but of T→P constraints in languages with postlexical tone; and secondly, that EA rich pitch accent distribution arises due to a purely *phonological* constraint (PWD→T). Chapters 7 & 8 investigate these predictions empirically in EA, exploring surface pitch accent alignment, and gradient as well as categorical reflexes of focus, respectively.

1.7 Chapter 7

The interaction of tone and prominence is argued in chapter 6 to be governed by positive markedness constraints, which are by definition ambivalent as to whether or not the tones whose distribution they constrain are present in the input or not (lexical or postlexical).

Thus, even in intonational languages, in which all tones are 'postlexical' and not present in the input, the tone bearing unit (TBU) may vary, since, according to the T→P fixed hierarchy, tone could in principle be required to associate to any level of the prosodic hierarchy.

In general the assumption in AM theory is that pitch accents are associated with the main stress foot of accented words, and that this association is inherited by the stressed syllable. Chapter 7 reports an investigation designed to confirm in particular whether the stressed syllable or the foot is the TBU in EA. In addition, the investigation also facilitates proposal of a formal phonological representation for EA rising pitch accents, and feeds into current debate in AM theory regarding the mechanisms regulating surface alignment of pitch accents, as a reflex of underlying association of phonological targets to elements of prosodic structure (cf. Ladd 2003).

Experimental data were elicited in which the prosodic weight of the stressed syllable of target words was systematically varied (CV vs. CVC vs. CVV), and reveals the precise patterns of alignment of L and H targets in EA rising pitch accents. These are broadly similar to results observed for Dutch, in that alignment of the H peak appears to vary with syllable type (Ladd et al 2000), and it is argued that EA rising pitch accents associate with the foot, rather than the stressed syllable.

In terms of formal T↔P analysis, this indicates that in EA, as well as in Dutch, the constraint requiring any tones to be associated to the head of the PWd (T→PWD) is ranked higher than DEP_{TONE}. The separation of T→P and P→T constraints, within T↔P theory, predicts what proves to be the case: that it is possible for Dutch and EA to share the property of having pitch accents which associate to the same TBU, but not to share the property of rich pitch accent distribution.

The chapter concludes by discussing the correct phonological representation of EA rising pitch accents and suggests that they are best represented as a single phonological object (a bitonal pitch accent: L+H*), phonologically associated with the foot as TBU.

1.8 Chapter 8

As discussed in chapter 5, a common view of pitch accent distribution has been that pitch accents are inherently focus-marking (cf. discussion in Ladd 1996:221 ff.). The claim that pitch accent distribution in EA arises due to a purely phonological constraint (PWD:T) raises the question: how is focus expressed in EA? Chapter 8 reports the results of an experiment designed to clarify if and how focus is marked in EA.

Target sentences were placed in contexts to manipulate both information focus and contrastive focus status of target words, since a distinction between these two types of focus has been argued to be relevant syntactically in Arabic (Moutouakil 1989, Ouhalla 1997, Kiss 1998). Qualitative analysis of this focus dataset, reported in chapter 3, showed that even when post-focal and 'given', EA words bear a pitch accent, and thus that there appears to be no categorical reflex of either information focus or contrastive focus in EA.

Quantitative analysis of the focus dataset reveals that there are gradient effects of contrastive focus (only), in the form of pitch range manipulation: pitch range is expanded in focussed words and compressed in post-focal words. After some discussion as to whether pitch range manipulation is a phonetic or phonological property, it is suggested that in EA pitch range manipulation is a gradient phenomenon, under the control of speakers.

The chapter concludes by discussing the typological implications of the apparent lack of any prosodic reflex of information focus in EA (neither categorical nor gradient), in the context of the syntactic properties of the language.

1.9 Chapter 9

The descriptive facts of pitch accent distribution in EA, as well as the specific theoretical explanation proposed in the thesis, are summarised in the concluding chapter. The main contributions of the thesis are briefly explored, as well as potential avenues of future investigation.

2 Literature review

2.0 Outline of the chapter

The aim of this thesis is to demonstrate a property of Egyptian Arabic (EA) that appears to have gone largely un-remarked (namely the density of its pitch accent distribution) and to identify the correct location of a language with such a property in the continuum of cross-linguistic prosodic variation. This chapter reviews the relevant background literature to this task.

The thesis is framed within the autosegmental-metrical theory (AM) of intonation. A fundamental concept in AM theory is the autosegmental nature of tone. This was argued for by Goldsmith (1976) with respect to lexical tones, which are said to be autosegmentally associated to elements of the metrical structure of the word (such as the mora or syllable). The insight which forms the basis of AM theory, is that intonational pitch, that is, 'postlexical' tones, can also be successfully analysed as a series of autosegmental pitch targets or tones, phonologically associated with positions in the prosodic structure of the utterance.

The basic properties of AM theory are set out in section 2.1 and in particular how they relate to the notion of 'the unity of pitch phonology' (Ladd 1996:147ff.). This aspect of the theory means that it can be used to analyse any language, regardless of whether pitch functions lexically and/or postlexically. AM theory is thus argued to better facilitate cross-linguistic comparison among potentially typologically different languages. EA is a stress accent/intonational language, in which intonation is expected to interact to some degree to be determined with syntactic and semantic structure, and theories of the syntax-phonology interface and its influence on prosodic structure are therefore also reviewed.

The framework of analysis used in the thesis is the theory of violable ranked constraints, Optimality Theory (OT, Prince & Smolensky 1993). Section 2.2 introduces the basic notions of OT, and the implementation of AM and prosodic phonology within it, with a survey of constraints that have been argued to govern prosodic structure and its interaction with other parts of the grammar.

A survey of prior work on EA intonation in section 2.3, as well as on the intonation of other Arabic dialects, motivates the initial research questions addressed by the thesis. A number of these studies have mentioned the close relationship between pitch and word-level prominence in EA, resulting in dense pitch accent distribution, but no prior study has sought to account for this property in the context of the relationship between prosodic structure and morphosyntactic structure.

This thesis thus fills a two-way gap in the literature: firstly, the typology of intonational and prosodic phonology lacks information about the reflexes of prosodic structure in Arabic dialects in general, and in EA specifically; secondly, whilst the segmental phonology and metrical phonology of EA has been much discussed, the properties of the phonological component of the grammar of EA above the level of the word has received relatively little attention.

The chapter concludes in section 2.4 by discussing practical issues that had to be addressed in collecting the corpus of EA speech data on which the thesis is based.

2.1 An autosegmental-metrical theory of intonation

The title of this thesis refers to 'intonational pitch accents'. The notion of a 'pitch accent' in an intonational language (defined as a pitch movement which displays association to a stressed syllable), is inherent to the Autosegmental-Metrical (AM) theory of intonation, within which the thesis is framed.

Due to an overlap of terminology, the term 'pitch accent' has other meanings, so that one might talk about a 'lexical pitch accent language', which refers to a language in which tone plays a partial role in the lexical specification of some morphemes. Pitch accent languages include Japanese and Swedish, in which the meaning of certain words depends on the *position* and/or *type* of pitch movement produced on or near the accented syllable of the word. EA is by no means a lexical pitch accent language, but an apparently straightforward intonational language of the most ordinary kind. The pitch contour associated with EA utterances contributes only postlexical meaning, at the level of the sentence or utterance, and not at the level of individual lexical items.

The claim of this thesis is that EA pitch accents are of typological and theoretical interest not for their function (they are clearly postlexical) nor their variety (they are mostly of a single type), but for their ubiquitous presence on every content word in every utterance. In a sense then EA pitch accents are not in themselves all that interesting, instead, it is the density of their distribution that is interesting, and apparently unexpected, within existing prosodic and intonational theory.

The remainder of this section sets out in greater detail what is meant by a pitch accent in the AM theory of intonation, and why this particular theory of intonation is adopted in this thesis as the best starting point for understanding what is interesting about EA pitch accents.

2.1.1 Pitch accents and boundary tones

A fundamental notion within AM theory is that the surface pitch contour spread across an utterance arises from a linear sequence of 'pitch events'. In a tone language some of these events will be lexically specified, but in an intonational language like EA, the tonal events are expected to be of two kinds only: i) pitch accents: pitch movements which are phonologically associated to stressed syllables; and, ii) edge tones: pitch movements which align to the edges of phrase-level constituents.

The 'affiliation' of surface pitch events to certain positions in the utterance is easiest to see when a parallel sequence of tonal events (pitch accent + edge tone) is assigned to utterances of different lengths. The following examples are taken from Ladd (1996:44). In (2.1a) below speaker B expresses incredulity that 'Sue' should take on a certain career, and the full sequence of tonal events is squeezed on to a single syllable: a rise to a high peak followed by a fall and then another rise. If the same incredulity is expressed over a sequence containing more syllables, as in (2.1b), the tones spread out across the extra syllable, but do not do so evenly - the high peak and following fall stay localised around the main stressed syllable and the rise stays localised at the end of the phrase. There are clearly two separate tonal events - a rise-fall pitch accent associated with the main stressed syllable, and a rising edge tone. In AM theory both (2.1a) and (2.1ab would receive the same analysis, a rise-fall pitch accent (such as H*+L) followed by an edge tone (H%)[1]:

[1] See section 2.1.5 above (footnote 9) for an alternative view of the tones making up this tonal sequence.

(2.1) a. A: I hear Sue's taking a course to become a driving instructor.

B: Sue!?

b. A: I hear Sue's taking a course to become a driving instructor.

B: A driving instructor!?

2.1.2 Autosegmental postlexical tone

The main reason for choosing to analyse the facts of EA intonation within the autosegmental-metrical (AM) theory of intonation is its ability to capture the notion of 'the unity of pitch phonology' (Ladd 1996:147ff.). A theory which can be used to analyse any language, regardless of whether pitch functions lexically and/or postlexically, is better equipped to facilitate insightful cross-linguistic comparison among potentially typologically different languages.

Intonation has been defined as "the use of suprasegmental phonetic features to convey 'postlexical' or sentence-level pragmatic meanings in a linguistically-structured way" (Ladd 1996:6). By this definition intonation is observed in all types of language across the continuum of variation between an archetypal 'tone' language and an archetypal 'stress' language. In a tone language, such as Mandarin or Thai, pitch is part of the lexical specification of some if not all morphemes (cf. Hyman 2001). In a stress language such as English or EA, pitch may feature among the phonetic correlates of word- or phrase-level metrical prominence, but pitch is not itself part of the lexical specification of any morphemes[2]. A tone language may thus display postlexical as well as lexical tones, alongside each other. Postlexical use of tone in tone languages includes pitch register or pitch range shifts, blocking of downstep at phrase boundaries and insertion of boundary tones at phrase edges (Yip 2002:271ff.).

The autosegmental nature of phonological tones (in tone languages) has been established for some time (Goldsmith 1976). Tones are known to behave differently from other lexically specified segments: tones display stability (the tone remains if structure is deleted) and mobility (tones move to available positions in structure), as

[2] The category of 'accentual' languages, which have stress, but also display tone lexically assigned to certain words, is here assumed to be a sub-type of tone language (Hyman 2001, Yip 2002).

well as the properties of spreading (more than one tone realised on a single segment) and floating (an unrealised tone). These properties together are argued to be best explained in terms of autosegmental association of the tones to elements of the metrical structure of the word, such as the mora or syllable (Yip 2002:72-77).

The insight which forms the basis of AM theory is that intonational pitch, that is, 'postlexical' tones, can also be successfully analysed as a series of pitch targets or tones, autosegmentally associated with the prosodic structure of the utterance[3]. In AM theory intonationally meaningful pitch movements are represented as a sequence of phonological tones which *associate* autosegmentally with prominent positions in metrical structure, and/or with the edges of metrically or prosodically defined constituents. Pitch movements may not phonetically *align* exactly to these anchor points, but their *association* is phonologically specified in terms of metrical structure.

These notions are of interest for the present study because, in a language in which every word is marked with pitch one could pursue a language-specific analysis whereby pitch is simply among the phonetic correlates of lexical stress in EA. The claim of AM theory, that the surface pitch contour of any utterance reflects a sequence of phonological tones, forces a more in-depth study of EA: if all tones are autosegmental phonological objects what accounts for the fact that EA has so many of them?

At this point it is worth clarifying some terminological choices. In this thesis the term *stress* is used to denote word-level prominence (lexical stress). An *accent* (or pitch accent) denotes a tonally marked prominence: the main stressed syllable of one or more words in a phrase or utterance which are additionally marked by a movement in pitch. Since in EA all stressed syllables will be found to bear an accent, these two terms (*stress* and *accent*) are in practice usually interchangeable for EA.

The term *nucleus* is used to denote the main prominence in an utterance, also called the *nuclear accent.* In some versions of AM theory the nucleus is assumed always to be the final accent in an intonational phrase[4]; this is not assumed here. Nonetheless in most

---

[3] This insight was demonstrated for English by Liberman (1975) and Pierrehumbert (1980), and for the interaction of lexical and postlexical accents in Swedish by Bruce (1977).
[4] For example, by Beckman & Pierrehumbert (1986, 1988) and within the ToBi transcription system for English (Beckman & Elam 1993).

cases nuclear accents are observed to be final, and thus the term *pre-nuclear accent* will always denote a non-phrase-final accent.

2.1.3 Pitch targets analysed as level tones

In AM theory all pitch events are defined using one of two level tones - either high (H) or low (L), with H is near the top of a speaker's pitch range, and L near the bottom. There is good evidence to show that speakers control the height of pitch targets: in multiple repetitions, speakers control the height of pitch peaks in a stress language like English as carefully as speakers of tone languages do (Ladd 1996:66).

The choice to analyse intonational pitch events using level tones is another important benefit of using AM theory for cross-linguistic study, since lexical and intonational tones can be analysed in the same way, using combinations of level tones[5]. This is indeed essential for analysis of intonation in 'mixed' languages which have some lexical use of tone (Gussenhoven 2000 example 12). In the following example from Roermond Dutch, the singular form of the word [ɛrm] 'arm' is signalled by the fact that it bears an Accent II extra H tone, alongside the usual declarative H* L% contour.

(2.2)
[Miene ERM^{II} zit aan miene handj vas] *Roermond Dutch*
| | |
%L H* H L%
my arm sits to my hand attached
'My arm is attached to my hand.'

Egyptian Arabic is uncontroversially assumed to be an intonational language, in which tone is used exclusively postlexically. Yet as we shall see, it shares the property of rich pitch accent distribution with lexical pitch accent languages such as Swedish. Within AM theory it is possible to examine these two facts in parallel and assess whether there is any link between them, rather than assuming them to be accidental.

[5] The main alternative view of intonation, known as the 'British school' of intonation, sees pitch contours as pitch 'configurations' (see inter alia O'Connor & Arnold 1961, Halliday 1967, Halliday 1970, for a useful comparison of the two systems see Cruttenden 1997).

2.1.4 Transcription of intonation

AM theory provides for a system for the transcription of intonational contours. Intonational tunes are seen as tonal sequences combining pitch accents and edge tones of various kinds, with all pitch events transcribed using H and L only. A set of standard notational devices used in AM theory is listed in (2.3) (Ladd 1996, Beckman & Elam 1993).

(2.3)	H	high target
	L	low target
	*	pitch accent (associated with the main stressed syllable of some words)
	-	phrase tone (associated with a phrase edge)
	%	boundary tone (associated with a phrase edge)
	!	downstep

Pitch accents may be composed of one target (monotonal) or at most two (bitonal) resulting in the set of possible pitch accents listed in (2.4) below (Face 2002:7). The star notation indicates which of the two tones in a bitonal accent is associated primarily with the stressed syllable[6].

(2.4) H* L* H*+L H+L* L*+H L+H*

As regards, edge tones, in early AM work on English, notation was proposed for two types, showing affiliation to the edge of prosodic phrases at different levels (Beckman & Pierrehumbert 1986): *boundary tones* [H%, L%] align to the edge of a full prosodic phrase called intonational phrase (IP), and *phrase tones* [H-, L-] align to the edge of an intermediate phrase (iP), nested within the larger phrase. Since an IP is composed on one or more iPs, the right edge of an utterance was argued always to bear a sequence of a phrase tone and a boundary tone (the right edge of both iP and IP coincide at the right edge of the utterance)[7].

The remaining symbol '!' is used to denote *downstep* which refers to phonological lowering of the F0 target level of an H tone. In Pierrehumbert (1980) and Beckman & Pierrehumbert (1986), downstep was argued to be triggered phonologically by any bitonal pitch accent, whilst other authors have argued that downstep is better analysed

[6] There is much debate regarding the exact role of the 'starred tone' in a bitonal accent, which is discussed in chapter 7 section 7.1.1.

[7] The need for an intermediate phrase level is disputed in some AM analyses (Grabe et al 1998). Some authors have argued for boundary tones at both the beginning and end (left and right edges) of intonational phrases (Grabe et al 1998, Gussenhoven 2005), or for a 'zero boundary' where the boundary tone maintains the pitch level of the last pitch accent (Grabe et al 1998).

as ‘an independent linguistic choice’ under the control of the speaker (see discussion of this debate in Ladd 1996:89ff.).

An influential AM theory transcription system for intonation is the Tones & Break Indices (ToBI) system, which was developed for General American English (GAE)[8]. English ToBI can be used to describe the intonation of GAE and probably also southern British English. However it cannot be used directly to describe the intonation of other languages since it is the result of a *phonological* analysis of a particular language, rather than a phonetic transcription system. The theoretical choices underpinning ToBI have however been successfully adapted and as a result AM-style phonological analyses exist of many languages, using a similar notation system (see Jun 2005c for a comprehensive survey).

This thesis implements an AM-style transcription system for EA intonation (discussed in detail in chapter 3) in which the tonal sequence is analysed using symbolic labels to represent pitch accents and boundary tones, and the correspondence between the tones and prosodic boundaries of different strengths is motivated in the text.

2.1.5 Metrical structure as the target of association of postlexical tone

As seen in the example in (2.1) above, the surface pitch contour of an utterance is composed of tones which cluster around certain positions in the metrical structure of an utterance, such as the stressed syllable or the edge of the phrase. In AM theory this surface clustering is analysed as evidence of phonological *association* with the heads or edges of constituents at different levels in a hierarchy of prosodic constituents.

As discussed above two levels of intonational phrasing have been proposed for English (Beckman & Pierrehumbert 1986, Beckman & Elam 1993): the intermediate phrase (iP) and the intonational phrase (IP), with the right edge of an iP marked with a phrase accent, either L- or H-[9], and the right edge of an IP is marked with (a phrase accent

---

[8] A full ToBI transcription includes an acoustic waveform (sound record), a fundamental frequency (F0) contour (pitch record) and four tiers of labels (tones, break indices, orthography, miscellaneous). The tones tier contains the symbolic labels representing pitch accents and boundary tones. The break indices tier contains symbolic labels marking the strength of the boundary between adjacent words (0 for the boundary between a word and a clitic, up to 4 for the boundary between two intonational phrases).

[9] In a ToBI analysis of English it is the phrase tone (L- or H-) which spreads inward from the phrase edge to fill the space between the nuclear accent (the last pitch accent) and the phrase end; hence the analysis of the contour in (2.1) above would be H* L-H%.

and) a boundary tone, either L% or H%. In English the iP is also identifiable because it is the domain of downstep, so that at the start of a new iP pitch range is reset.

A more highly articulated hierarchy of prosodic constituents is proposed within the theory of Prosodic Phonology, which seeks primarily to explain parallels and mismatches between the prosodic and syntactic representations of an utterance (Nespor & Vogel 1986, Selkirk 1986, Inkelas & Zec 1995). The key contribution of early work in prosodic phonology was to demonstrate that the suprasegmental phonological representation of utterances is hierarchically organised, rather than consisting of a linear sequence of segments interspersed with boundary markers (as was the case in Chomsky & Halle 1968). The prosodic representation thus motivated is known as the 'Prosodic Hierarchy'.

The hierarchical nature of prosodic representation is due to a notion of prosodic well-formedness which has come to be known as the Strict Layer Hypothesis (SLH) (Selkirk 1984:26): "a category of level *i* in the hierarchy immediately dominates a (sequence of) categories at level *i*-1". The hypothesis is, therefore, that prosodic representation will have the following properties:

(2.5) Well-formedness of prosodic representation according to the SLH

layered $* X^{i}$ | X^{i+1}	no constituent dominates a constituent at a higher level	exhaustive $* X^{i}$ | X^{i-2}	no constituent will be undominated by a constituent at the next higher level
headed $* X^{i}$ |	every constituent dominates a constituent at the next lower level	non-recursive $* X^{i}$ | X^{i}	no constituent will dominate a constituent at the same level

Well-formed prosodic structure is thus hypothesised to be a hierarchy of domains in which every domain at a particular level is composed exclusively of domains at the next level down in the hierarchy[10].

[10] Some aspects of this hypothesis have been challenged on empirical grounds, particularly the claim that prosodic structure is non-recursive (Ladd 1996:237ff.).

There have been a number of proposals as to the exact number of layers of prosodic constituents that are present in the prosodic hierarchy[11]. The set of constituents adopted as a working hypothesis for EA in the present thesis is shown in (2.6), together with notational equivalents used elsewhere in the literature, and the element of syntactic structure from which each is mapped (if applicable) (after Selkirk 2005a).

(2.6) The Prosodic Hierarchy

constituent		*equates to:*[12]	*maps from:*
Utterance	U		
Intonational Phrase	IP		a root sentence or sentence-external clause
Major Phonological Phrase	MaP	phonological phrase/intermediate phrase	a maximal projection (XP)
Minor Phonological Phrase	MiP	accentual phrase	a syntactically branching constituent (two PWds)
Prosodic Word	PWd	phonological word	a morphosyntactic word (lexical)
Foot	Ft		
Syllable	σ		
Mora	μ		

As noted above, a primary goal of Prosodic Phonology has been to determine the nature of the mapping relation between morphosyntactic structure and the prosodic representation. Of the above levels in the hierarchy, four are argued to be 'syntactically grounded' (Selkirk 2005a), in that they can be defined with respect to some aspect of syntactic structure.

The Intonational Phrase (IP) is usually thought to be mapped from one of two types of syntactic element (Nespor & Vogel 1986:188-9): 'root sentences' (defined as any sentence (S) which is not dominated by a node other than S: Emonds 1976), and clauses external to the root sentence, such as parenthetical expressions, non-restrictive relative clauses, tag questions and vocatives[13].

---

[11] Selkirk (Selkirk 1981b) proposed six constituents: syllable, foot, prosodic word, phonological phrase, intonational phrase, and utterance; Nespor & Vogel (1986) motivate an additional constituent level, the clitic group, between the prosodic word and the phonological phrase, which is however not adopted here.
[12] Throughout the thesis, when the work of other authors is quoted I will use the author's *original* notation and terminology, with equivalency to my own notation and terms provided in brackets or in a footnote.
[13] Selkirk (2005a) has recently argued that IP is mapped from a single syntactic category, 'Comma Phrase'(Potts 2003, Potts 2002), which encompasses both root sentences and sentence-external clauses.

A Major Phonological Phrase (MaP) is mapped from the maximal projection of a lexical category (an ‘XP’). This level of constituency, more widely termed the Phonological Phrase, has been the subject of much research, regarding the exact nature of the mapping relation from XP to MaP[14]. Conceptions of the relation that have been observed to be valid cross-linguistically have included sensitivity of phrasing to the relation between a lexical head and its complement (Nespor & Vogel 1986), to the edges of XPs (Selkirk 1986), and to the need to keep the XP within a single MaP (Truckenbrodt 1999, Truckenbrodt 1995). These nuances are discussed in greater detail in section 2.2.2 below, in the context of their instantiation as violable constraints on prosodic structure.

The reason for using the term ‘Major’ Phonological Phrase is in order to accommodate addition of the next constituent level down to the hierarchy: the Minor Phonological Phrase (MiP). The MiP is tonally marked in Japanese, and has been shown to be an independent level of the hierarchy from the MaP (Kubozono 1993, Poser 1984), and equates to the ‘Accentual Phrase’ level in purely tonal analyses (Beckman & Pierrehumbert 1986). It has been shown that MiP maps systematically from a *branching* syntactic phrase, that is, a phrase composed of two Prosodic Words (PWds) (Kubozono 1993, Selkirk et al 2003)[15].

The last of the constituent levels which is mapped from morphosyntactic structure is the Prosodic Word (PWd), which maps from morphosyntactic words in lexical (but not functional) categories (Selkirk 1996). It equates to the ‘Phonological Word’ sometimes used elsewhere in the literature.

Whilst in some cases prosodic constituency and syntactic constituency are the same, this has been shown to be not always the case. Thus prosodic structure is argued to be related systematically though not isomorphically to syntactic structure. Some authors argue that phonological processes cannot refer directly to syntactic structure, but only to the hierarchy of prosodic constituents which acts as the mediating structure between

[14] Selkirk & Kratzer (2004) have recently argued for a re-analysis of the mapping between MaP and syntactic structure, based on the notion of phase/phase edge, working within the Minimalist Framework, after Chomsky (2001).
[15] Whilst this constituent is not adopted by all authors, the fact that it is necessary in Japanese, following the hypothesis of Nespor & Vogel (1986) that all constituent levels of the hierarchy are present in all languages, means that it is included in the hierarchy here as a working hypothesis. Relevance of MiP for analysis of EA is discussed in depth in chapter 5.

syntax and phonology, known as the Indirect Reference Hypothesis (Selkirk 1984, Hayes 1989, Truckenbrodt 1995)[16].

The remaining constituent levels of the hierarchy are defined independently of syntactic structure, but have been shown to act as the domain of phonological processes (Nespor & Vogel 1986). The mora, syllable and foot are in addition well-motivated as prosodic constituents in their own right (McCarthy & Prince 1996, McCarthy & Prince 1990, Hayes 1995). The status of the utterance is less certain, although motivated by Nespor & Vogel on the grounds of semantic cohesion across strings encompassing more than one root sentence (Nespor & Vogel 1986:211ff.)[17].

All of these constituent levels are included in the version of the prosodic hierarchy adopted here, because of the importance of the notion of prosodic *prominence* in this thesis. Prosodic prominence can be formalised by means of the notion of a *Designated Terminal Element*, which is relevant to all levels of the hierarchy, not just those levels which are syntactically grounded.

In a prosodic domain 'X', which is composed of one or more domains 'Y' at the next level below it in the hierarchy, the notion of *relative metrical prominence* determines that there is an asymmetric relationship among the one or more Y-level constituents which are grouped together into X (see interalia: Liberman & Prince 1977, Prince 1983, Selkirk 1984, Halle & Vergnaud 1987, Hayes 1995). This most prominent constituent Y within X is the *DTE* of X, and can also be described as the *head* of X. Within a PWd, for example, the most prominent foot is the DTE or head, and is assigned prominence (marked by stress in English).

Relative prominence is assigned in this way at all levels of the hierarchy. The choice of which constituent at a given level is chosen as head is language-specific, but has been shown at MaP level and above be to be most often either leftmost or rightmost, at the beginning or end of the phrase, respectively (Nespor & Vogel 1986). In a bracketed grid representation (after Halle & Vergnaud 1987)[18], the DTE of an Utterance is therefore the mora which is dominated by the chain of heads up through the levels of the prosodic

---

[16] But see Chen (1990) for discussion of an alternative view.
[17] It may be that sequences of root clauses phrased together are in fact phrased into an enlarged IP, perhaps due to constraints on the size of IPs.
[18] This could also be represented by means of a prosodic tree.

hierarchy, as shown in (2.7) below. Note that in English the head of phrase level constituents is rightmost.

(2.7) The Designated Terminal Element (DTE) (after Hayes 1995:369)
The DTE of the Utterance (▲) is the mora which is dominated by the chain of heads up through the levels of the prosodic hierarchy.

U	(					x	)
IP	(					x	)
MaP	(		x	)	(	x	)
MiP	(		x	)	(	x	)
PWd	(x	)	(x	)	(x)	(x	)
Ft	(x	)	(x	)	(x)	(x	)
σ	(x)	(x)	(x)	(x)	(x)	(x)	(x)
μ	x x	x	x x	x	x x	x x	x
	Bel-	gian	far-	mers	grow	tur-	nips
						▲	

Nespor & Vogel (1986) assumed that all constituent levels of the hierarchy are present in all languages, even if no surface reflex is observed of a particular constituent level in the language. They argue this on theoretical grounds, on the basis that a theory that requires all languages to have a specific set of phonological units is stronger than one that does not. They also argue that since the hierarchy is defined in terms of mapping from other components of the grammar then the absence of any constituent would appear to imply no interface with that part of the grammar in that language, which seems unlikely. For the same reasons, this thesis adopts as a working hypothesis the view that all constituent levels of the hierarchy are present in EA, even if no surface reflex of prosodic constituency at a particular level is observed.

Empirical evidence for prosodic constituency comes from a wide range of phonological and phonetic phenomena[19]. Segmental processes have been shown to be sensitive to prosodic constituents of different sizes, as demonstrated extensively by Nespor & Vogel (1986), as also have tonal processes such as tone sandhi and high tone distribution (see Yip 2002:116ff. for a summary). Rhythmic evidence can give clues to phrasal constituency, such as stress retraction in Italian and English, which applies within but not across prosodic domains at different levels (Frota 2000). Fine-grained phonetic detail has also been shown to vary consistently with levels of the prosodic hierarchy, so that duration of the last constituent in a phrase is lengthened, with different degrees of

[19] Selkirk (1986) queries the reliability of evidence from certain postlexical rules. She sees a distinction between rules that are part of the phonetic implementation, which tend to be gradient in effect, tempo-sensitive and apply in variable domains, and rules which are part of phonology proper.

lengthening at different levels (the higher the level, the greater the lengthening: Wightman et al 1992). Finally, intonational evidence from the position of edge tones and pitch register reset (limiting the application of downstep) has been shown to be sensitive to syntactically-defined prosodic boundaries in many languages including Japanese (Selkirk & Tateishi 1991), Bengali (Hayes & Lahiri 1991) and European Portuguese (Frota 2000).

A pilot study for this thesis explored possible evidence for prosodic constituency in EA and found that the most reliable cues to prosodic boundaries were the distribution of edges tone and local pitch register range reset (Hellmuth 2004). Other authors have noted that some phonological processes are sensitive to prosodic constituency, but it proved impractical to use these as reliable tests: Watson (2002) reports a number of segmental processes such as sibilant assimilation and voicing assimilation which apply within the phonological phrase (=MaP), but investigation of these showed a wide range of speaker variation, and it proved impractical to use them as reliable tests[20].

This thesis adopts the view that there is a single prosodic structure to which all phonological processes including intonation are sensitive[21]. Thus the constituent referred to as iP in purely intonational analyses of English is here assumed to equate to a unit of the syntactically grounded prosodic hierarchy (MaP), and that the IP of both hierarchies are one and the same. There is good evidence to support this 'integrated' view of prosodic structure (Hayes & Lahiri 1991, Selkirk & Tateishi 1991, Frota 2000), but other studies have also found mismatches, which show for example apparent violations of the Strict Layer Hypothesis (Ladd 1996:244ff., Dresher 1994)[22]. Such mismatches are however amenable to analysis, and indeed predicted to occur, within a theory of phonology based on the notion of conflicting constraints on phonological representations. The next section describes such a theory, Optimality Theory, and sets out how the key concepts of AM theory and prosodic phonology are implemented in the form of violable constraints in this thesis.

[20] El Zarka (1997) notes that retraction of secondary stresses is sensitive to prosodic domains, but this was not pursued due to controversy over the status of secondary stress in EA (see Hayes 1995:72).
[21] Other authors have argued that the domain of application of postlexical rules is purely tonally defined, such as Jun (1996) for Korean.
[22] In some cases different types of evidence suggest different phrasing generalisations (Gussenhoven & Rietveld 1992), and Gussenhoven (2004:167) points out that there are conflicting assumptions which render the iP/MaP parallel problematic.

2.2 A theory of ranked violable constraints

AM theory, and this thesis, is concerned with the interaction between phonological tones (pitch accents) and metrical structure (the hierarchy of prosodic constituents). In the thesis this interaction is modelled within Optimality Theory, a theory of ranked violable constraints (Prince & Smolensky 1993, McCarthy & Prince 1995, McCarthy 2002), which is introduced briefly in section 2.2.1 below, with discussion of why the theory is suitable for analysis of EA pitch accent distribution in an AM framework. Sections 2.2.2-2.2.3 go on to examine in greater detail the types of constraints that are known to govern different aspects of prosodic structure.

2.2.1 Modelling the interaction of conflicting constraints on prosodic structure

Optimality Theory is an output-oriented theory and therefore particularly suitable for modelling the properties of tone in a 'unity of pitch phonology' conception of intonation: the theory is well-equipped to characterise how the grammar might treat phonological objects similarly regardless of their origin or function. In addition, treatment of interface conditions on prosodic structure (the mapping from syntax) alongside issues internal to the prosodic representation itself (prosodic well-formedness) is also highly conducive to analysis within OT, which is able to model the interaction of conflicting requirements arising in different parts of the grammar.

In OT the grammar is conceived of as an evaluation metric, of potential output forms against an input form (Prince & Smolensky 1993). The decision as to which candidate output form is selected as optimal is based on an evaluation of all possible candidates against a set of constraints on output representation. Output forms are generated by GEN and under the 'Richness of the Base' hypothesis any output form is a logically possible correspondent for any input form. In practice however the most interesting candidates for analysis are those which differ minimally from the input form (McCarthy 2002).

Constraints are of two different kinds: faithfulness constraints penalise forms which are not faithful to the input form in some respect[23]; markedness constraints penalise forms which display typologically marked properties. The constraint set CON is deemed to be universal, but the relative ranking of constraints is language-specific. Low-ranked constraints may well be violated by surface forms, since it is the forms that satisfy the

[23] Faithfulness constraints are formulated as a correspondence relationship between linguistic forms (McCarthy & Prince 1995).

most highly-ranked constraints that are evaluated as optimal. The effects of a highly-ranked constraint will often be seen more often in the surface linguistic forms of a language, whereas the effects of a low-ranked constraint may well be seen only rarely (if some higher ranked constraint is itself outranked). An output form may violate every constraint in the constraint set but still be selected as optimal if it minimally violates the most highly ranked constraint(s).

Another advantage of using OT in the present thesis is that it is a theory which lends itself to typological comparison. The notion that constraints are universal means that any new constraint proposed in an analysis must by definition be present in the grammar of all languages, though its effects may be rarely seen in a particular language if it is low-ranked. If a subset of constraints can be shown to account for the surface patterns in a particular language, then the prediction of the theory is that every possible permutation of those constraints, in a *factorial typology*, will yield a possible human language (McCarthy & Prince 1993 chapter 6, McCarthy 2002).

Whilst most constraints may be ranked freely with respect to all other constraints, those which encode hierarchical relations have been argued to be in fixed ranking. Two ways of expressing hierarchical relations may result in a fixed ranking (McCarthy 2002): by *harmonic alignment*[24] of natural prominence scales (such as the Sonority Hierarchy), or by exploiting existing *stringency relations*[25] among linguistic forms (such as the stringency relation that holds between heads and non-heads within a prosodic constituent).

2.2.2 Constraints on prosodic structure

There has been much research on the properties of the prosodic representation and on the type of constraints which govern the occurrence and distribution of prosodic constituents. These are set out below in three sections, taking constraints on the prosodic structure itself first, then on its relationship with syntactic structure, and finally on the rhythmic properties of resulting constituents.

[24] Harmonic alignment was defined in Prince & Smolensky (1993:136) as the derivation of pairs constraint hierarchies from the combination of a pair of prominence scales, one of which is binary.
[25] If two constraints A and B stand in a *stringency relation* then the violations of A will always be a proper subset of the violations of B: A "imposes a more stringent test" than B does (McCarthy 2002:20).

2.2.2.1 The Strict Layer Hypothesis: prosodic domination constraints

The SLH discussed in section 2.1.5 above, has been argued to decompose into four 'prosodic domination' constraints (Selkirk 1996), as in (2.11) below.

(2.11)	LAYEREDNESS	No C^i dominates a C^j, where j>i. *eg: no σ dominates a Ft*
	HEADEDNESS	Any C^i must dominate a C^{i-1} (except if C^i - σ) *eg: a PWd must dominate a Ft*
	EXHAUSTIVITY	No C^i immediately dominates a C^j, where j< i-1 *eg: no PWd immediately dominates a σ*
	NONRECURSIVITY	No C^i dominates C^j, where j = i. *eg: no Ft dominates aFt*

These constraints are freely ranked and are violable. Thus the problem of minimal violation of some aspects of the SLH, such as the problem of apparent prosodic recursion argued for by Ladd (1996), are more amenable to analysis. Of the above constraints, Selkirk has suggested that LAYEREDNESS and HEADEDNESS are universally undominated, as there appear to be no counterexamples (Selkirk 1996). In contrast, prosodic structures which minimally violate EXHAUSTIVITY and NON-RECURSIVITY are observed (Truckenbrodt 1995)[26].

Relative prominence within prosodic constituents can be argued to arise due to a constraint requiring the presence of a head at either the right or left edge of each constituent (Truckenbrodt 1995:119):

(2.12) ALIGN (α, R/L, HD_α, R/L):
Align each right edge of a prosodic constituent α with a grid-mark that heads that α[27].

This constraint is adopted here at all levels of the hierarchy, and in the spirit of the hypothesis that constituents are present even if not prosodically marked (Nespor & Vogel 1986), similarly we hypothesise that heads of constituents are present, even if not prosodically marked (which implies that the family of ALIGNα,HD_α constraints are unviolated). As for the *position* of the head in prosodic constituents in EA, by analogy

[26] Examples of minimal violation of EXHAUSTIVITY and NON-RECURSIVITY are discussed in chapter 5 (section 5.4.2) in the context of the mapping from morphosyntactic structure to prosodic structure, at the level of the word (Selkirk 1996).

[27] This is an Alignment constraint which relies on gradient degrees of violation to evaluate potential candidates (McCarthy 2003). In an effort to appeal only to constraints incurring categorical violations, one might argue instead for a pair of constraint families: OBLIGATORYHEADα (requiring every constituent at a particular level to have head), (compare "MetricalHeadedness" in Selkirk 2005a:39), and ENDRULE(α)L/R (requiring heads to be in the leftmost/rightmost position in the constituent).

with other dialects of Arabic which have been shown to have rightmost phrasal stress (Benkirane 1998, Chahal 2001), the working hypothesis is that the head will be rightmost[28] in MiP, MaP, IP and Utterance. Prominence within the PWd and below is discussed in section 2.3.1 below in reference to EA stress assignment[29].

2.2.2.2 The mapping from syntactic structure: interface constraints

Early work in Prosodic Phonology showed that languages varied in the nature of the mapping relation between prosodic phrasing at the level of the phonological phrase (=MaP) to the internal structure of syntactic maximal projections (XPs). Nespor & Vogel (1986) argued that the mapping relation reflected the syntactic *head-complement relation*, and was thus determined in part by the direction of syntactic recursion in a particular language and by cross-linguistic variation in the availability of a re-structuring option. Selkirk (1986) argued that the mapping relation reflected sensitivity of phrasing to the *edges* of syntactic maximal projections (XPs).

(2.13) The mapping of phonological phrases (=MaP) from syntactic phrases (XP)

a.	Right -recursive languages (head initial)			
	$[[X^0$		$[\ Y^0\ ]_{YP}]_{XP}$	
	(		)	Chi Mwi:ni (Selkirk 1986)
	(	)	( )	Ewe (Clements 1978)
b.	Left-recursive languages (head final)			
	$[[Y^0\ ]_{YP}$		$X^0\ ]_{XP}$	
	(		)	Korean (Cho 1990)
	(	)	( )	Japanese (Selkirk & Tateishi 1991)

The languages illustrated in (2.13) above are analysed as follows:

[28] The term rightmost/leftmost here refers to the position of the head in a representation based on phonetic transcription of the word; the phonetic transcription is written from left-to-right (rather than in Arabic orthography, which is written from right-to-left).

[29] There is probably also a sister constraint to ALIGN α,HD_α, which would take the form ALIGN HD_α,α (requiring every instance of a head of a constituent to be dominated by a constituent at that level). Selkirk (2005a) has indeed argued for a constraint HEAD-OF-α (requiring every head of a constituent to be dominated by a *distinct* instance of a constituent at that level). The effects of such a constraint do not appear to be at issue in the EA data here and are therefore not pursued further.

(2.14) Relation-based vs. edge-based mapping from syntactic structure.

	relation-based mapping	*edge-based mapping*
Chi Mwi:ni	a right recursive language, with obligatory restructuring	phonological phrases mapped from the right edge of maximal projections
Ewe	a right-recursive language, which disallows restructuring	phonological phrases mapped from the left edge of maximal projections
Korean	a left-recursive language which allows restructuring	phonological phrases mapped from the left edge of maximal projections
Japanese	a left-recursive language which disallows restructuring	phonological phrases mapped from the right edge of maximal projections

Ghini (1993) argued that the cases which Nespor & Vogel analysed by means of variation in availability of 'restructuring' could be successfully reanalysed within the edge-based view, by introducing sensitivity of phrasing to prosodic weight: the phrasing facts in Italian display a preference for phrases of uniform prosodic weight, and if uneven, for phrases of increasing size through an utterance. Indeed Truckenbrodt (1995:70-72) has shown that there are cases which can only be captured by the edge-based plus prosodic weight approach.

Nonetheless even this combined conceptions of the syntax-phonology interface is still insufficient to account for the full range of phrasing facts observed in some other languages. Hale & Selkirk (1987) showed that phrasing in Tohono O'odham (TO, formerly known as Papago) was sensitive to an additional aspect of the syntactic representation, namely whether or not an XP was itself 'governed by' a lexical head: only the edges of XPs which are not themselves part of a larger XP in TO trigger a phonological phrase (=MaP) boundary.

Within OT these three factors: XP edges, XP integrity and prosodic weight are expressed by means of separate constraints, and cross-linguistic variation is ascribed to language-specific rankings among the constraints. Constraints on prosodic weight are treated in section 2.2.2.3 below, and constraints relating phrasing to XPs are treated here.

Sensitivity of phrasing (at the MaP level) to the left or right edge of an XP is expressed by means of alignment constraints (Selkirk 2000, McCarthy & Prince 1993):

(2.15) ALIGNXP, R: Align (XP,R; MaP,R)
For each XP there is a MaP such that the right edge of XP coincides with the right edge of MaP.

ALIGNXP, L: Align (XP,L; MaP,L):
For each XP there is a MaP such that the left edge of XP coincides with the left edge of MaP.

The relative ranking of these constraint will determine whether phrasing is sensitive to the left or right edge of XPs[30].

The lexical government parameter proposed by Hale & Selkirk (1987) for Tohono O'odham is re-analysed by Truckenbrodt (1999, 1995) as a violable constraint:

(2.16) WRAPXP: Each XP is contained in a phonological phrase (=MaP).

The relative ranking of ALIGNXP and WRAPXP in a language will determine the surface phrasing patterns of a language. Selkirk (2000) notes that the two constraints may fulfil different perceptual roles, with ALIGNXP playing a demarcative function and WRAPXP a cohesive function.

2.2.2.3 Prosodic well-formedness : rhythmic constraints

As noted above, Ghini argued that an edge-based analysis should be augmented with sensitivity to the prosodic weight of resulting constituent phrases.

Prosodic phrasing had also been shown to be sensitive to whether or not a syntactic category is branching or not (that is, whether or not is composed of more than one element). In the example in (2.17) from English, cited in Inkelas & Zec (1995), evidence for prosodic constituency comes from a rule of stress retraction (the 'Rhythm Rule') which applies between two words within a single phonological phrase (=MaP) but not between two words across a phonological phrase boundary.

(2.17) English Rhythm Rule

a) John pérseveres gládly.	vs.	John persevéres gládly and diligently.
b) Rabbits réproduce qúickly.		Rabbits reprodúce véry quickly.

[30] Truckenbrodt (1999:228 fn11) discusses the possibility of sensitivity to both left and right edges of XPs, and the potential relevance of this option to understanding EA phrasing is discussed in chapter 5 (section 5.3.1).

In the relation-based mapping from syntax provision was made for restructuring of phonological phrases if a syntactic complement was branching in certain languages, whilst in the end-based version of the theory sensitivity to branching was not accounted for directly. Zec & Inkelas (1990) propose an alternative mapping from syntax in which branchingness, called *syntactic sisterhood*, is the key feature of syntactic structure which is mapped into prosodic structure.

Selkirk (2000) instead proposed that branchingness effects are the result of constraints on the well-formedness of prosodic structure, for example requiring a constituent to be minimally branching, that is, 'binary'. The interaction between prosodic well-formedness constraints and interface constraints on the mapping between syntactic and prosodic structure results in branchingness effects on surface prosodic phrasing, and can be modelled in OT. For example, Selkirk (2000) proposes an analysis of English in which a branching phrase is preferred. She argues that this is due to phonological well-formedness constraints on the size of MaPs; these constraints favour productions in which as many phrases as possible are binary. Thus in (2.18), (2.18a) and (2.18b) are preferred to (2.18c) (which is only possible in non-neutral contexts).

(2.18) a. (She lóaned her róllerblades to Róbin)$_{\mathrm{MaP}}$
b. (She lóaned her róllerblades)$_{\mathrm{MaP}}$ (to Róbin)$_{\mathrm{MaP}}$
c. *(She lóaned)$_{\mathrm{MaP}}$ (her róllerblades)$_{\mathrm{MaP}}$ (to Róbin)$_{\mathrm{MaP}}$ *in neutral context*

The well-formedness constraint is formulated as a requirement that each constituent be composed of two constituents at the next layer down in the hierarchy, and Selkirk implements it at the MaP level[31]:

(2.19) BINMAP
A Major Phrase consists of two Minor Phrases.

Similar accounts of branchingess effects have been proposed for a variety of languages including Japanese (Selkirk & Tateishi 1988, Selkirk et al 2004), and a number of Romance languages (Elordieta et al 2003, Prieto 2005b, Sandalo & Truckenbrodt 2002). The relevance of branchingness effects to an understanding of EA pitch accent distribution is discussed in detail in Chapter 5.

[31] Selkirk's (2000) diagnostic for MiP phrase status in English is the presence of a pitch accent (which she indicates with an acute accent).

2.2.3 Bringing tone and prosodic structure together

As outlined in section 2.1 above, the fundamental claim of Autosegmental-metrical (AM) theory is that the surface pitch contour of a language reflects the phonological association of pitch events (tones) to positions in prosodic structure (heads and edges of prosodic constituents). If such association is indeed *phonological*, and thus part of the phonological component of the grammar, then in OT it should find expression in violable constraints.

A common technique in AM theory has been to model the intonational system of a language as a finite-state system, with choices made from among a fixed inventory of possible pitch accents and edge tones (Pierrehumbert 1980). Much of the work carried out in AM theory has therefore been to establish on the basis of empirical evidence what the correct inventory of possible pitch events is, in a particular language (see the papers in Jun 2005c for numerous examples). Whilst there has also been discussion of some variation in which phrases occur in different languages, it appears that the nature of the association relation between tones and prosodic structure has been assumed to be the same in all languages.

It is possible that the phonological grammar also contains a specific mechanism, or in OT, a constraint, to introduce an association relation between tones and prosodic structure. Some aspects of cross-linguistic prosodic variation may therefore be due to the ranking of such a constraint relative to other constraints on prosodic structure.

This thesis seeks to account for the distribution of EA pitch accents within a particular theory of the relation between tone and positions of metrical prominence (referred to as tone↔prominence theory, T↔P theory) and which seeks to account for the facts of EA pitch accent distribution in an analysis that is also sympathetic to the notion of the 'unity of pitch phonology'. These issues are discussed in detail in chapter 6. Note that within the scope of this thesis it is not possible to test the predictions of the analysis against the full range of potential cross-linguistic prosodic variation. Rather, the thesis seeks to test the predictions of T↔P theory in-depth against the empirical facts of one language: EA.

The next section reviews what is known already about the prosody of EA and of other spoken dialects of Arabic.

2.3 The prosody of Egyptian Arabic

Egyptian Arabic (EA) is the dialect of Arabic spoken in Cairo, Egypt, and by educated middle class Egyptians throughout Egypt.

The metrical phonology of EA is well-described, and has been the subject of much research (see Watson 2002 for a comprehensive summary and review of prior work). A description of the facts of EA word stress assignment is provided in section 2.3.1, together with brief discussion of other metrical properties such as rhythm and clash resolution.

The prosody of EA above the level of the word has received comparatively much less attention, and this situation is paralleled across most spoken dialects of Arabic. Sections 2.3.2 and 2.3.3 give a brief outline of previous work on the intonation of other spoken dialects and of EA itself, respectively. Section 2.3.4 describes a series of pilot studies which were formative in defining the research question addressed in this thesis.

2.3.1 The metrical phonology of EA

This section provides background information about the metrical phonology of EA, looking at the assignment of word-stress, rhythmic properties and strategies employed in EA for avoiding stress clash.

2.3.1.1 The metrical phonology of EA: word stress assignment

The dialect of Arabic spoken in Cairo assigns primary word-stress as shown in (2.20) below (data and generalisations from Langendoen 1968).

(2.20) Stress assignment in Egyptian Arabic (EA)
Stressed syllable indicated in **bold** type.

a.	a final 'superheavy' syllable (CVCC or CVVC) is stressed	dar**a**bt ?a9m**aa**l
b.	a penultimate heavy syllable (CVV or CVC) is stressed	kat**a**bta kit**aa**ba
c.	if both final and antepenult are light syllables then either the penult or the antepenult is stressed, whichever is an even number of syllables from the first heavy syllable in the word or, if there are no heavy syllables in the word, from the beginning of the word.	makt**a**ba k**a**taba
d.	In disyllables the penult is stressed unless the final syllable is 'superheavy'.	r**a**?aa q**aa**lat

The generalisation made in (2.20c) is seen most clearly in the EA pronunciation of Classical Arabic words, which have long sequences of light syllables (Langendoen 1968:102):

(2.21)	?adwiyat**u**hu	'his medicine'	šajar**a**tun	'tree'	(stress on the penult)
	?ink**a**sara	'it broke'	šajar**a**tuhu	'his tree'	(stress on the antepenult)

The colloquial EA version of these words lose word-final case and verbal markers resulting in shorter sequences of light syllables; they may also undergo vowel syncope producing 'new' heavy syllables, which attract stress:

(2.22)	?adw**i**yatu	'his medicine'	š**a**gara	'tree'
	?ink**a**sar	'it broke'	šag**a**rtu	'his tree'

However, as McCarthy (1979:446) points out, even in EA colloquial pronunciation the operation of (2.20c) above means that a heavy antepenult does not automatically attract stress, as it would in other dialects of Arabic:

(2.23) madr**a**sa 'school' EA m**a**drasa 'school' Palestinian Arabic[32]

The key generalisations in EA stress therefore are: attraction of stress to heavy syllables word-medially but not word-finally; attraction of stress to word-final superheavy syllables; and, the characteristic EA displacement of stress from a heavy antepenult when it is followed by two light syllables.

Hayes (1995) has argued that cross-linguistically the key variables underpinning stress assignment are: metrical foot type (a left headed *trochee* which is syllable- or mora-counting, or a right-headed *iamb*), direction of foot construction (from the left or right edge of the word), and word-level prominence (an *End Rule* assigning prominence to the head of the leftmost or rightmost foot). In combination with other optional restrictions regarding extrametricality or construction of 'degenerate' feet, Hayes suggests that this Metrical Stress Theory can capture the basics of stress assignment in any language. Languages may also vary slightly in the domain of stress assignment, between the morphosyntactic word or the Prosodic Word, but is has been shown that

[32] Abu Salim (1983:96).

stress in EA assigned within the Prosodic Word, which includes all affixes (Watson 2002).

Working within Metrical Stress Theory, Hayes (ibid. pp67-71) analyses EA stress by means of the *moraic trochee* as its basic foot template. A moraic trochee consists of either two light syllables of which the first is strong, or a single heavy syllable. For EA Hayes invokes consonant extrametricality[33], and states that foot construction is left-to-right in moraic trochees; word-layer construction follows 'End Rule Right' resulting in a right-headed word-layer constituent. He notes further that EA has a total ban on degenerate feet, which means that a single light syllable 'left over' at the right edge of the word after foot-construction has taken place is not assigned foot status alone.

```
(2.24)   (        x         )
         (x  )(  x    . )
         ? i n  k a s a r a                    Hayes (1995:70)
```

There is some debate in the literature as to whether there is any secondary stress in EA. If the foot is taken as the conditioning context of vowel syncope (a high vowel is deleted in the weak syllable of a foot) then there is evidence for the presence of feet to the left of the main stress foot in EA (Kenstowicz 1980). However it is not clear whether such stresses are systematically marked with correlates such as duration, amplitude or pitch, and Hayes (1995:71) notes that a foot-based analysis of secondary stress would fail to account for the distribution of secondary stresses reported by Weldon (1980)[34].

2.3.1.2 Rhythmic properties of EA

Heliel (1977) is a study of the rhythmic properties of EA, with the aim of ascertaining whether or not EA is a syllable- or stress-timed language (in the definition of Roach 1982). On the basis of an instrumental study of read narratives Heliel concludes that EA is stress-timed because the distance between stressed syllables tends to isochrony whereas the distance between syllables does not[35].

---

[33] He also states that *if* final long vowels are pronounced long in EA then mora extrametricality is also required to account for their not attracting stress; in fact final long vowels are only pronounced long in EA in a few exceptional cases.

[34] El Zarka (1997) also reports secondary stress in EA.

[35] Heliel (1977) also provides a brief but thorough description of EA intonation, in order to explore potential effects of intonation on rhythm in EA, which will be referred to at intervals throughout the thesis.

Mitchell (1969:156) describes EA as stress-timed but in a ‘quantitative’ fashion, by which he means that although EA is stress-timed, it does not fully reduce all unstressed vowels, as occurs in English: “non-prominent syllables in this variety of Arabic are not subject to ‘reduction’ in the manner of their English counterparts and must be given their due rhythmic weight. In this respect, Egyptian Arabic differs from predominantly ‘syllable-timed’ Moroccan varieties.. and may be classified as of a third type neither stress- nor syllable-timed, for which the label quantitative might be appropriate”

The search for a phonetic correlate of the distinction between stress-timing and syllable-timing has been fraught (Roach 1982); however, recent studies have proposed a quantitative rhythmic measure based on the relative duration of vowels and consonants within an utterance (Grabe & Low 2004, Ramus 2002). A recent study using this technique compared six dialects of Arabic (Ghazali et al 2002), and found them all to fall in the same range of quantitative scores as those observed in languages known to be stress-timed such as English. However there was considerable variation among the different dialects, with EA falling in the middle of the continuum of variation.

2.3.1.3 Stress clash resolution strategies in EA

A clash is defined as adjacent prominences at the same level of the metrical grid (Nespor & Vogel 1989)[36]. A range of clash resolution strategies are observed cross-linguistically, with effects on both F0 alignment and duration. The most common effects are stress-shift (accent-shift), stress-deletion (accent-deletion) and ‘beat-insertion’.

Languages have been observed to employ different strategies at different levels of the prosodic hierarchy, to resolve clashes between adjacent prominences. For example, in English the most common strategy used to resolve a clash between adjacent words (W) within a phonological phrase is stress-shift (accent-shift) in the first word: that is, in W1 in a sequence: |[W1 W2]|; a clash between words *across* a phrase boundary |W1][W2| is resolved by means of beat-insertion after W1. Italian uses the same strategy across a phrase boundary, but prefers stress-deletion (accent-deletion) in W1 as a means of resolving a phrase-internal clash (Frota 2000). There is also evidence to suggest that languages may also vary in the degree of clash which can be tolerated; in Italian only a

[36] There is no clear consensus in the literature regarding the nature of clash, with authors roughly divided between a stress-based analysis, in terms of adjacent metrical prominences, and an accentual analysis in terms of adjacent pitch accents. Since in the present study these two factors reduce to the same thing in EA (if as claimed a pitch accent marks every word-level metrical prominence) these differences are not explored further here (a summary is found in Frota 2000 chapter 3).

clash between strictly adjacent stresses (with 0 intervening syllables) will trigger clash resolution, whereas in English a clash of either 0 or 1 intervening syllables between stresses will trigger stress-shift (Nespor & Vogel 1989).

The examples in (2.25) below illustrate that languages have different strategies for resolving the clash: in English the problem is solved by moving the stress leftwards in W1, called *stress shift*; Italian removes the offending stress altogether, called *beat deletion*.

(2.25) Clash resolution in English and Italian (Frota 2000:116)

a.	3		*	*		*		*	
	2	*	*	*		*	*	*	
	1	*	*	*		*	*	*	
		thir	*teen*	*men*	→	*thir*	*teen*	*men*	

b.		*	*			*		
	2	* *	*		* *	*		
	1	* *	*		* *	*		
		sara	*fatto*	→	*sara*	*fatto*	‘(it) will be done’	

There is no existing literature available on clash resolution strategies in EA. Mitchell (1962:28) notes that there are occasions when a clash of 1 intervening syllable may trigger what appears to be stress-deletion (accent-deletion):

> “standing alone, both *kit**aa**b* ‘book’ and *far**i:**d* ‘Farid’ (proper name) have their prominent syllable, but in *kitab far**i:**d* ‘Farid’s book’ it is possible for the prominent syllable of the second word only to stand out; no long vowel appears in a non-prominent syllable .. hence *kitab*”.

In a pilot study for this thesis (Hellmuth 2005), a strictly adjacent clash context inadvertently formed part of the experimental materials, and a survey of subjects’ production strategies in dealing with this clash maybe of interest. The noun phrase [hi'laal '?azra?] *crescent blue* ‘blue crescent’, was recorded four times each by six EA speakers (in different information structure contexts). In all 24 tokens both words were accented, matching the findings of the present study. In 10 of the 24, the alignment of the pitch accent in the second word ['?azra?] appears to be unchanged and the clash was therefore tolerated. In the remaining 14 cases however, visual inspection of the pitch track, as well as the auditory impression, suggests that the accent has been shifted rightwards in W2 [?az'ra?], to some extent. Of these latter 14 cases, 3 also include a degree of lengthening of the final syllable of the first word [hi'laal]. There were no effects of clash on the alignment or position of stress in the first word.

This evidence is limited but suggests that whilst stress-deletion is avoided in EA, stress-shift in W2 and beat-insertion after W1 may be available as clash resolution strategies. The remarks in Mitchell (1962) mentioned above suggest that not only strictly adjacent stresses but also a clash of 1 intervening syllable could count as a clash in EA. The relevance of clash to EA pitch accent distribution is discussed in chapters 6 and 7.

2.3.2 The intonation of spoken Arabic dialects

A chronological overview of the study of Arabic intonation charts a progression from relative lack of interest in the phonology of the language ‘above the level of the word’, through descriptive studies and studies working within the British school of intonation, to early instrumental studies and the first studies in AM theory.

The segmental phonology of classical Arabic was described with great accuracy by the traditional Arabic grammarians[37], notably Sibawayhi (1990, cf. Al-Nassir 1993). Stress and intonation were not described in similar depth however (Suleiman 1999), other than in the context of poetry (Alhawary 2003). An exception to this is Ibn Jinni (1986:370-3), who notes how suprasegmental factors such as loudness and duration can be used to generate the desired pragmatic interpretation:

> “You may praise a person and say ‘kaan wallahi rajulan’ ([*was by-God a-man*] ‘he was indeed a man’). Here you emphasise the word ‘allaah’ [*God*] by lengthening the ‘l’ and making it louder. This is as if you are saying he was a virtuous, brave, generous person, or the like… You may also say ‘sa?alnaahu fawajadnaahu insaanun’ ([*we-asked-him and-we-found-him a-person*] ‘we have asked and found him a person’). You slow your speed and make the sound louder on the word ‘?insaan’ [*person*]. This is as good as saying that he is kind, open-handed or the like…”.

However the prosodic properties of ordinary, non-emphatic, speech seems to have gone largely undescribed[38].

---

[37] The traditional Arab grammarians were writing in the 8th-10th centuries CE (1st-3rd centuries AH); see Suleiman (1999) for a critical summary.

[38] Alhawary (2003) suggests that this may be due in part to the fact that word-stress assignment in classical Arabic is fully predictable.

A number of descriptive studies were made of Arabic dialects within the British school of intonation (following O'Connor & Arnold 1961), including Iraqi Arabic (Ghalib 1977) and Riyadhi (Saudi) Arabic (Badawi 1965). In a pedagogically-oriented study, Mitchell (1993) appears to be the first author to offer a comparative overview of intonation in ordinary speech in a variety of Arabic dialects, taking in Egyptian, Jordanian, Libyan, Lebanese, Syrian and Moroccan dialects.

AlHarbi (1991) is a study of Kuwaiti Arabic working within a functional view of intonation (Quirk et al 1964, Crystal 1969). He offers an account of the mapping between syntactic & semantic structure (‘clauses’) and intonationally defined prosodic phrases (‘intonation groups’). He found that phrasing was quite variable, and was affected by ‘a very large number of situational factors’ but nonetheless that: “speakers paragraph their flow of speech, by means of intonation, at grammatically relevant points” (Al-Harbi 1991:179). This was probably the first study to explore, in an Arabic dialect, the mapping between intonationally defined prosodic phrasing and syntactic structure.

Early instrumental studies on the intonation of Arabic dialects include Al-Ani (1970) on Iraqi Arabic and Al-Rammuny (1989) on Jordanian Arabic, and there are also a number of excellent descriptive accounts based on instrumental data, such as Ingham (1994) for Najdi Arabic, and Al-Khalifa (1984) for Kuwaiti Bedouin Arabic. The first autosegmental-metrical (AM) study of an Arabic dialect was Chahal (2001) for Lebanese Arabic.

2.3.3 Prior work on the intonation of EA

Turning to work on the intonation of Egyptian Arabic in particular, these follow the same pattern as that observed for other dialects, with descriptive, ‘functional’, instrumental and AM studies all represented.

Harrell (1957:17ff.) describes EA intonation firstly in terms of the relation of loudness (amplitude) to phrasing. His term microsegment seems to be the equivalent of the Prosodic Word, in the terms employed in the present thesis. He notes that prominence within a phrase is not always marked in EA: “in a phrase of more than one microsegment, the stresses in the various microsegments are frequently approximately equal in loudness... there are also phrases which have no phrase stress (i.e. the stresses of the various microsegments are of approximate equal loudness) despite wide internal pitch intervals". His description of intonation focuses on global intonation contours (across whole phrases). He observes that there are two main patterns – 'sustaining' and 'descending', with two types of descending pattern (single and double peak). The shape of global intonation contours across whole sentences in EA is discussed in detail in chapter 3.

The pedagogical study mentioned above, Mitchell (1993), gives a very detailed description of some aspects of EA intonation, down to small details of the shape and alignment of pitch movements to stressed syllables. Similarly, the reference grammar of Gary & GamalEldin (1981) explores some aspects of the interaction of intonation with syntactic and semantic structure in a systematic way. Both of these studies are referred to throughout the thesis in order to provide background information relevant to each particular chapter.

The first instrumental study of EA intonation was Abdalla (1960) who compared absolute F0 values in Hertz of successive pitch peaks in a phrase, and proposed an inventory of phonemic pitch contours, observed in various different utterances types, including declaratives, imperatives and questions. More recent instrumental studies of EA intonation include Norlin (1989), who compared F0 contours in neutral and non-neutral utterances, Rifaat (1991), who proposes an algorithm for modelling the global F0 contours of the EA pronunciation of Modern Standard Arabic (MSA), and Ibrahim et al (2001) who determine the pitch range and register properties of global pitch contours in EA declaratives and questions. Reference will be made to all of these studies in the thesis as a comparison to the contours and pitch range properties observed in the thesis corpus.

Finally there have been two studies of EA intonation within the AM theory of intonation. Both analyse the EA pronunciation of Modern Standard Arabic (MSA), from radio broadcasts, in the case of Rifaat (2004), and from a corpus of recorded speech in the case of El Zarka (1997). The findings of these studies are described in detail in chapter 3 in the context of the model of intonation proposed in the thesis.

At this stage the most relevant point to note about the literature on EA intonation set out above is that, whilst some of these studies do make reference to the fact that there is

often a pitch movement on every word in a phrase in Arabic (cf. Harrell above), none have noted that this is perhaps an unusual property, nor have sought to offer an explanation for it. This thesis therefore seeks to contribute to our growing understanding of EA intonation in two ways: firstly, by establishing what the facts of EA pitch accent distribution are, and, secondly, by means of analysis of these facts in a framework amenable to cross-linguistic comparison (the AM theory), to offer an explanation for this typologically unusual property.
The next section briefly describes the rationale and results of a series of pilot studies which hinted at the potential theoretical gains to be made from a direct study of pitch accent distribution in EA.

2.3.4 Background to the research question of the thesis

Three pilot studies were carried out in order to begin to establish the properties of EA intonation and identify areas of fruitful research. Each reproduced an existing study carried out on another language.

The first pilot study collected focus related data, with two speakers of EA (Hellmuth 2002a, Hellmuth 2002b). The speakers read scripted sentences containing repeated words which would be thus be 'given' in context, as well as items of low semantic weight such as indefinite pronouns, in an attempt to elicit words in contexts conducive to 'de-accentuation'. The sentences were direct translations of parallel data reported for English and other languages in Ladd (1996 chapter 5). The speakers were also asked to respond using a scripted target sentence to a series of wh-questions designed to elicit focus on different parts of the target sentence (modelled on Chahal 2001). The main outcome of this first study was to note that the speakers failed to deaccent given words or words of low semantic weight, and that, in a sentence with focus on the initial word, words after the focus were not de-accented. The second pilot study also elicited focus-related data, but using a game scenario in order to elicit semi-spontaneous speech with varying information structure (Hellmuth 2005). The methodology was closely modelled on that of Swerts et al (2002), and again, this second study found no evidence of de-accenting.

The third study was designed in order to establish what types of phenomena might mark prosodic phrase boundaries in EA (Hellmuth 2004). A set of specially constructed SVO sentences was designed, in which the length (number of words) and syntactic

complexity of both subject and object was systematically varied. The aim was to create long sentences which would be conducive to insertion of a sentence-internal phrase boundary. The methodology was closely modelled on that of Elordieta et al (2003), but in addition contexts for a segmental sandhi rule of epenthesis were placed across potential phrase boundary positions. This study found that very few phrase boundaries were inserted and those that occurred were marked tonally. In addition, it was notable that even in these long utterances every word in the sentence bore a pitch accent.
As a result of these various small-scale pilot studies, it seemed that the distribution of pitch accents in EA was consistently dense, and resistant to de-accenting even in non-neutral contexts and long utterances. A systematic study of EA pitch accent distribution and of its interaction with other areas of EA grammar (such as information structure) was therefore designed and implemented.

The next section describes some of the issues that had to be addressed in collecting speech data in a spoken dialect of Arabic.

2.4 Issues in data collection

Diglossia is used to denote 'community bilingualism', where all of the members of a speech community have full command of two dialects or languages, which are usually kept distinct by function (Romaine 2002). This unusual language situation is the norm in the Arab world, and in Egypt means that all speakers acquire and use the spoken dialectal variety (which has no standardised written form), at home and in informal contexts; they learn to speak, read and write the standard form, MSA, at school, which is then used throughout life in formal settings (Holes 1995).

Although use of written Arabic in experimental materials carries with it the risk of eliciting not the dialectal form but a more standardised register of speech, the thesis uses read speech (from written prompts) in the two experimental investigations. In the alignment investigation (described in chapter 7) written prompts were seen as the best way to get a large number of parallel tokens, and the facts of alignment in 'lab' speech have been shown to be consistent with those observed in semi-spontaneous speech (Lickley et al 2006). In the focus investigation (described in chapter 8) written prompts were used because the subtle nuances of information structure were very difficult to elicit by other means (see discussion of this problem in chapter 8 section 8.2.2).

In an effort therefore to keep register interference to a minimum in the present study, EA lexical items and spelling conventions were used as much as possible in the written prompts in order to elicit colloquial productions. Although the written form of EA is not standardised, Egyptians use a colloquial written form in informal documents such as personal letters, and are used to reading it in informal publications such as cartoons. Prompts written using colloquial orthographical conventions have also been shown to be a reliable method for eliciting read speech in Moroccan Arabic (Siemund et al 2002).

Some examples of lexical items exclusive to EA, and of spellings which are incorrect in MSA but correct in EA, are listed below in (2.26) and (2.27). The written prompts used for the thesis were checked for authenticity by an Egyptian teacher of EA before use.

(2.26) Examples of EA lexical items used in the datasets to elicit colloquial register

EA	*MSA*	*gloss*
[mobayl]	--	mobile telephone
[nounou]	[Tifl]	baby
[SaaG]	---	piastres
[biyu?9ud]	[yijlis]	he sits/he is sitting
[ruHna]	[ðahabna]	we-went
[bitaa9]	[Haqq]	belonging to
[?awwi]	[jiddan]	very
[?illi]	[?allaði]	which/that (relative clause marker)

(2.27) Examples of EA spelling conventions used to elicit colloquial register[39]

EA spelling	*transliteration of EA*	*MSA spelling*	*transliteration of MSA*	*gloss*
قولتله	"?uultiluh"	قلتله	"?ultiluh"	I-told-him
يوناني	"yuunaani"	يناني	"yunaani"	Greek
شوفنا	"šuufna"	شفنا	"šufna"	we-saw
ليه	"liih"	له	"lihi"	to-him
صغيّر	"SuGayyar"	صغير	"SaGiir"	small

A total of fifteen speakers participated in the recordings; of these there was just one speaker who attempted at first to read the sentences in an MSA-like register. In this instance the recording was halted and he was requested to produce the sentences in EA register, which he was then able to do without difficulty. All of the other speakers produced EA register without needing to be prompted. The informal content of many of the target sentences and paragraphs (particularly of filler paragraphs) made for natural,

[39] The most common difference between EA and MSA spelling is in vowels; only long vowels are written in MSA, whereas even a short vowel is often written in EA (using the MSA grapheme for a long vowel).

idiomatic productions in most cases. There was much discussion amongst speakers as to the correct EA spelling of certain words, which confirms that although EA speakers are able to read in EA spelling, the EA written form is by no means yet fully conventionalised.

Another potential source of interference would be from a second language that speakers may know, such as English. In order to reduce potential interference from second languages to a minimum speakers were recruited at a private English school, among students in classes at pre-intermediate level or lower in English; none had any proficiency in any other language besides Arabic. Fifteen speakers participated in total, nine male and six female. All were mother tongue speakers of EA, born and raised in Cairo, aged between 21-34 years, and none had any auditory or speech production difficulties. Recordings were made on the school premises and the speakers were paid a small fee for their participation.

2.5 Summary and conclusion

This thesis adds EA to the range of languages for which prosodic theory must account by increasing our knowledge of EA sentence phonology. Specifically, the thesis presents distributional and experimental evidence to support classification of EA as a stress-accent language in which, however, pitch accent distribution is sufficiently different from that reported in other stress accent languages as to require explanation. A new typological category is required to describe EA, and to explain EA, a more finely grained articulation of the grammatical relationship between phonological tone and prosodic structure is proposed.

The first step in this task is to demonstrate the distribution of intonational pitch accents in EA and is undertaken in chapter 3.

3 Pitch accent distribution in Egyptian Arabic

3.0 Outline and aims

The purpose of this chapter is to set out distributional evidence in support of the generalisation that in EA there is a pitch movement (pitch accent) associated with every Prosodic Word.

Section 3.1 reviews what is known already from the literature regarding the intonation of EA in general, and specifically what is known about the distribution of pitch accents.

Section 3.2 briefly sets out the methodology used in a corpus survey of EA speech recordings. The contents and source of the corpus materials is described, as also the notational system used during auditory transcription, and the assumptions on which this notational system is based.

Section 3.3 presents the results of the survey in which it is found that in EA a pitch accent is observed on (almost) every content word. This is true in a variety of contexts which in other languages would be conducive to pitch accent 'deletion', such as in long rhythmic utterances and in non-neutral contexts. The results also suggest that the generalisation holds of semi-spontaneous and spontaneous speech.

Section 3.4 discusses the fact that in the overwhelming majority of cases it is the same type of pitch accent that is observed on every PWd in EA: a rising pitch movement. An inventory is drawn up of the pitch accents and boundary/phrase tones observed during transcription, and is compared to alternative models of EA intonation in the literature, proposed for EA productions of Modern Standard Arabic (El Zarka 1997, Rifaat 2004).

The co-occurrence of rich pitch accent distribution with high frequency of one pitch accent type has been noted in other languages such as Spanish & Greek (Jun 2005b) and in section 3.5 it is argued that 'accent every word' languages are a typologically valid category within the range of cross-linguistic intonational variation.

3.1 Background: what is known about EA pitch accent distribution

As outlined in section 2.3.3, a range of prior work exists on EA intonation, carried out for different purposes and based on varying theoretical assumptions. This section presents a synthesis of the evidence that can be gained from these various sources about EA pitch accent distribution. In most cases the facts about pitch accent distribution are reported in a discussion of some other facet of EA prosody, ranging from 'phrasal stress' to nucleus placement and marking of emphasis, and are here re-interpreted in terms of AM theory.

Whilst no prior study specifically aimed to document pitch accent distribution, studies of EA made for other purposes include revealing comments to the effect that EA has "a tendency to accent all words" (Mitchell 1993:230) and that "in the unmarked case the lexical stress of each word will in continuous speech be stressed" (Heliel 1977:125). Similarly, in her AM-framework instrumental study of the EA pronunciation of Modern Standard Arabic (MSA) El Zarka (1997:356) cites Rifaat's (1991:175) description of the same dialect, and confirms that the same generalisation holds in her own corpus data; namely that in neutral declaratives "every phonological word .. receives a pitch accent"[40].

A number of descriptions suggest that the prominence of all but one word in a phrase is 'demoted' but that the relation between pitch and word-level stress remains "largely predictable"(Gary & Gamal-Eldin 1981:125). In his AM-framework instrumental study of the EA pronunciation of MSA, Rifaat observes that whilst phrase-medial words are 'de-stressed' by one of his speakers, nonetheless "stressed syllables are always associated with higher F0 than unstressed syllables" and that "all pre-final stressed syllables are associated with a large F0 movement.. an indicator of word stress" (Rifaat 1991). Similarly, Abdalla (1960), in an early instrumental study on EA, suggests that degrees of phrasal stress are discernable, but that "stress, quantity and fundamental frequency.. function together" (Abdalla 1960:19).

These comments create the overall impression that, whilst a percept of phrasal prominence is possible in EA, the stressed syllables of phrasally *non-prominent* words are also tonally marked. EA phrasing is discussed in greater detail in chapter 5 but the

[40] My translation of: "Jedes phonologische Wort (und fakultativ auch Funktionswörter) erhalten einen Akzentton" (El Zarka 1997:356)". Variable accentuation of function words is treated in detail in chapter 5.

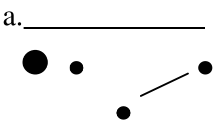
initial working hypothesis is that phrasal constituency can be determined independently from effects on pitch register, with declination 'reset' at phrase boundaries[41].

Another source of evidence about EA pitch accent distribution comes from patterns of EA prosody in non-neutral contexts. Mitchell (1993) notes that the 'nucleus', in the sense of the main or focal prominence of the utterance (as defined in the British school of intonation, e.g. O'Connor & Arnold 1961), can be located in different places in the sentence to alter the meaning without changing the word order. Mitchell cites the following sentence, in which it is just as possible to locate the nucleus on [?itneen] 'two' or [gineeh] 'pounds', as on [maSri] 'Egyptian' (Mitchell 1993:230):

(3.1) ?itnéen ginéeh máSri
two pounds Egyptian
'Two Egyptian pounds."

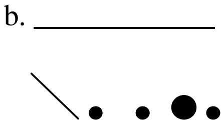
In examples where movement of the nucleus is possible, Mitchell does not specify whether or not following material is de-accented, as it would be in English. In another example that Mitchell gives, reproduced in (3.2) below, Mitchell's O'Connor & Arnold-style 'tadpole' notation, indicates that in (3.2b) there is lexical stress on the medial stressed syllable of the post-nuclear time adverbial [delwa?ti] 'now' (the syllable is marked with an enlarged dot), but not a pitch movement; in contrast there is a pitch movement on the same syllable in (3.2a) (the syllable is marked with a sloping line) (Mitchell 1993:224-5):

(3.2) a.



naazil dilwa?ti
'Are you (ms) coming down now?'

b.



naazil dilwa?ti
'I'm/He's coming down now.'

This notation appears to suggest that de-accenting does take place; however, it may be that the notation designed for English is not flexible enough for transcription of the facts of EA. The conventions of 'tadpole' notation were conceived for English and do not provide for pitch movements other than in nuclear or pre-nuclear position. To mark a pitch movement in post-nuclear position you would have to use either another nucleus-type figure or figures designed to denote pre-nuclear pitch movements (e.g. 'stepping

[41] For evidence of pitch register reset as a reflex of prosodic phrasing in EA see Hellmuth (2004).

head'). The notation could be adapted in this way, but would result in something that is predicted to be (and is) unnatural in English.

In a pilot study undertaken for this thesis (Hellmuth 2002a), an SVO sentence ([muna Hamet naala min liina] 'Muna protected Nala from Lina"), was elicited under varying focus conditions with two speakers. Wh-questions were used to elicit either broad focus or narrow focus on any one of the three sentential arguments (all proper names), but without changing word order. The experimental methodology was modelled closely on work by Chahal (2001) on Lebanese Arabic. Chahal reports a variety of strategies used by her speakers to express narrow focus, which include early placement of nuclear main prominence, and de-accenting of items following the nucleus (Chahal 2001:171ff.). In contrast the pilot study indicated a very different picture for EA. Although the results of one speaker had to be discarded due to a large number of unnatural renditions, auditory analysis of the tokens of the remaining speaker showed no post-focal de-accenting whatsoever: all of the words in all of the sentences bore pitch accents, regardless of focal context (Hellmuth 2002a).

A further pilot study used a game scenario to elicit short semi-spontaneous utterances in controlled focus contexts, and again found no post-focal de-accenting in EA (Hellmuth 2005)[42]. These studies suggest that even in non-neutral contexts words tend to bear pitch accents in EA. At first glance this contrasts with a statement made in a recent study by Rifaat of colloquial EA (Rifaat 2005), who suggests that there is greater deaccentuation in EA than in the EA pronunciation of MSA (cf. his earlier studies (Rifaat 1991, Rifaat 2004)). However Rifaat represents the pitch movements on 'de-accented' EA words by means of a L* type accent, which is consistent with the impression that, even when 'de-accented', all words in EA bear a pitch accent.

To confirm these impressions from the literature, a qualitative corpus survey was carried out using auditory transcription to document the distribution of pitch accents in EA, across a variety of contexts and speaking styles. The methodology used during the survey is described in section 3.2 below, and the results are reported in section 3.3. As a by-product of the survey, the transcriptions were then used to elaborate a preliminary AM model of EA intonation, which is presented and discussed in section 3.4.

[42] The prosodic reflexes of information structure and focus in EA are investigated in detail in chapter 8, which explores the claim that gradient pitch range manipulation marks focus in EA (Norlin 1989).

3.2 Methodology: corpus survey

The purpose of the corpus survey is to document in detail the distribution of intonational pitch accents in EA sentences, across a variety of contexts and speaking styles.

3.2.1 Materials

The corpus selected for detailed auditory transcription comprised both read and (semi-) spontaneous speech materials.

The read speech materials were a subset of a larger corpus of read speech sentences collected from fifteen EA speakers, for use in experimental investigations in chapters 7 (the 'alignment' corpus) & 8 (the 'focus' corpus). The general property of rich pitch accent distribution was observed throughout all of the read speech materials during editing of the speech recordings.

A subset of data were selected for closer inspection as a representative sample of the full dataset. This involved choosing a subset of speakers (the six speakers who had undertaken all of the various recording tasks, 3 female and 3 male) and a subset of sentences (chosen so as to have a variety of syntactic and semantic structures). By design, the sentences recorded for chapter 8 (which investigates the prosodic reflexes of focus in EA) contain a systematically varied range of information structure, so the read speech materials yield evidence regarding pitch accent distribution in both neutral and non-neutral contexts.

Five of the speakers (2 female, 3 male) also provided recordings of a narrative folk tale read three times from a written text, with readings interspersed with other unrelated tasks. At the end of the recording session they were then asked to re-tell the folk tale from memory. These recordings are referred to as read and re-told narratives respectively. The second reading of the narrative and the re-told version were transcribed for each of the five speakers. This 'narrative corpus' thus yields evidence of pitch accent distribution in both read and semi-spontaneous speech styles.

Finally, a spontaneous telephone conversation extracted from the LDC Call Home Egyptian Arabic corpus (Karins et al 2002) was submitted to auditory transcription. A conversation between two female speakers was selected, since it was difficult to interpret the pitch track in conversations between male speakers due to their reduced pitch range. This material thus provides evidence of pitch accent distribution in EA in fully spontaneous conversational speech.

The full set of materials included in the transcription corpus are listed in (3.3).

(3.3) Materials included in the corpus survey (listed by recording *ID code*).

corpus section:	speech style:	materials:	# of speakers:	recording ID codes:
align corpus	read speech	8 syntactically varied sentences	6 speakers (x 3 repetitions)	*111101, 111203, 112209, 112312, 121114, 121317, 212120, 212121*
focus corpus	read speech	1 x SVO sentence in 10 different focus contexts	6 speakers (x 3 repetitions)	*121, 122, 123, 124, 221, 223, 321, 323 421, 423*
read narratives	read speech	1 x folk tale	5 speakers	*fna2, fsf2, meh2, miz2, mns2*
retold narratives	semi-spontaneous speech	1 x folk tale	5 speakers	*fna4, fsf4, meh4, miz4, mns4*
LDC corpus	spontaneous speech	spontaneous telephone conversation	2 speakers	*4862A, 4862B*

3.2.1.1 The align corpus

The 'align' sentences were selected from among a dataset collected in order to investigate the alignment of pitch targets in EA pre-nuclear pitch accents (chapter 7). For that investigation the target words were placed in carrier sentences designed to create sentences that were as natural as possible, in order to compensate for the lexical infrequency of some of the target words. There were 24 target words, yielding 24 sentences which were mixed with distractors and pseudo-randomised, and the full set read aloud three times each by 15 speakers of EA (6 female, 9 male).

During quantitative analysis of the resulting 1080 sentences it was noted that rich pitch accent distribution was the norm throughout the dataset: in all of the 1080 sentences there was a pitch accent on every content word, across sentences and across speakers. Nonetheless of the 24 target sentences 8 were selected for closer auditory transcription, and are listed in the table in (3.4) below.

(3.4) Align sentences chosen for auditory transcription

111101	Zahar	namaš	9ala	gism	il	walad	wa	9arifna	innu	-l-Hasba
	appeared	spots	on	body-	-the-	boy	and	we-knew	that-it	the-measles

"Spots appears on the boy's body and we knew that it was measles."

111203	al	asmaa?	wa	nimar	it	tilifoon	bititnisi	bi	s-sur9a
	the	names	and	numbers	the-	telephone	forget-themselves	with	-the-speed

"Names and telephone numbers are easy to forget."

112209	HaSalit	9ala	minHa	min	is	sifaara	9ala šaan	tiruuH	tidris	fi	?amriika
	she-obtained	at	grant	from	the	embassy	in-order-to	she-goes	she-studies	in	America

"She got a grant from the embassy to go and study in Amercia."

112312	miš	mumkin	id	duxuul	taani	ba9d	il	xuruug
	not	possible	the	entry	secondly	after	the	exit

"It is not possible to come in again after leaving."

121114	fii	maani9	kibiir	bayn-i	wa	bayn	id-diraasa	-l-9ulya	wa	huwwa	l-filuus
	there-is	obstacle	big	between-me	and	between	the-study	the-high	and	it	the-money

"There's a big obstacle between me and higher education and that's money."

121317	šufna	nuunu	SuGayyar	?awwi	fi	-l-	mustašfa
	we-saw	baby	small	very	in	the	hospital

"We saw a tiny baby in the hospital."

212120	il	walad	da	minamrad	xaaliS	ma	šuft	-š	zayy-uh
	the	boy	that	rebellious	very	NEG	I-saw	NEG	like-him

"That boy is very rebellious, I've never seen anyone like him."

212121	9amm-i	mimangih	nafsuh	?awwi	ba9d-ma	gayy	min	barra
	my-uncle	boastful	of-himself	very	after	he-came	from	abroad

"My uncle has been full of himself since he came back from overseas."

The 8 sentences were selected because they include a range of common syntactic structures, including coordination, embedded clauses, construct state (*iDaafa*) constructions, negation and nominal (null copula) sentences.

For each of the 8 ‘align’ sentences, 18 tokens were transcribed (6 speakers x 3 repetitions) yielding detailed transcription of 144 sentences. The six speakers were chosen from the full set of 15 speakers because they also participated in recording of the focus data collected for chapter 8. Transcriptions of all 144 sentences are provided in Appendix A (A.1-A.8). The align sentences dataset contains 792 potentially accentable content words (nouns, verbs, adjectives, adverbs).

3.2.1.2 The focus corpus

The focus sentences were selected from among a dataset collected in order to investigate the prosodic reflexes of focus in EA (see chapter 8). For that investigation two lexically distinct target sentences, together with extensions of them involving initial cleft and/or final negative continuation, were placed in frame paragraphs designed to manipulate the information status of certain words within the target sentence with regard to both presentational and contrastive focus (see section 8.2.2). There were 10 combinations of possible frames and sentence type for each lexical set, and these were pseudo-randomised and read aloud three times each by 6 speakers of EA (3 female, 3 male). The full focus dataset therefore contained 2 x 10 x 6 x 3 = 360 sentences.

During analysis of the full dataset for chapter 8, a pitch accent was observed on every content word throughout the dataset, regardless of focus context. The presence of a local pitch maximum on words falling after a contrastive focus, which were themselves ‘given’ in context (*target* words), was systematically checked to determine whether or not a local F0 maximum occurred during target words, and thus whether or not such target words were ever ‘de-accented’. The target word in each token was labelled as an interval using Praat 4.2 and the automatic pitch maximum identification function used to decide whether a local F0 maximum occurs within (or near to) the target word. When this method is used on unaccented function words the local maximum is identified as being at the start of the word because pitch simply falls steadily throughout the word; it was seen as being a practical and unambiguous way to determine whether a F0 maximum occurs or not, avoiding labeller bias. The absence of an F0 maximum would be interpreted as an instance of de-accenting; however, there were no such instances: there were no post-focal 'given' words which did not bear a pitch accent.

A subset of the focus dataset was submitted to detailed auditory transcription, by using the 72 renditions of just one target sentence shown in 3.5 below. Transcriptions of all 72 sentences are provided in Appendix A (A.9-A.12). The focus sentences dataset contains 288 potentially accentable content words (nouns, verbs, adjectives, adverbs).

(3.5) Focus sentence selected for full auditory transcription

121-124	mama	bitit9allim	yunaani	bil-layl
	mum	learns	Greek	in-the-evening

"Mum is learning Greek in the evenings"

3.2.1.3 The narrative corpus

The narrative [guHa wa bayaa9 il mooz] "Goha and the banana seller" is a traditional folk tale featuring a 'wise fool' character 'Goha'. The version used for recordings was taken from a textbook for learners of EA (Abdel-Massih 1975), which has the advantage of providing texts written using the conventions of EA orthography. These differ notably from those of Modern Standard Arabic (MSA), particularly with respect to the spelling of words containing long vowels, so that speakers immediately recognise that EA is being elicited, rather than MSA.

This particular story was chosen because it features a marketplace bargaining dialogue which includes a sequence of differing prices and monetary units. This generates a context in which items are textually given in the text, in a very natural and unforced way. The full text of the story is provided in Appendix A (A.13).

The written version of 'Goha and the banana seller' contains 211 words which could readily be classified as content words (nouns, verbs, adjectives and adverbs). The narrative was read three times each by five speakers (2 female/3 male). One repetition from each speaker (their second) was submitted to detailed auditory transcription. This yields 1055 potentially accentable content words in the read narratives corpus. The speakers' retold versions of the story varied in length from 114-158 content words each. Full transcriptions of the read narratives from all speakers, and a sample transcription of a re-told narrative (speaker *fna*) are provided in Appendix A (A.14-A15).

3.2.1.4 The spontaneous speech (LDC) corpus

The Linguistic Data Consortium Callhome Egyptian Arabic Speech Supplement corpus (Karins et al 2002) comprises recordings and partial (textual) transcriptions of 20 telephone conversations[43]. Participants were aware that their speech was being recorded but were not given any particular subject matter to talk about, and were paid a small fee for their participation. The text of a portion of each conversation is provided with the corpus. One such portion, from a conversation between two female speakers, was submitted to detailed auditory transcription here. The selected conversation portion contains a total of 119 + 315 content words for the two speakers respectively.

3.2.2 Transcription system

The notation used during auditory transcription follows the assumptions of the autosegmental-metrical theory (AM) of intonation (Beckman & Pierrehumbert 1986, Pierrehumbert & Beckman 1988, Ladd 1996), in which pitch contours are analysed into a sequence of low (L) and high (H) pitch targets, associated with the heads or edges of prosodic constituents (see section 2.1).

In particular the working hypothesis during transcription was that the basic EA pitch accent is rising, and such accents were provisionally notated as LH*. The exact phonological representation of this pitch accent is explored in detail in chapter 7 (section 7.4.2). Any pitch movements which did not appear to follow the normal pattern (a rising movement, localised roughly within the stressed syllable of the accented word) were noted, and are discussed in detail in section 3.4 below.

In order to avoid misinterpretation of instances of pitch perturbation and pitch tracking errors, a combination of auditory impression and visual inspection of the pitch track was used throughout transcription. In cases where it was not clear from visual inspection of the pitch track whether there was a meaningful rising pitch movement within a particular word or not (and where there were no potentially confounding pitch track errors or perturbations), the word was highlighted within the working window in Praat 4.2 and the pitch maximum identification function used to determine whether or not a local pitch maximum occurred within the word. As noted above, this function was used systematically on all target words in the focus corpus.

[43] I am grateful to the Department of Linguistics, UMass (Amherst) who provided me with access to their LDC corpus materials during an academic visit in Spring 2004.

As regards edge phenomena, again as a working hypothesis, it was assumed that possible boundary and phrase tones in EA would include:

(3.6)	H%	L%	indicating the right edge of an Intonational Phrase (IP)
	L-	H-	indicating the right edge of a Major Phonological Phrase (MaP)

The correct analysis of IP-final pitch accents is not uncomplicated, since these tend to resemble a falling rather than rising pitch movement. One analysis would classify IP-final pitch accents as a qualitatively different pitch accent type, involving perhaps a HL sequence, and restricted to ‘nuclear’ IP-final position. This type of analysis has been argued to best account for the facts of a number of European languages including varieties of Italian (Grice et al 2005) and European Portuguese (Frota 2000). Similarly, Rifaat (2004) analyses phrase-final pitch accents in Egyptians’ pronunciation of Standard Arabic as a qualitatively different pitch accent.

An alternative analysis of IP-final pitch accents however sees the final falling movement as the result of (very) early peak alignment in a standard LH* pitch accent, with early alignment of the H peak due to tonal crowding from IP-final boundary tones as well as proximity of the strong prosodic boundary (IP). Boundary effects on peak alignment of this latter kind have been observed for many languages including Lebanese Arabic (Chahal 2001) and Spanish (Prieto et al 1995).

The latter analysis is assumed here as a working hypothesis, and IP-final pitch accents are notated as LH*, even if there is an early peak and therefore essentially falling pitch through the word in question. Evidence from the survey which supports a LH* analysis of IP-final pitch accents is discussed in section 3.4 below. Nonetheless the primary purpose of the present investigation is to establish the *distribution* of pitch accents, and subsequent re-analysis of IP-final ‘nuclear’ pitch accents as a qualitatively different pitch accent type would not detract from the overall generalisations claimed here.

In addition to these pitch events and properties, which may reasonably be assumed to be phonologically relevant, other aspects of the pitch contour were also transcribed, which may or may not be phonologically relevant, but the detail of which were deemed potentially to be phonetically or phonologically relevant. These included possible cues to juncture or prominence, such as pause, lengthening, increased/reduced pitch excursion, and the presence of level pitch throughout a word or morpheme.

Finally, for each word or morpheme which was not associated with a rising pitch movement, the probable direction of cliticisation was transcribed (either leftwards to the preceding content word, or rightwards to the following content word). In reality the direction of cliticisation was frequently hard to establish, so notation represents only a best estimate[44]. The full set of notation marks used is illustrated in (3.7) below.

(3.7) Notation used during auditory transcription

LH*	pitch accent
H-/L-	phrase tones
H%/L%	boundary tones
>	cliticises rightward
<	cliticises leftward
---	level pitch throughout word
=	lengthening
~	pause
↑	H peak in expanded pitch range
↓	H peak in compressed pitch range
→	suspension of downstep (H peak at same height as previous peak)

3.2.3 Research questions during transcription

The aim of the transcription is to provide answers to two main research questions:

1. What is the distribution of pitch accents? (Is every content word accented?)
2. Are any pitch accent types observed other than rising (LH*) pitch accents?

The second of these questions arises due to comments in Jun (2005b:447), regarding a possible correlation between rich pitch accent distribution and small pitch accent inventory size in European languages (cf. section 3.5 below).

In addition two secondary research questions were addressed during transcription in order to generate a preliminary model of EA intonation, so that the facts of EA as observed in the corpus can be verified against other descriptions and models of EA[45]:

3. What global pitch contours are observed (e.g. declaratives vs. questions)?
4. What combinations of phrase and boundary tones are observed?

[44] The direction of cliticisation of function words in EA is discussed further in Chapter 5 section 5.4.4.
[45] During transcription the treatment of function words was also noted, and the results of this analysis are reported in Chapter 5 (section 5.4.1).

The results of the corpus survey transcription are set out below, treating each research question in turn, starting with generalisations about accentuation of content words in section 3.3.

3.3 Pitch accent distribution in EA - corpus survey results

3.3.1 Treatment of content words

The results of the transcription provide striking distributional evidence for the generalisation that in EA there is a pitch accent on every content word in the dataset; across all contexts and speech styles, over 95% of content words in EA are accented. A summary of the results is provided in the table in (3.8).

(3.8) Counts/percentages of unaccented content words in the corpus (all speakers).

	# content words	# unaccented content words	% accented content words
align sentences	792	6	99.2%
focus sentences	288	0	100%
read narratives	1055	31	96.8%
re-told narratives	686	29	95.7%
conversation	434	8	98.1%
Total	*3255*	*76*	*97.9%*

Whenever there was a borderline case (accented vs. unaccented) it was counted as unaccented, so these distributional counts represent the most conservative estimate, from the point of view of a null hypothesis that EA accents every content word.

The following syntactic categories were classified as content words: nouns, verbs, adjectives and adverbs. The verb [kaan] ‘to be’ can function as an auxiliary verb or copula verb in Arabic and was counted as a function word in both of these roles. The verb [raaH] ‘to go’ is also used in EA with auxiliary function, but was counted as a function word only when used in this sense; when used as a verb of motion it was classified as a content word (some speakers used [raaH] in the retold narratives). Prepositions were classified as function words. This included prepositional modifiers such as [Gayr] ‘except/other than’ and [laHsan] ‘in case’.

There were comparatively fewer unaccented content words in the read sentences than in longer stretches of speech, whether read or spontaneous. Nonetheless the number of unaccented content words in narrative and conversational contexts was still extremely low. This contrasts with a distinction observed in Spanish between the distribution of pitch accents in speech collected under ‘laboratory’ conditions (‘lab speech’) and spontaneous speech (Face 2003). Face found approximately 70% accented content words in spontaneous speech in Spanish, compared to a distribution in Spanish lab speech which is similar to that observed here in EA. The results of the present survey for EA however suggest that highly populated pitch accent distribution is found in both lab and spontaneous speech in EA.

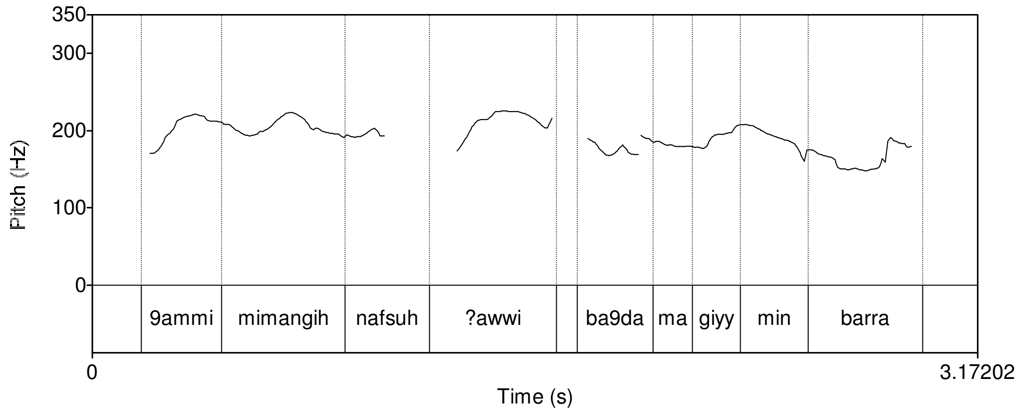
Looking at the results in a little more detail, we find that the generalisation (that every content word is accented) holds across all speakers: there was no speaker who left content words unaccented particularly more than others. The table in (3.9) below shows the actual counts of unaccented words expressed as a proportion of the number of potentially accentable content words, together with the percentage of accented content words across the whole corpus, by speaker.

(3.9) Actual counts of unaccented content words and total percentage of accented content words across the whole corpus, by speaker.

	faa	*fna*	*fsf*	*meh*	*miz*	*mns*	*‘A’*	*‘B’*
align	1/132	1/132	2/132	0/132	1/132	1/132	-	-
focus	0/12	0/12	0/12	0/12	0/12	0/12	-	-
read	-	6/211	7/211	1/211	8/211	10/211	-	-
re-told	-	6/133	1/134	10/158	9/114	3/147	-	-
LDC	-	-	-	-	-	-	5/119	3/315
TOTAL (%)	99.3	97.3	98.0	97.9	96.2	97.2	95.7	99.1

The corpus thus provides evidence that EA has highly populated pitch accent distribution, across a variety of contexts and speech styles. Sample pitch tracks and transcriptions are provided below in Figures 3.1-3.4 (stressed syllables are underlined in the transcription).

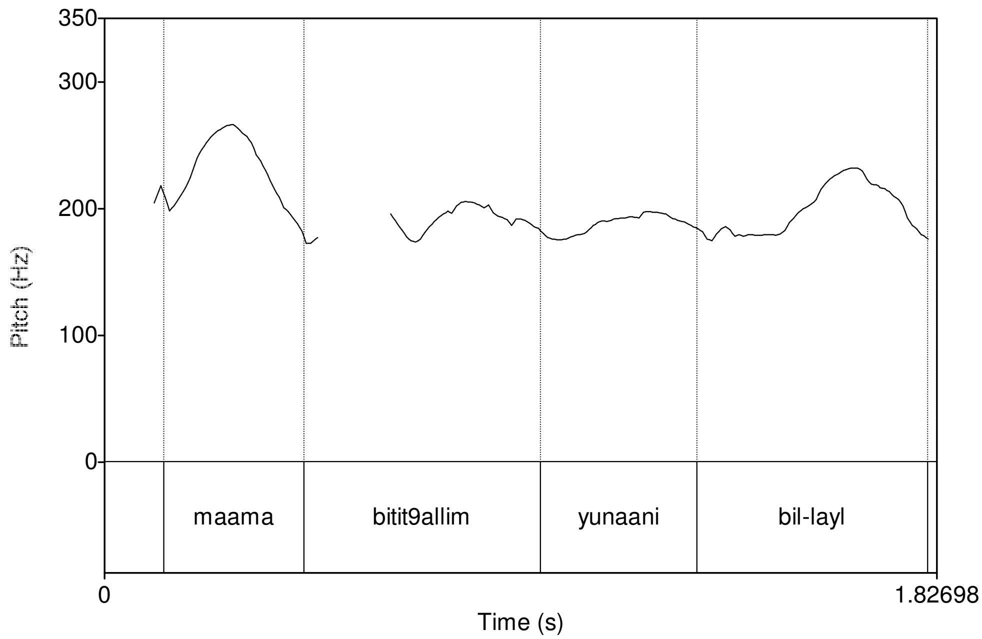
Figure 3.1 Sample read neutral declarative (*212121faa1* from the align corpus).



9amm-i	mimangih	nafs-uh	?awwi	ba9d	ma	gayy	min	barra
uncle-my	boastful	himself	very	after	that	he-came	from	overseas
LH*	LH*	LH*	LH* H-	>	>	LH*	<	!LH* L-L%

212121: "My uncle has been full of himself since he came back from overseas."

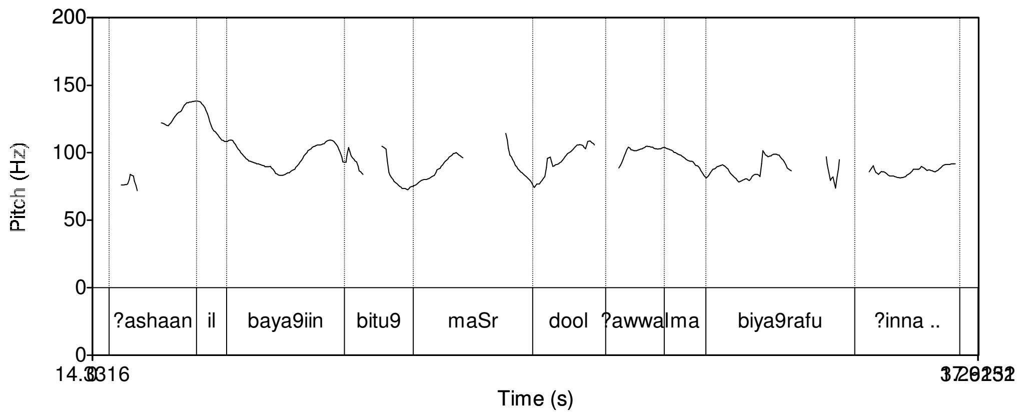
Figure 3.2 Sample read non-neutral declarative (*123faa1* from the focus corpus).



maama	bitit9allim	yunaani	bi-	-l-	layl	
Mum	learns	Greek	in-	-the-	night	
LH*	LH*	LH*	<	<	LH*	L-L%

123: "Mum is learning Greek in the evenings" (contrastive focus on [maama])

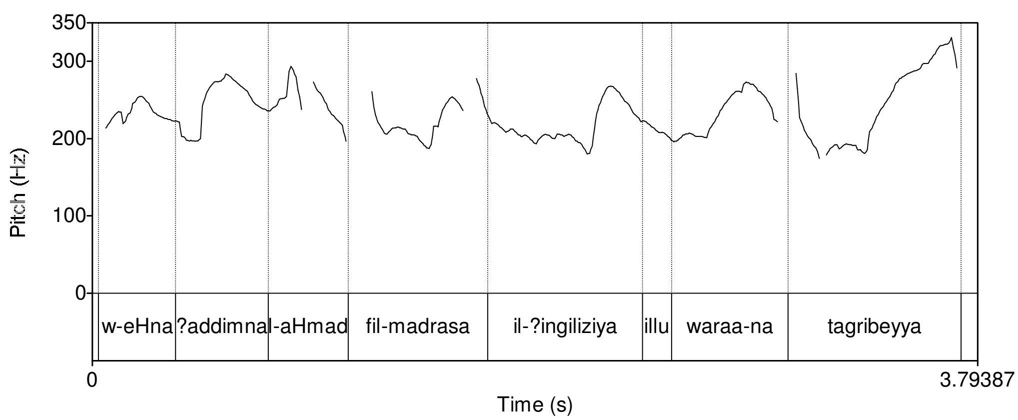
Figure 3.3 Sample extract from a retold narrative (*mns4*).



?ašaan	il	bayaa9iin	bitu9	maSr	dool	?awwal	ma	biya9rafu	?inna..
because	the	sellers	belonging	Cairo	them	first	that	they-know	that..
LH*	<	LH*	<	LH*	LH*	LH*	<	LH*	<

mns4: "Because the traders in Cairo, as soon as they know that....."

Figure 3.4 Sample extract from the spontaneous speech (LDC) corpus (*4862B*).



wi	iHna	?addimna	li-	?aHmad	fi	il-	madrasa	il-	ingliziyya ..
and	we	applied	for	Ahmed	in	the	school	the	English
>	LH*	LH*	>	LH*	<	<	LH*	<	LH*
						...illi	waraa-na	il-	tagribiyya
						that	behind-us	the	near
						<	LH*	<	LH* H-H%

4862B (330.53-334.27): "And we have applied for (a place for) Ahmed at the English school right behind us."

3.3.2 Instances of unaccented content words in EA

Across the whole corpus, approximately 2-4% of content words were unaccented. This section examines particular categories and contexts which seem to favour non-accenting of a content word. These include 'utterance-peripheral' items, words of high frequency, and modifiers, as well as words occurring in fast renditions of certain speakers.

The set of utterance-peripheral items comprises 12 tokens across the whole corpus, and of these 9 are ‘reporting verbs’ occurring in the read or retold narratives. For example [?aal-luh] ‘he said to him’, [?ul-luh] ‘say to him’ and [?aaluu-luh] ‘they said to him’ are consistently unaccented by speaker *miz* in the retold narrative. These verbs are peripheral to the dialogue which carries the narrative of the story. The remaining tokens which seem to be utterance-peripheral are all instances of the word [maani9] ‘obstacle’ in the opening phrase of an align corpus sentence (*121114*):

(3.10)

fii	maani9	kibiir	bayn-i	wa	bayn	id-diraasa	-l-9ulya
there-is	obstacle	big	between-me	and	between	the-study	the-high
					wa	huwwa	l-filuus
					and	it	the-money

“There’s a big obstacle between me and higher education and that’s money”

The word [maani9] is produced with either no pitch accent, or a very compressed pitch accent, in four tokens, by three different speakers (*faa3*/*fna3*/*fsf1*/*fsf2*). The auditory impression of the way this opening phrase is produced is that of a high ‘anacrusis’, and compression (or lack of accentuation) of [maani9] lends prominence to the following adjective [kibiir] ‘big’. The compression could however also be a by-product of a clash between adjacent stressed syllables in [fii maani9] (stressed syllables underlined), causing undershoot of the leading L target of the LH* pitch accent on [maani9].

There were a small number of unaccented content words which can be analysed as being high frequency words, either because they fulfil a discourse function rather than a lexical function in context, or because they form part of an idiomatic phrase. For example, the discourse particle [ya9ni] ‘well/I mean’ (lit. ‘it means’) is usually unaccented, by all speakers throughout the corpus, even though grammatically it takes the form of an inflected finite verb. The noun [nahaar] ‘day’ is unaccented by speaker A in the LDC corpus in the idiom [ya nahaar abyaD] lit. ‘oh white day’ (which equates roughly to ‘what wonderful news!’). The verb [xalli] ‘keep’ in the set phrase [xalli baalak] ‘take care’ (lit. ‘keep your wits’) is unaccented by most speakers (in the read and retold narratives)[46].

[46] Compare also unaccented [xad] ‘take’ in [xad baalak] ‘take care’ in one token from speaker *meh*.

Both utterance-peripheral and high frequency words could be said to be of low semantic weight and thus perhaps susceptible to ‘de-accenting’. Ladd (1996 ch6) demonstrates that languages differ with regard to de-accenting of items of low semantic weight, in that Germanic languages such as English or German tend more towards de-accenting of such items than Romance languages such as Italian or Romanian. In EA there are a tiny number of such instances which suggests that EA should be classified with the Romance languages in Ladd’s two-way typological grouping (see further discussion of this typology in chapter 5 section 5.1.1). The presence of some such tokens could be taken as evidence that accenting of low semantic items is a tendency rather than a rule in EA (and presumably also in the Romance languages). Alternatively, these occasional instances of de-accenting could result from the fact that some speakers in the corpus had a reasonable command of English resulting in a minority of Germanic-like renditions[47].

The next set of unaccented content words which can perhaps be grouped together are ‘serial’ verbs and pre-head modifiers. These are not of low semantic weight, but occur in a structurally weak position. This set includes instances of an unaccented first verb in a ‘serial’ verb construction, in which one might consider the first verb to be playing a functional rather than lexical role. The example is the verb [fakkar] ‘he thought/decided’ in the phrase [fakkar yinzil maSr] ‘he decided to go to Cairo..’. Similarly there are a small number of unaccented modifiers which precede their syntactic head instead of occurring in unmarked head-modifier order, or quantifiers occurring as the first element in a construct state (*iDaafa*) genitive construction:

(3.11) Examples of unaccented modifiers observed in the corpus survey.

unaccented word	*context*		*in:*
[Tuul] ‘all’	[Tuul 9umruh] ‘all of his life’	iDaafa	*fsf2/mns2*
[kiilu] ‘kilo’	[kiilu mooz] ‘a kilo of bananas’	iDaafa	*fsf2*
[kulla] ‘all’	[kulla Haaga] ‘every thing’	iDaafa	*mns4*
[taani] ‘second/next’	[taani yoom] ‘the next day’	pre-head	*4862A*
[aaxir] ‘the end’	[aaxir disembir] ‘at the end of December’	iDaafa	*4862B*
[zayy] ‘like’	[zayy iš-ša??a] ‘like the (other) flat’		*4862B*

There were a small number of words which were unaccented despite falling in full argument positions in many cases, but occurred in a section of noticeably fast speech.

[47] I would tend to prefer the former explanation, since in recordings of English sentences containing items of low semantic weight, collected for a pilot study, a non-trivial number of accented tokens were found.

Note however that in a previous study which specifically elicited speech rate contrasts all content words were accented even in fast speech rates (Hellmuth 2004).

(3.12) Examples of words which were unaccented in fast renditions

unaccented word	*in:*
[taani] 'again'	*112312miz3*
[bayyaa9] 'seller'	*mns2*
[xamsa] 'five'	*mns3*
[SaaG] 'piastres'	*mns3*
[xamas] 'five'	*miz3*
[mooz] 'bananas'	*meh4*
[guHa] 'Goha'	*meh4*
[Haaga] 'something'	*fna4*

Finally, there remains a set of words which were transcribed as unaccented but for which there is no obvious explanation. This set comprises 34 words which were unaccented in one or at most two tokens each, occurring mostly in the read and retold narratives (examples are listed in Table 3.8 below). Despite there being no apparent explanation for non-accenting of these words, they nonetheless represent less than 1% of the total corpus.

(3.13) Examples of words which were unaccented with no obvious explanation.

unaccented word	*in:*
[mustašfa] 'hospital'	*121317mn3*
[tištiri] 'you buy'	*fsf2/mns2*
[faat] 'passed by'	*fna2/miz2/mns2*
[waaHid] 'a' (lit. 'one')	*fna2*
[kiilu] 'kilo'	*fsf2*
[ra?yak] 'your opinion'	*fna2/fsf2*
[la?] 'no'	*fna2/miz2*
[?ahu] 'here'	*4862A*
[Hoosa] 'trouble/mess'	*4862A*
[lissa] 'soon'	*4862B*

3.3.3 Summary

Having established the generalisation that in EA the overwhelming majority of content words bear a pitch accent, the next section outlines the properties of both global and local pitch movements observed in the corpus survey, in order to establish a working AM model of EA intonation.

3.4 A model of EA intonation

This section provides a brief overview of the global pitch contours observed in different sentence types in the corpus (section 3.4.1), then gives a more detailed description of the properties of local pitch movements observed on accented words (section 3.4.2). These are used to motivate a model of EA intonation, within the autosegmental-metrical (AM) theory of intonation, which is then compared to other competing AM models that have been proposed for EA (section 3.4.3).

3.4.1 EA global pitch contours observed in the corpus survey

A large proportion of the corpus consists of declarative sentences, elicited for other purposes as individual carrier sentences for embedded target words (the align section of the corpus), or forming part of a narrative paragraph or folk tale (the focus and narrative sections of the corpus). For the most part these sentences were produced by speakers in one of two ways: either as a plain declarative or as a 'non-final' declarative (ending with a continuation rise expressing non-finality, leading into a following sentence). The global pitch contours observed in such cases are described below and compared to those in existing descriptions and analyses of EA intonation.

There were only a small number of non-declarative contexts, such as yes-no questions, in the corpus (in the narratives and LDC sections of the corpus) but these are also described and then compared to existing accounts.

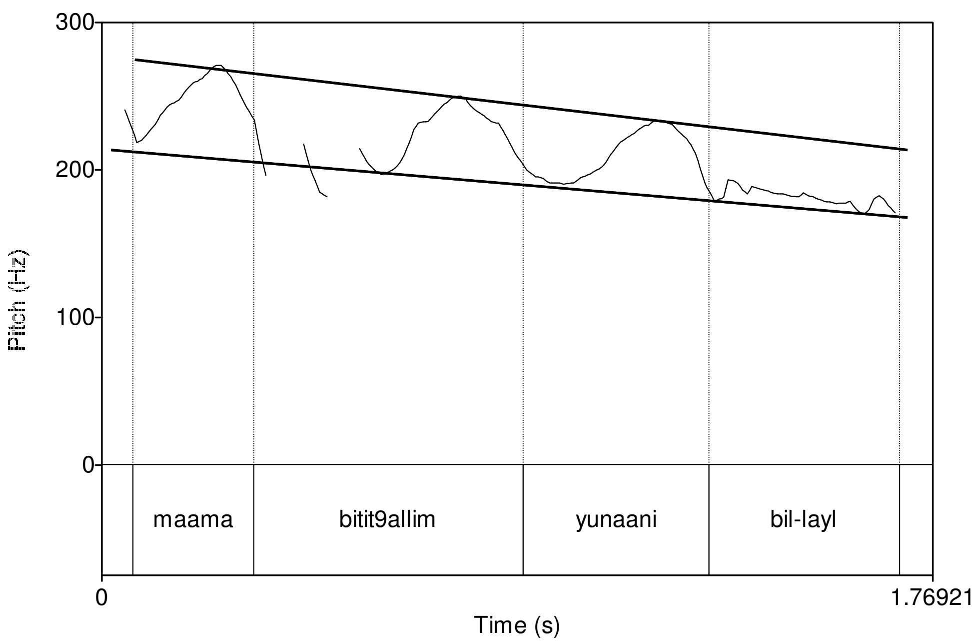
3.4.1.1 Declarative sentences

A typical EA declarative intonation contour shows an overall falling pattern, with a rising pitch accent localised around the stressed syllable of each content word. Plain declaratives end with falling pitch notated as a L-L% phrase-/boundary-tone sequence.

The height of the peaks and valleys of pitch accents on subsequent content words fall steadily throughout the utterance. This is assumed here to be an effect of declination, rather than phonological downstep, because the peak of the final pitch accent is often very much lower than would normally expected in the declination sequence, and this latter phenomenon is analysed as being phonological, with such pitch accents notated !LH* in the auditory transcription. A similar effect has been observed in English, and is known as 'final lowering' (Liberman & Pierrehumbert 1984).[48]

Both declination and final lowering can be seen in the example of a standard declarative given in Figure 3.5, which shows a declarative sentence from the focus corpus, elicited in a neutral context. Approximate register lines, superimposed on the pitch contour, serve to illustrate the falling height of both high (the top register line) and low (the bottom register line) turning points in subsequent pitch accents through the sentence. The pitch peak of the final accent is considerably lower than predicted by the slope of the top register line.

Figure 3.5 Neutral declarative showing declination & final lowering (*122fna2*).

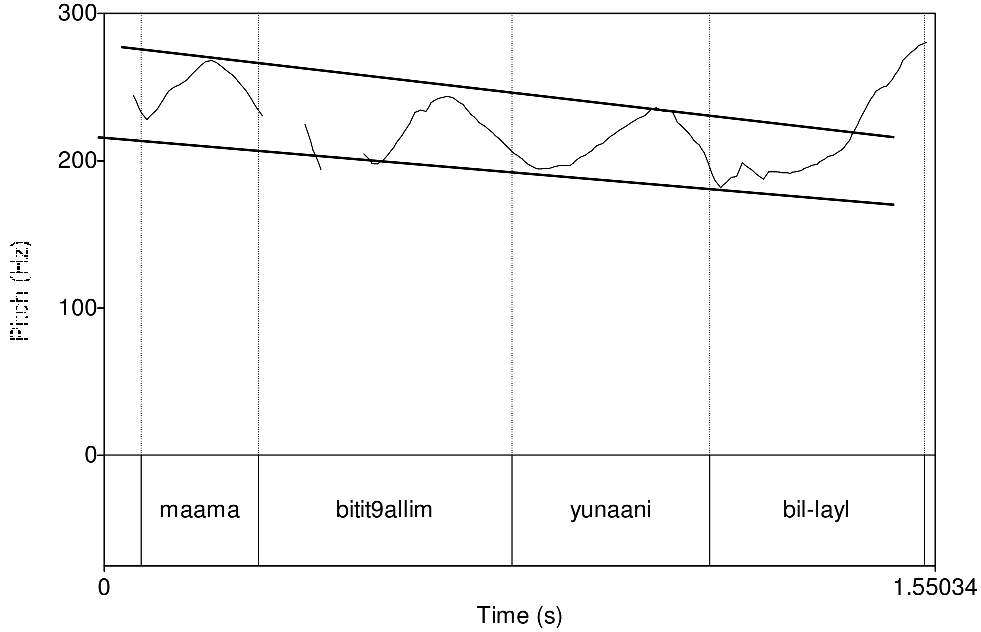


maama	bitit9allim	yunaani	bi-	-l-	layl	
Mum	learns	Greek	in-	-the-	night	
LH*	LH*	LH*	<	<	!LH*	L-L%

122: "Mum is learning Greek in the evenings" (neutral context)

[48] A potential argument against this view is the fact that declination appears to be very much under the control of speakers, so that the declination sequence can be suspended and successive peaks expressed in varying pitch range (see for example Figure 3.2 above). Such effects are analysed as changes in pitch register/range and are discussed in chapters 5 & 8 in the context of cues to phrase boundaries and expression of focus.

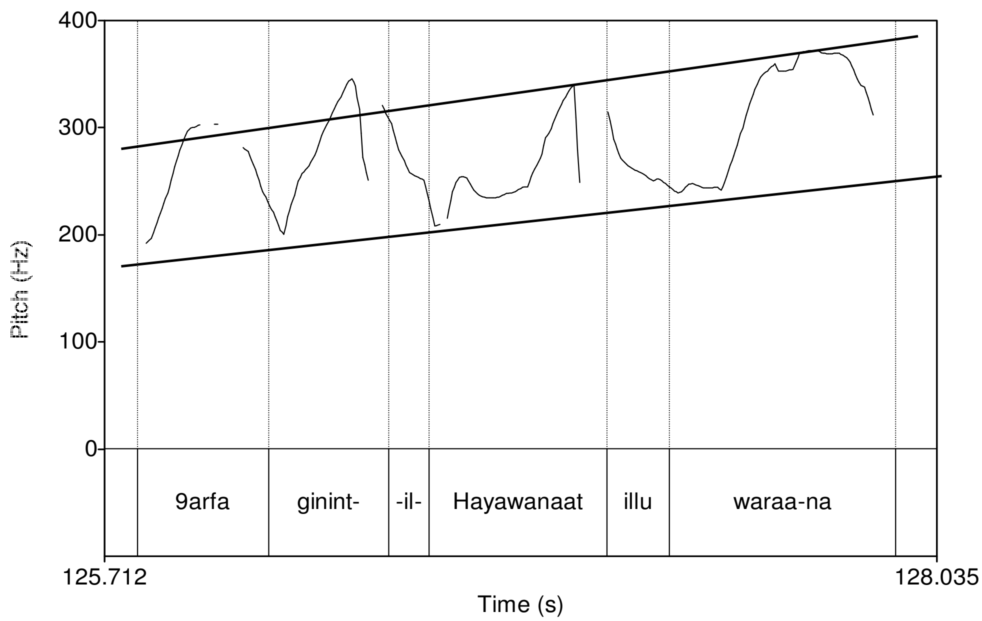
Figure 3.6 Neutral declarative showing a continuation rise (*122fna1*).



maama	bitit9allim	yunaani	bi-	-l-	layl	
Mum	learns	Greek	in-	-the-	night	
LH*	LH*	LH*	<	<	!LH*	H-H%

122: "Mum is learning Greek in the evenings" (neutral context)

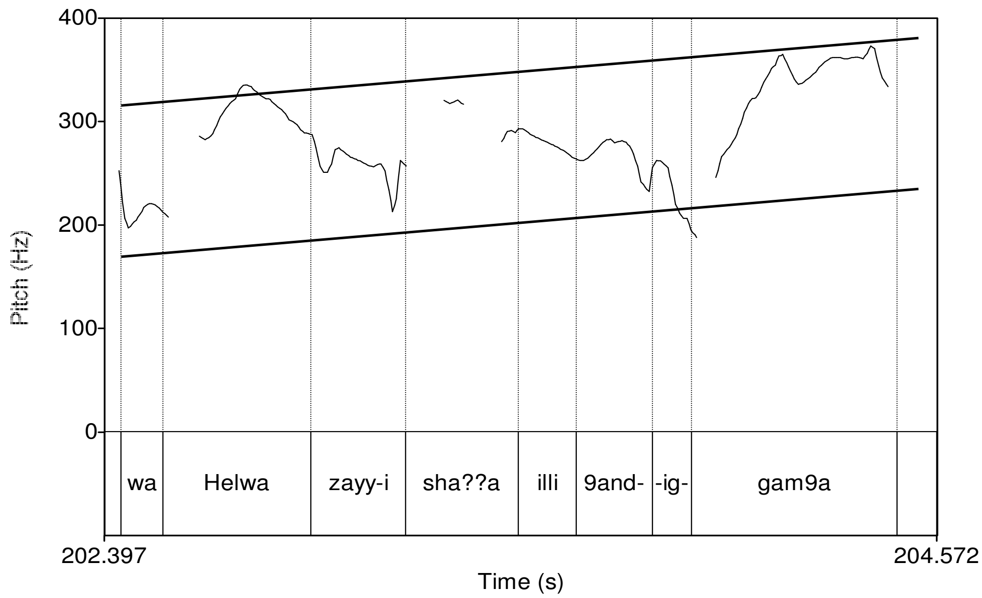
Figure 3.7 Yes-no question showing rising pitch register (*4682B*).



9arfa	gineena	il-	Hayawanaat	illi	waraa-na
knowing (f.)	garden-	-the-	-animals	that	behind-us
LH*	LH*	<	LH*	<	LH* H-H%

4682B (449.77-451.95): Do you know the zoo behind us (i.e. behind our house)?"

Figure 3.8 Yes-no question showing rising pitch register (*4682A*).



wi	Hilwa	zayy	il-	ša??a	illi	9and	il-	gam9a
and	nice	like	the	flat	that	at	the	university
>	LH*	<	<	LH*	<	<	<	LH* H-H%

4682A (526.40-528.50): "Is it as nice as the flat near the university?" [49]

3.4.1.2 Continuation rises

Non-final declarative sentences display the same declination across subsequent pitch accents as observed in plain declarative sentences. Non-finality is however expressed by rising pitch at the end of the sentence, notated as a H-H% phrase-/boundary-tone sequence. The continuous rise in pitch between the last pitch accent and the edge tones (H-H%) in this type of sentences is argued here to be evidence in favour of analysis of final pitch accents in EA as having the same (rising) phonological specification as all other pitch accents: LH*. If final pitch accents were phonologically specified as a falling accent (such as H*L or HL*) then we should see evidence of a fall in pitch to a low pitch target before the final boundary rise (the H-H% combination).
An example of a continuation rise is provided in Figure 3.6, which shows another neutral-context token from the focus corpus (*122fna1*).

3.4.1.3 Questions

There are a small number of yes-no questions (YNQs) in the spontaneous conversation section of the corpus (LDC)[50]. These all take the form of 'declarative questions', in

[49] Note that the pitch level of the LH* peak on the word [ša??a] 'apartment, flat' in Figure 3.8, is not clear due to the voiceless initial segment and perturbation from a particularly creaky geminate glottal stop [?].

which the sentence has the syntactic structure of a statement, and the question status of the utterance is expressed only by prosodic means.

The examples of declarative questions in the corpus show pitch accents on all content words, with rising register lines for both H and L turning points through the utterance followed by a final rise (H-H%). The features of YNQs are illustrated in two examples of YNQs from the LDC section of the corpus (*4682B/4682A*) shown in Figure 3.7-3.8.

3.4.1.4 Comparison to global pitch contours reported in the literature on EA

The global pitch contours observed during transcription of the corpus are here compared to those observed by other authors in various instrumental studies on EA.

In a small production study with one speaker, Norlin (1989) elicited declarative sentences and declarative questions (with no syntactic question-marking)[51], and found that (neutral) declaratives showed continuous declination throughout the sentence. In another instrumental study, Rifaat (1991) found also that in declarative statements the pitch height of H and L turning points fall through successive stressed syllables.

In declarative questions, Norlin found that declarative questions started at the same F0 level as declarative sentences but showed no declination with global F0 "more or less horizontal" (p48) before a final rise.

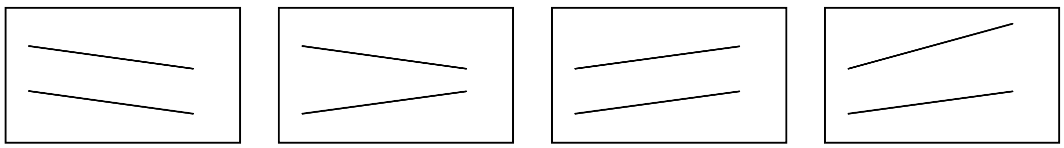
Ibrahim et al.(2001) report slightly different results for declarative questions in a production study of elicited lab speech. They provide accurate 'linear trendlines', calculated mathematically from the F0 contour in declaratives, and three types of question: WHQs (containing an overt wh-word), YNQs (starting with a question word), and declarative questions (declarative syntax distinguished intonationally). Upper/lower trendlines were calculated on all points lying above/below a global trendline calculated from all F0 values in an utterance using the least error squares method.

Ibrahim et al. find that declarative sentences show declination as observed by other authors, with both upper and lower trendlines falling throughout the sentence. All of the

[50] The only instance of a wh-question in the LDC conversation (*4682B* 542.33-544.07) is used to scold a child and functions more as an exclamation than as a real question, and is thus deemed unrepresentative.
[51] Norlin also elicited questions and statements in focus contexts, and these findings are discussed in chapter 8.

three question types showed a rising lower trendline, but only YNQs and declarative questions showed a rising upper trendline (with greater upward slope in the upper trendline in declarative questions than YNQs). In contrast, WHQs had a falling upper trendline, which when combined with the rising lower trendline resulted in narrowing pitch range through the sentence. These generalisations are illustrated in schematised form in (3.14) below.

(3.14) Schematised upper and lower F0 trendlines in EA (based on Ibrahim et al 2001).



declarative sentence WHQ YNQ declarative question

The authors suggest that WHQs contain the most syntactic cues to question status (an overt in-situ question word) and so prosodic cues are lessened, or are of lesser importance. In YNQs there are fewer syntactic cues to question status so prosodic cues are enhanced. Prosodic cues are the strongest of all in declarative questions, which have no syntactic cues to question status.

In his extensive corpus survey of broadcast Modern Standard Arabic, recorded from Egyptian radio, Rifaat (2004) notes that final-rising pitch in a phrase (a final LH pitch accent, in his notation) is used consistently to indicate incompleteness. This contrasts with declination throughout the utterance, which “is one of the major tools to indicate completeness” (Rifaat 2004:10). He also notes instances of ‘final lowering’ in which the final pitch accent of the utterance is produced with a significantly lower peak than expected from declination alone.

El Zarka (El Zarka 1997:355ff.) reports similar results in her study of MSA intonation, as produced by EA speakers, including the fact that final pitch accents may be realised either in a lower pitch range (final lowering) or with an early peak. She reports continuous declination in MSA WHQs, though with a slightly higher initial pitch level at the start of a question (compared to the start of a statement); this high pitch is realised on the wh-word, which is always sentence-initial in MSA and is always accented. In MSA YNQs, El Zarka also reports declination through the sentence (with both upper and lower trendlines sloping downwards, followed by final rising pitch at the end of the utterance.

Comparing these findings in the literature with those observed in the present corpus study, as observed by all other authors, most declaratives observed in the corpus showed continuous declination throughout the utterance. Final lowering is observed by both Rifaat (2004) and El Zarka (1997) in the EA pronunciation of MSA, as it is also in the present corpus. The examples of questions from the corpus, discussed in section 3.4.1.3 above, are declarative questions, and display rising 'trendlines' through successive L and H turning points, matching the findings of Ibrahim et al. (2001).

The next section reviews the properties of local pitch movements, localised around the stressed syllables of content words, in order to determine the number of pitch accent types in EA.

3.4.2 EA pitch accent types observed in the corpus survey

The vast majority of pitch movements observed during auditory transcription of the corpus were rising pitch movements, localised around the stressed syllable of each content word. The exact alignment properties of these 'standard' rising pitch targets is investigated quantitatively in chapter 4 (and in further detail still in chapter 7).

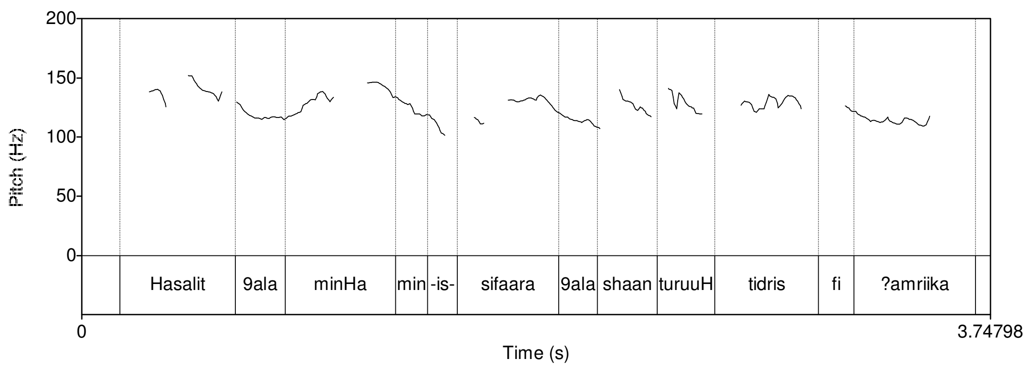
The properties of the remaining small number of potentially 'non-standard' pitch accents, as observed during detailed examination of the pitch track and spectrogram, are set out here. Such cases fall into four categories: i) pitch accents showing an unusual local pitch contour after the stressed syllable, which are analysed as instances of an inserted phrase tone; ii) an unusual pitch contour between two content words (possible absence of L turning point), which are analysed as undershoot of the L pitch target due to tonal crowding; iii) pitch accents showing an unusual local pitch contour before the stressed syllable; and, iv) pitch accents in sentence final ('nuclear') position.

3.4.2.1 Non-standard local pitch contour after the stressed syllable

There were a small number of words in which pitch continues to rise after the end of the stressed syllable, instead of immediately starting to fall again towards the next pitch accent (on the stressed syllable of the following content word). The pitch rise in these cases can however be analysed as resulting from an internal phrase boundary, so that the rising LH* pitch accent is followed by a high phrase tone (H-).

For example, both instances of this non-standard local pitch contour in the align section of the corpus occur in a position where it would be plausible to insert a phrase boundary, or indeed in a position where other speakers did insert a more salient boundary in their renditions of the same sentence. An example of the latter type is illustrate below in Figure 3.9, which shows a LH*H- combination on the word [minHa] ‘grant’ (in *112209mns3*) in a position which other speakers mark with a H- and also lengthening and/or pause.

Figure 3.9 Non-standard local pitch contour after the stressed syllable of [minHa] ‘grant’, analysed as indication of a H- phrase tone (*112209mns3*).



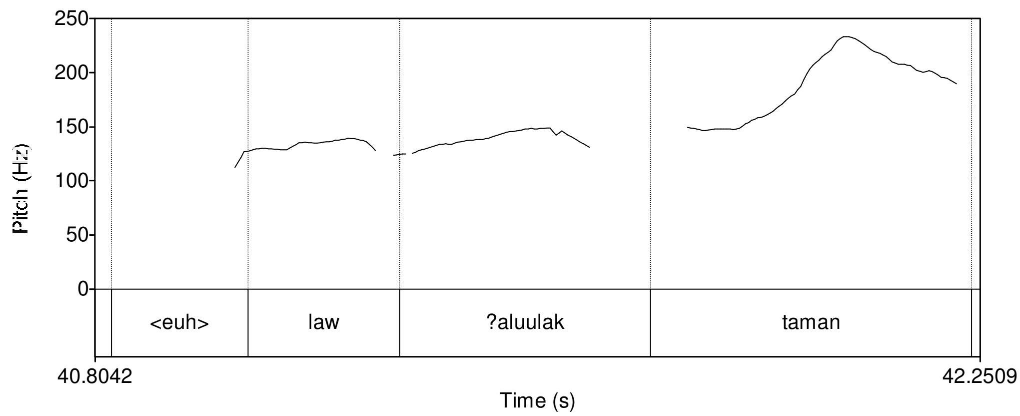
HaSalit	9ala	minHa	min	is	sifaara	9ala	šaan	tiruuH	tidris	fi	?amriika
LH*	---	LH*	<	<	LH*	<	LH*	LH*	LH*	<	!LH* L-L%

112209: “She got a grant from the embassy to go and study in America.”

Similarly there are cases where pitch falls after the stressed syllable to an ‘elbow’ which coincides with the right edge of the word, rather than falling gradually across all intervening unstressed syllables until the next pitch accent (the stressed syllable of the following content word). Some can again be analysed as instances of L- tone, indicating that the accented word falls at the right edge of an internal phrase of some kind. However there are one or two others which fall in a position in which it is unlikely that a L- tone would be inserted, such as in the case illustrated in Figure 3.10 below[52]; however, there are too few tokens to determine what other factors may be at issue.

[52] El Zarka (p.c.) also observed configurations of this kind in her EA pronunciation of MSA data.

Figure 3.10 Non-standard local pitch contour after [law] ‘if’ (*mns4*).

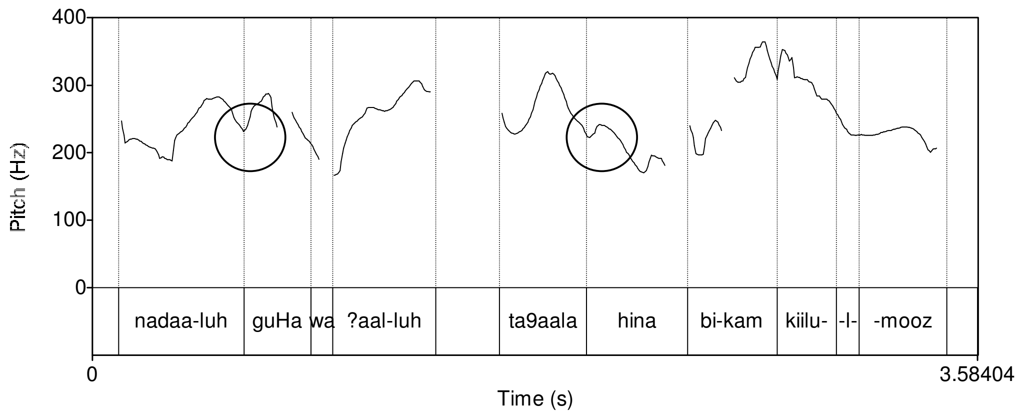


law ?aaluu-lak taman
LH* ? LH* LH* H-H%
mns4: ‘If they tell you a price....’

3.4.2.2 Non-standard local pitch contour between two content words

As in the anacrusis context mentioned above, there are cases where the pitch valley between two accented words is smaller than might be expected. This could either be as a result of undershoot of the leading L target of the second LH* pitch accent, or it could be evidence for a distinct H* pitch accent. At present I retain the view that these are cases of undershoot, since there are a very small number of instances, and they tend to occur in contexts where there are a small number of unstressed syllables between accents, i.e. undershoot could arise as tonal repulsion from an upcoming pitch accent. There are two such cases in the example provided in Figure 3.11 below. In this speech extract the leading L of the LH* pitch accent in two words, [guHa] ‘Guha’ and [hina] ‘here, does not reach the same level low of pitch as the L target in other words, nor at the level that might be expected from a steadily descending lower register line drawn through successive L targets in the utterance.

Figure 3.11 Undershoot of leading L due to tonal crowding (*fna2*).



nadaa	-luh	guHa	wa	?aal-	luh	ta9aala	hina
called	to-him	Guha	and	said	to-him	come	here
LH*	<	H*	<	LH*	< H-	↑LH*	!LH* L-L%
bikam	kiilu	l-	mooz				
how much	kilo	the	bananas				
↑LH*	LH*	<	!LH* H-L%				

fna2: 'Guha called to him and said: "Come here! How much is a kilo of bananas?".'

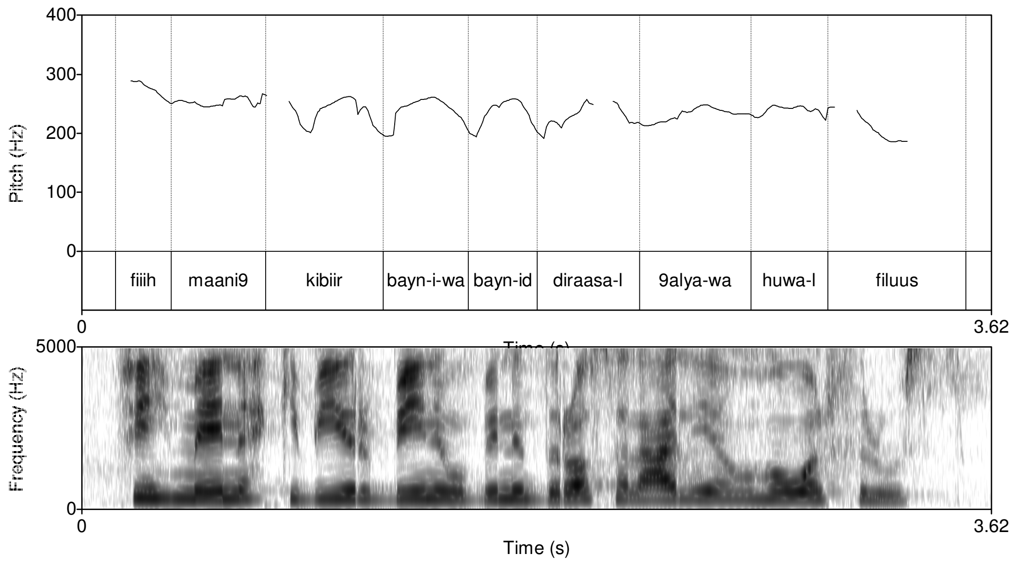
3.4.2.3 Non-standard local pitch contour before the stressed syllable

A potential exception to the LH* pitch accent type is a small number of cases in which the leading L appears to be aligned with the start of the word rather than with the onset of the stressed syllable. There are 12 such cases; however, 11 of them occur in instances of the word [diraasa] 'study' in a sentence from the align section of the corpus (*121114*). The other case of early alignment of the leading L target is in the word [SuGayyar] 'small' (in align sentence *121317*). An example of early alignment of the L target in the word [diraasa] 'study' is provided in Figure 3.12 below.

It is striking however that these are the only such cases in the dataset. It would be plausible to think that, in the case of /diraasa/, application of vowel syncope has caused the onset of the word and the stressed syllable to coincide: [draasa]. EA has a highly productive process of vowel syncope affecting high vowels /i and /u/ in monomoraic syllables, provided that the resulting consonant cluster has an upward sonority slope (Watson 2002:70-72).

However as can be observed in Figure 3.12, which is typical of all the instances of early L alignment on [diraasa], a vowel is clearly visible between the burst of the [d] and the following [r] trill which precedes the long stressed vowel. Nonetheless the L valley turning point appears to coincide with the onset of the word not the onset of the stressed syllable[53]. These examples could in principle indicate some kind of marginal edge alignment of the L leading tone of EA LH* pitch accents. However there are only a very small number of cases observed in the present corpus, in which the properties of the local contour pattern in all other respects with the standard LH* rising pitch accent. As a result these cases are set aside and are not deemed to constitute evidence of a different pitch accent type.

Figure 3.12 Example of early alignment of leading L target (*121114fsf1*).



fii	maani9	kibiir	bayn-i	wa	bayn-	id-diraasa	-l-	-9ulya...
there-is	obstacle	big	between-me	and	between	the-study	the	high
LH*=	LH*	LH*	LH*	<	<	LH*		<
					...wa	huwwa	-l-	filuus
					and	it	the	money
					<	LH*	<	!LH* L-L%

121114 "There's a big obstacle between me and higher education and that's money."

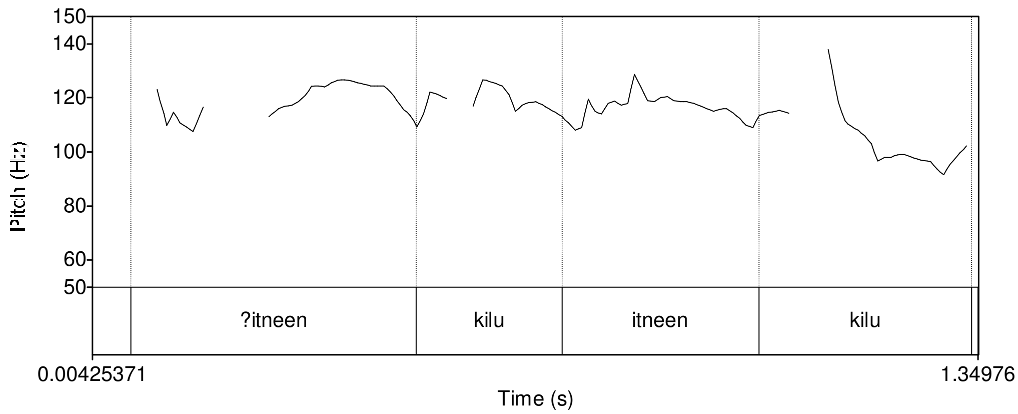
3.4.2.4 Non-standard local pitch contour on sentence-final content words

Finally, there are cases in the align section of the corpus in which sentence-final pitch accents appear to be falling rather than rising. As discussed in section 3.2.2 above, these could plausibly be analysed as a different type of pitch accent (i.e. a falling pitch accent), since it is not unusual for languages to distinguish between 'pre-nuclear' (non-final) and nuclear (final) accents. However there are also cases of sentence-final pitch accents which are clearly cases of a rising accent with an early peak, presumably arising because the sentence-final accent is adjacent to a strong prosodic boundary and thus

[53] In addition a syncope based explanation could not apply in the case of /SuGayyar/ which would have an initial cluster with falling sonority.

subject to leftward shift of the peak as in Lebanese Arabic and Spanish (Prieto et al 1995). Figure 3.13 shows a sequence of two phrases from a retold narrative, which together form a single IP (the example occurs during the bargaining section of the story). The word [kilu] ‘kilo’ is repeated; in the first (non-final) instance the word bears a LH* pitch accent with usual alignment; in the second (phrase-final) instance the peak is shifted leftwards, so that there is falling pitch through most of the word.

Figure 3.13 Example of falling pitch/early peak in a final pitch accent (*miz4*).



<euh>	itneen	kiilu	itneen	kiilu
	two	kilos	two	kilos
---	LH*	LH*	LH*	LH* L-L%

mns4: ‘Two kilos! Two kilos!’.

Crucially, as described in section 3.4.1.2 above, there are no instances whatsoever in the corpus of falling final pitch accents preceding high boundary tones. For this reason I continue to assume that apparently falling sentence-final pitch contours can be decomposed into a standard LH* pitch accent with an early peak, followed by a phrase-final and IP-final boundary tone combination such as L-L%.

3.4.2.5 Summary: EA pitch accent types

Overall then, even allowing for these marginal cases, the overwhelming majority of pitch accents in the align corpus are rising pitch accents in which the rise is aligned to the stressed syllable of the word. All pitch accents observed are therefore analysed as tokens of a single phonological object: LH*.

It is striking not only that there are so many pitch accents in EA (one on every content word) but that it appears to be the same pitch accent type used to mark each word.

3.4.3 An AM model of EA intonation based on the corpus survey

3.4.3.1 Proposed inventory of pitch accents and edge tones in EA

Based on the survey of global and local pitch contours observed during auditory transcription of the corpus, I propose the inventory of phonological pitch accents and edge tones for EA shown in (3.15).

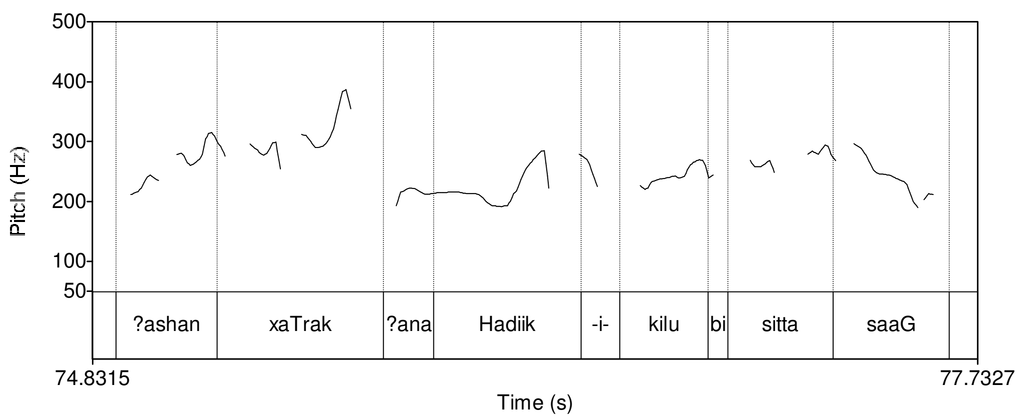
(3.15)	LH*		'default' pitch accent, on every content word
	H%	L%	indicating the right edge of an Intonational Phrase (IP)
	L-	H-	indicating the right edge of a Major Phonological Phrase (MaP)

The most common phrase and boundary tone combinations observed in the corpus were L-L% and H-H%. Nonetheless a few examples of H-L% and L-H% were also observed, as described below, which suggests that phrase and boundary tones may freely combine in EA:

(3.16)	L-L%	declarative	
	H-H%	continuation rise	
	H-L%	mid-level	used in reported speech ('open-ended')
	L-H%	fall-rise	signifies reproach/irony (rare)

A H-L% boundary tone, which sounds like an open-ended, mid-level final tone, appears in a few instances in the read and re-told narratives. It is particularly common in cases of reported speech, and is found at the end of the section of indirect speech, as in the example in Figure 3.14 below.

Figure 3.14 Example of H-L% 'open-ended' boundary tone (*fsf4*).

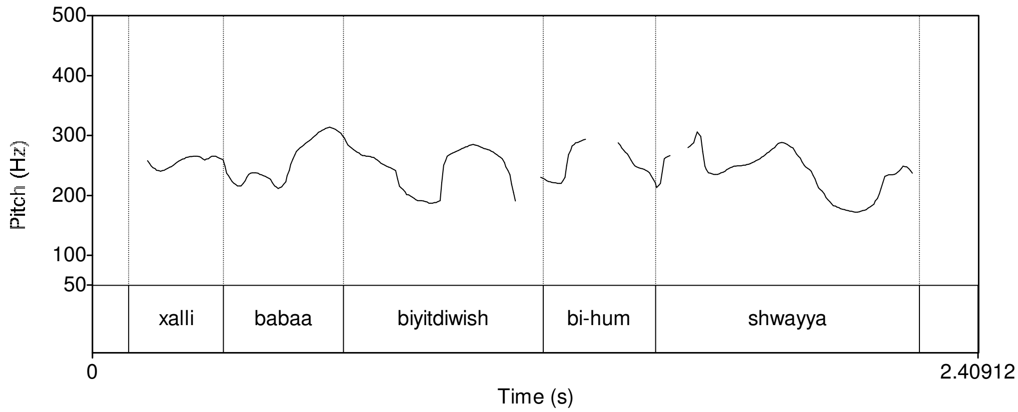


9ašaan	xaTrak	?ana	Haddiik	ik-	kiilu	bi	sitta	saaG
in-order	sake-your	I	will-give-you	the	kilo	for	six	piastres
LH*	LH* H-	---	LH*	<	LH*	<	LH*	LH* H-L%

fsf4: 'For your sake, I will give you the kilo for six piastres'.

The L-H% fall-rise boundary tone combination is much less common and was only observed in the spontaneous speech (LDC) corpus. As illustrated in Figure 3.15 below this boundary tone combination, together with the preceding rising pitch accent (LH*) results in a rise-fall-rise at the end of the phrase. Chahal (2001:162) reports that use of the same combination in Lebanese Arabic is deemed 'foreign' and is thought to be a borrowing from English (in which it is commonly used to express emphasis).

Figure 3.15 Example of L-H% 'reproach' boundary tone (*4682A*).



xalli babaahum yitdiwiš biihum šiwayya
LH* LH* LH* ↑↑LH* LH* = L-H%

4682A (389.84-392.03): 'Let their father look after them for a while!'

3.4.3.2 Comparison to other AM models of EA and EA productions of MSA

Four prior analyses of EA or of EA productions of Modern Standard Arabic (MSA) have been proposed within the autosegmental-metrical (AM) framework.

In the earliest of these, Rifaat (1991) proposes an inventory of two pitch accents for the EA pronunciation of MSA: a rising LH pitch accent on all pre-final stressed syllables, and a falling HL pitch accent on final stressed syllables. He notes that the shape of the LH pitch movements on pre-final stressed syllables are "quite redundant and they appear to carry no distinctive information". In a later paper, based on a large corpus survey of the same dialect, Rifaat (2004) refines his definitions of the two basic pitch accents, and adds a further two marginal pitch accents to the inventory:

(3.17) Pitch accent inventory (Rifaat 2004)[54]:

[' denotes association to a stressed syllable; # denotes a phrase boundary]

'H	pitch accent occurring in all positions
'HL#	falling pitch accent occurring only in utterance-final position
'L	very infrequent accent occurring before or after a focussed 'H
LH#	utterance-medial continuation rise or utterance-final YNQ rise

The last of these, LH#, fulfils the function within Rifaat's model of a boundary tone combination in the present thesis. Rifaat chooses to model EA without the use of any boundary tones or edge tones which he states are predictable from the properties of preceding pitch accents. In order to describe all of the contours he observes however he includes the LH# tone as a pitch accent in the inventory, even though it does not associate with a stressed syllable, and its distribution is limited to phrase-/utterance-edges. The 'HL# accent is similarly restricted in its distribution to phrase-/utterance-edges, and Rifaat's (2004) notation thus encodes his (1991) distributional statement, that falling HL accents are reserved for final stressed syllables.

Rifaat proposes the 'L pitch accent in order to capture instances of "de-accentuation or flattening of stress groups", even though such cases are rare and can be predicted, being observed only before or after an "over-accentuated 'H" (Rifaat 2004:7). In the present model (proposed here in section 3.4.3.1 above) such instances would be expressed as instances of a standard LH* accent produced in compressed pitch range[55]. As noted in section 3.1.1 above, Rifaat's analysis of the pitch movements on 'de-accented' EA words by means of a context-specific 'L accent, could be argued to be consistent with the notion that there *are* pitch accents on these words (and thus the perception of prominence), expressed in such contexts within a very compressed pitch range.

This leaves the 'H pitch accent as the most common accent in Rifaat's model, which he describes as "a default or unmarked accent" (Rifaat 2004:8). This parallels the central finding of this chapter that in EA there is only one pitch accent occurring in non-final positions. Rifaat argues that the default EA pitch accent is a monotonal 'H accent. This is a departure from his earlier description of the accent on pre-final stressed syllables as rising LH (Rifaat 1991); however, in his new model Rifaat argues for a notion of pitch accents as 'peak features'. In this conception the 'H tone has a peak aligned "at the middle of the stress group": in pre-final position the peak shifts rightwards to yield a LH

[54] Rifaat has proposed the same four accent inventory for colloquial EA, in a recent paper (Rifaat 2005).
[55] The phonetics and phonology of EA pitch range manipulation is explored in detail in chapter 8.

accent; in final position the peak shifts leftwards to yield a HL accent. This seems slightly to contradict Rifaat's inclusion of 'HL# as an independent pitch accent in the inventory alongside 'H, and could further be argued to lend support to analysis of all pitch accents (final and non-final) as being of a single phonological specification.

This latter view is adopted by El Zarka (1997:235ff.) in her study of the EA production of MSA. She argues for a single pitch accent occurring in all positions, both final and non-final, based on a survey of her own corpus materials[56]. She argues that the 'only stable element' of the unmarked pitch accent is the H peak itself, which she suggests therefore is the element phonologically associated with the stressed syllable: H*.

Nonetheless El Zarka observes that the EA pitch contour is characterised by falling and rising pitch between successive H* accents, and thus that there must be intervening L targets that must feature in the analysis. She notes that after the H* peak the pitch contour falls steadily over all intervening unstressed syllables, until the next accented syllable, and proposes that this span is a Tonal Domain ("Tondomäne")[57]. In her analysis the L targets between H* pitch accents are neither trailing nor leading tones that form part of the a pitch accent, but rather a L edge tone marking the right edge of the Tonal Domain. The Tonal Domain is not a constituent of the prosodic hierarchy but rather a purely tonally defined object. It consists of the span between accented syllables, but this span is foreshortened by an intervening prosodic boundary (IP, PP) (El Zarka 1997:250). Unaccented syllables after a prosodic boundary are included as pre-accentual syllables in a following Tonal Domain.

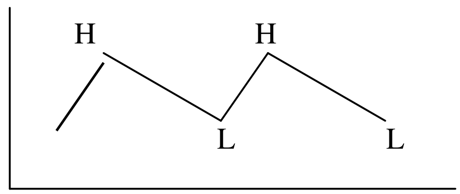
An example of the resulting tonal association is provided in the following example from MSA in (3.18) below (El Zarka 1997:243 example 9.3; stressed syllables <u>underlined</u>):

```
(3.18)  and   where   FUT.-you go   in-Cairo
        wa    ?ay.na  sa-taz.ha.bii fil-qa.hi.ra                MSA
        σ     σ*  σ   σ σ  σ* σ     σ σ σ*  σ
              |         |  |          | |   |
              H         L  H          L H   L
              \_________/  \__________/ \___/    ←tonal domains
        'Where will you go in Cairo?'
```

[56] El Zarka considers but rejects proposal of an L* tone, and argues that instances of low tone on stressed syllables in her corpus are cases of pitch range compression due to final lowering (El Zarka 1997:251-2).
[57] This equates uncontroversially to a parallel tonally-relevant domain proposed in the literature and described variously as the 'stress group' (Bruce 1982) or the 'foot' (Halliday 1967).

El Zarka's analysis results in a surface contour with sharply rising pitch before a stressed syllable, and a gradual pitch slope after the stressed syllable (with slope determined by the number of intervening stressed syllables, unless a prosodic boundary intervenes). This is illustrated in (3.19) below (El Zarka 1997:244 Figure 9.10).

(3.19)



El Zarka accounts for the rise before an initial H* by suggesting that the beginning of the first Tonal Domain in an utterance is characterised by a rise to the first peak, in order to mark the metrical prominence of the phrase-initial syllable. This conception of EA tonal structure predicts the same surface realisation of the pitch contour as observed in the present thesis, and indeed as predicted by the model proposed in section 3.4.1.2 above. El Zarka's H*L H*L sequence in (3.19) above, with a rise to the phrase-initial Tonal Domain, would be analysed in the present model as LH* LH* L-L%.

The final part of El Zarka's model are boundary tones. She notes four possible phrase-/utterance-final contours: rising, level, falling or low-falling which are analysed using a single boundary tone (H%) for the rising contour, and modifications to the basic pitch accent for other contours as shown in (3.20) below (El Zarka 1997:267 Figure 9.25).

(3.20)

H*L H%	rising
H*L H^{*--}	level
H* ↑L	falling[58]
H*L	low falling

3.4.4 Summary

The model proposed here for EA makes the following key claim: that EA has a default pitch accent which is found on all accented syllables. This view is shared by other authors as regards distribution (there is a default accent, at least non-finally for Rifaat) and peak salience (the stable element of the pitch accent that associates to the stressed

[58] The H*L H^{*--} (level) and H*↑L (falling) final combinations may not be contrastive (p.c. El Zarka).

syllable is the H peak). The views diverge regarding representation of intervening L targets between accented syllables (for El Zarka).

The goal of this thesis is not only to establish the facts of EA pitch accent distribution but also to identify the place of EA in the range of typological variation. The model proposed here appeals only to elements of autosegmental-metrical representation that are well-motivated in other languages, and thus facilitates cross-linguistic comparison in the remainder of the thesis.

In his (2004) study, Rifaat argues persuasively that the Arabic intonation system is both structurally and functionally ‘simple’. To capture this salient property of EA, Rifaat suggests a model without recourse to certain elements of standard AM theory (such as boundary tones). The model presented here represents EA intonation with the most minimal subset possible of (arguably) universal elements (one pitch accent, two phrase tones and two boundary tones), thereby respecting the simplicity of the system, whilst facilitating cross-linguistic comparison.

Indeed the claim of this chapter is that the very real simplicity of EA intonation is rooted in two facts: that not only is every content word accented in EA (there is no variation in the distribution of pitch accents across utterances) but each content word also bears a pitch accent with the same phonological specification: a default LH* tone.

3.5 Discussion: intonational typology

The results of the corpus survey yield the generalisation that in EA every PWd bears an intonational pitch accent. It is also striking that almost exclusively, and certainly in pre-nuclear positions, it is the same pitch accent type that is used to mark each word.

Other languages have been reported to share these properties. In a typological survey of 21 languages analysed in AM frameworks, Jun points out that Spanish and Greek have a pre-nuclear (non-final) accent “on almost all content words, and further, that the type of pitch accent is basically the same (L*+H for Greek)” (Jun 2005b). These are exactly the same two co-occurring properties established here to hold of EA.

Other languages for which descriptions report both highly populated pitch accent distribution and a predominant pre-nuclear pitch accent type are Northern European

Portuguese (Vigario & Frota 2003), varieties of Italian (Grice et al 2005), Tamil (Keane 2004) and Danish (Grønnum 1983, Gussenhoven 2004:223ff.).

Jun points out a parallel correlation in other languages between less populated pitch accent distribution and greater variety of pitch accent types (e.g. in English & German). She goes on to note that the formulation of her typological survey cannot capture "differences between stress languages that differ in the frequency and the type of postlexical pitch accent" (Jun 2005b:447).

Surveys of pitch accent distribution do not feature in most descriptions of intonation, and can only be captured in AM notation via insertion of additional pitch accents in notation. A number of approaches to intonational typology have been suggested, and it is an open question where the issue of pitch accent distribution would fit among the categories of variation used in existing conceptions of the range of typological variation.

Ladd (1996:119) for example proposes a four-way characterisation of intonational typological variation, shown in (3.21).

(3.21) Ladd's (1996:119) taxonomy of differences between intonational languages.

1	semantic differences	differences in the meaning or use of phonologically identical tunes
2	systemic differences	differences in the inventory of phonologically distinct tune types, irrespective of semantic differences
3	realisational differences	differences of detail in the phonetic realisation of what may be regarded phonologically as the same tune
4	phonotactic differences	differences in tune-text association and in the permitted structure of tunes

Ladd elaborates on the final category of phonotactic, or *distributional*, differences and notes that this allows for (and predicts) variation in the "permitted phonotactic distribution of an element of the system" (1996:120).

Gussenhoven (2004:275) sets up a comparison of intonational features across three intonational languages (French, English and Bengali). The categories of variation that he includes in this survey of intonational features are listed in (3.22) below. The listing specifically includes a count of the number of pitch accents per phonological phrase, and of the frequency of de-accentuation.

(3.22) Gussenhoven (2004:275) survey of intonational features.

PP = Phonological Phrase; IP = Intonational Phrase

1	number of pitch accents per PP?
2	PP-based readjustment of pitch accents?
3	boundary tones on PP?
4	boundary tones on IP?
5	obligatory IP-final boundary tone?
6	bitonal IP-final boundary tones?
7	number of prenuclear pitch accents
8	number of nuclear contours
9	contour HLH
10	contour LHL
11	frequent deaccentuation

Extreme variation in the density of distribution of pitch accents should be noticeable under either of these surveys. However (by chance) neither survey includes a language that has correlated rich pitch accent distribution and sparse pitch accent inventory, as noted by Jun and as established here for EA .

As Jun notes, whilst there may be a plausible functional explanation for rich pitch accent distribution, there is at present no formal means of capturing this new typological category within the AM framework:

> "In the case.. where pitch accent occurs at a regular interval (i.e. on almost every content word), with a similar type of pitch accent, each of the accents would provide a cue for a word boundary, functioning similarly to the Word boundary tone in Serbo-Croatian of the Accentual Phrase boundary tone in Korean. ... The perceptual equivalence of word segmentation, whether marked by the head tone or by the edge tone of the unit, is not captured in the [AM] model" (Jun 2005b:447).

Jun suggests that rich pitch accent distribution may serve as a perceptual cue at the word-level. This can be achieved by marking the head of the word - with pitch marking the stressed syllable (as in EA) - or by marking word edges - as in Korean or Serbo-Croatian.

The next chapter pursues this suggestion by investigating the properties of EA pitch accents in the context of word-prosodic typology, with the aim of clarifying what types of phonetic cues are used in EA to mark words.

3.6 Conclusion

This chapter has demonstrated the empirical basis of the central claim of this thesis: that EA has very rich pitch accent distribution, with a pitch accent occurring on every content word. This was shown to be true from a survey across a variety of speech styles.

In addition, EA also has the property of marking each accented word with the same pitch accent type. A detailed survey of pitch movements localised around stressed syllables was provided to support this view, as well as a formal model of EA intonation which proposes a single default pitch accent in the EA pitch accent inventory. This model was compared to other analyses of EA within the AM framework.

The correlation of rich pitch accent distribution and use of a single pitch accent type appears not to be unique to EA. The addition of EA to the list of languages which share these two properties suggests that EA is a useful testing ground for Jun's (2005b) suggestion that in such languages pitch may be used as a cue at the word level.

This hypothesis is pursued further in the next chapter which investigates the phonetic correlates of word-level prominence in EA.

4 Word level prominence in Egyptian Arabic

4.0 Outline and aims

Having established in chapter 3 the generalisation that in EA a pitch movement is observed on every Prosodic Word (PWd), the purpose of this chapter is to clarify the status of this word-level pitch marking, in order to identify the position of EA in the spectrum of word prosodic typology.

Section 4.1 sets out the theoretical background to the chapter, exploring categories of word prosody types that have been suggested and the properties associated with them. The experimental studies described in the body of the chapter aim to establish which of these properties hold of EA. Prior work on word level prominence in EA and in other spoken Arabic dialects is also reviewed in this section and results in the working hypothesis that EA is a stress-accent language.

Section 4.2 describes the rationale, methods and results of an experimental study on the alignment of individual low (L) and high (H) pitch targets in the rising pitch movements observed on EA PWds. The target words in the dataset are trisyllabic, with stress falling on the medial syllable (eg [mi'malmil]). If either of the L or H pitch targets align to one edge of the trisyllabic target word, then it might support analysis of the pitch movements observed on every PWd in EA as markers of one or more edges of the PWd, or some other PWd-sized prosodic constituent. Alternatively, if the L or H pitch targets align with the edges of the word-medial stressed syllable, this would support analysis of EA pitch movements as pitch accents associated with the prosodic head of the PWd (the stressed foot). This latter analysis is consistent with the hypothesis that EA is a stress accent language, and is the analysis which the results of the experiment support.

Section 4.3 describes the rationale, methods and results of a small post-hoc experimental study of potential non-tonal acoustic correlates of word level prominence in EA, namely, duration and intensity. These were compared in segmentally parallel stressed vs. unstressed syllables, which were word-initial in two test words from a single target sentence (from the alignment section of the corpus). If these non-tonal acoustic correlates were not significantly different between stressed and unstressed syllables, then EA would be best analysed as a non-stress accent language (Beckman 1986, Ladd 1996). In fact however, the results of the experiment suggest that non-tonal acoustic correlates do vary significantly between stressed and unstressed syllables, and this is consistent with the hypothesis that EA is a stress-accent language, as is widely assumed in the literature.

Section 4.4 discusses the position of EA as a language which does not fit neatly into existing word-prosodic categories. This is because it displays properties at the word-level which are consistent with analysis of EA as a stress accent language, yet, unlike 'archetypal' stress accent languages such as English, these tonal and non-tonal correlates appear systematically on every PWd (as demonstrated in chapter 3). The chapter concludes by arguing that density of pitch accent distribution reflects the need for an additional parameter of prosodic variation at the level of the word.

4.1 Theoretical Background

4.1.1 Assumptions in the literature that EA is a stress-accent language

This chapter seeks to establish the nature of word level prominence in Egyptian Arabic. Superficially, the answer to the question might seem to be obvious, since EA word-stress has been the subject of extensive phonological research, regarding the positional distribution of accent within words (see, inter alia: Harrell 1957, McCarthy 1979, Broselow 1976, Mitchell 1952, Watson 2002).

All of these studies have assumed without controversy that EA is a language which has salient word-level prominence. The studies are based on impressionistic judgements about the position of prominence within EA words, and there is in general only limited discussion of which correlates of word level prominence go together to create the percept of 'stress' in EA (the substance of any such discussion is reviewed in section 4.1.x below). Thus it is widely assumed that EA, and indeed that 'Arabic' in general, is a stress-accent language: as Watson (2002:79) puts it: "Arabic is a language with word stress".

By default, these authors are also assuming that pitch is not used to convey lexical contrasts in EA. This is made explicit in grammars of the language based on standard typological question-based survey techniques such as Gary & Gamal-Eldin (1981), which states that pitch is distinctive in EA only on the utterance level (Gary & Gamal-Eldin 1981:125). Indeed, since stress is assigned 'cyclically' in EA, within the PWd which includes all affixes, there are no examples whatsoever of accentual minimal pairs in EA. In contrast in 'non-cyclic' dialects such as Palestinian Arabic (PA), stress does not shift after affixation which can result in accentual near minimal pairs (Abu-Salim 1983:94) (stressed syllables marked in **bold** type):

(4.1) No stress shift under affixation in EA (compared to PA)[59]:

		EA	PA
'cow'	/baqara(t)/	[**ba**?ara]	[**ba**gara]
'your cow'	/baqara(t)-ak/	[ba**?ar**tak]	[**ba**gartak]

Based on the assumptions in the literature then, the overall aim of this chapter is to test the hypothesis that EA is indeed a stress-accent language, which is thus expected to share properties of word-level prominence with other stress-accent languages such as English.

4.1.2 Word level prominence and intonational typology

The rich pitch accent distribution observed in EA raises the possibility of a further alternative analysis of consistent word-level tonal marking.

In autosegmental-metrical (AM) theory, the surface prosody of an utterance is the result of combining the relevant contributions from different levels of the prosodic hierarchy. So the prosody of an utterance comprises both word-level (lexical) prosody and phrase-level (postlexical) prosody. At each level prosodic marking comprises cues to the head and/or the edges of each constituent at that level[60].

Jun (2005b) points out that consistent tonal marking at the level of the PWd, or an accentual phrase (AP) (a constituent frequently co-extensive with a PWd), is found in a number of languages, but that the origin or function of the word-level tonal cue may vary (Jun 2005b:431):

> "a phrasal tone, which marks a PWd or an AP, can be found in languages with lexical pitch accent (eg Japanese, Serbo-Croatian), stress (eg Chickasaw, Farsi) or with no lexical specification (eg Korean)."

In Jun's typology, there are three possible sources of consistent word-level tonal marking: lexical pitch accent, stress-accent, or Korean style non-accentual tonal

[59] The final [t] of /baqara(t)/ is part of the feminine marker 'ta-marbuta' and is unpronounced when not linked to a following genitive. Both [?] and [g] are common dialectal variants of Classical Arabic [q].
[60] See discussion in chapter 2 section 2.1.

marking. The first of these, lexical pitch accent, we may safely set aside, since there is no indication whatsoever that pitch plays a role in the lexical specification of any morphemes in EA. The second option, that EA is a stress-accent is our current null-hypothesis, as assumed in the wider literature. The third option, of Korean style consistent AP-edge marking, has yet to be ruled out however for EA.

Seoul Korean, the standard dialect of Korean, is a language which is argued to have neither lexical pitch accent nor lexical stress accent (for arguments in favour of a non-stress analysis of Korean see Jun 2005a). The peaks and valleys of the surface pitch contour of an utterance in Korean are thus argued to arise from the distribution of phrase edges. In Jun's (1996, 2005a) analysis of Korean, two prosodic constituents are consistently marked with tonal correlates : the accentual phrase (AP) and intonational phrase (IP)[61]. The AP is of interest for our current purposes because the majority of APs in Korean have been shown to be co-extensive with a single content word, that is a PWd (Schafer & Jun 2002, Jun 2003). An AP may be enlarged to incorporate two PWds if these latter are composed of a small number of syllables, and under the influence of increased speech rate (Jun 2003). The key identifying tonal cue to the AP in Seoul Korean is a double-rise tonal sequence (LHLH) realised across the whole AP (and showing association at both right (initial) and left (final) edges of the phrase, depending on the number of syllables on the phrase) (Jun 2005a:206-7).

There is a key difference then between Korean and a stress-accent language such as English. In a stress-accent language the tonal marking is unambiguously associated with the prosodic head (i.e. the stressed syllable) of the relevant prosodic constituent, whereas, in Korean, tone marks the constituent as a whole and displays association to its edges.

In order to confirm that EA is truly a stress-accent language, as widely assumed, it is necessary to establish the association properties of EA's ubiquitous pitch accents. In AM theory, it is generally assumed that the surface *alignment* of the pitch contour is a reliable indication of the underlying phonological *association* of tones to prosodic targets (c.f section 7.1.1, and Ladd 2003 for a summary and discussion)[62].

---

[61] These compare to the Minor Phrase (MiP) and Intonational Phrase (IP) respectively, in the version of the prosodic hierarchy assumed in the present thesis.
[62] See Xu & Liu (2005) however for a different view.

During auditory transcription the alignment of pitch targets in EA pitch movements was noted to be aligned closely with the edges of the stressed syllable, except in a small number of cases. The purpose of the first experimental investigation described in this chapter (in section 4.2) is to confirm this observation quantitatively by measuring the alignment properties of pitch targets in word-medial stressed syllables.

4.1.3 Word level prominence and word prosodic typology

From the point of view of word-prosodic typology, the key question to ask about any language is whether it is tonal or 'accentual' (see inter alia: McCawley 1978, Hyman 2001, Yip 2002).

These two categories can be clearly distinguished by their 'definitional features', according to Hyman (2001). The definitional feature of a tone language is the fact that the function of pitch in the language is (lexically) distinctive: in tonal languages tone is paradigmatic. In contrast, the definitional property of an accentual language is the fact that the function of pitch in the language is contrastive, marking out a single obligatory syllable as most prominent among the other syllables of the word: accent is syntagmatic. By this definition, given the absence of pitch-related lexical contrasts in the language, it is not difficult to classify EA as an accentual language in Hyman's terms.

Beckman (1986) has however argued persuasively for a further distinction among accentual languages, between 'stress accent' and 'non-stress accent', exemplified in the contrast between languages such as English and Japanese. In Japanese, accent is accompanied by (melodic) pitch features only, whereas in English an accented syllable is optionally marked using pitch, but always displays other (dynamic) correlates of stress: increased duration, increased intensity and more extreme formant values.

In the pair of languages which Beckman studied, non-stress accent occurred in a language (Japanese) in which use of pitch was lexically contrastive. However Ladd (Ladd 1996:155ff.) points out that the phonological parameter "lexical vs. postlexical use of pitch" is logically independent of the phonetic parameter "stress vs. non-stress accent". As a result we see languages in each of the four predicted typological categories shown in (4.2):

(4.2) Parameters in word prosodic typology (Ladd 1996:156 Fig 4.49).

		Phonetic typology	
		stress	*non-stress*
Lexical parameter	*lexical*	Swedish	Japanese
	postlexical	English	Bengali

As discussed already, there is no indication that pitch in EA is lexically contrastive; however, Ladd's typology illustrates that we cannot assume without further investigation whether EA should be classified as a postlexical stress-accent language, with both dynamic and melodic correlates of accent like English, or as a postlexical non-stress-accent language, with melodic correlates of accent only, like Bengali.

The next sections set out what is already known about the phonetic correlates of word level prominence in EA and in other Arabic dialects (section 4.1.4), and then a survey of methods used in other studies to disambiguate melodic and dynamic correlates of word-level prominence (section 4.1.5).

4.1.4 Phonetic correlates of word stress in EA and other Arabic dialects

The patterns of EA word-stress assignment, analysed in seminal papers on the topic (such as: McCarthy 1979, Kenstowicz 1980, Selkirk 1981a, Hayes 1981, Hayes 1995), are based on Mitchell (1960, reprinted in Mitchell 1975:75-98). In that paper, Mitchell states that the main phonetic features which identify word-level prominence in EA are threefold: relative "stress or force" compared to other syllables in the word, higher pitch and the fact that it bears a "kinetic or moving (falling) tone" as compared to the 'static' tones on non-prominent syllables (Mitchell 1975:94 fn2). The examples given are however of words pronounced in citation form which suggests that the 'falling kinetic tone' described may include both word-level and utterance-level tones (both pitch accents and boundary tones), since the word in isolation forms an utterance by itself. Nonetheless Mitchell's description suggests that both increased dynamic 'force' and higher melodic pitch combine to create the percept of word-stress in EA. Watson (2002:79ff.) also reports newly elicited EA word-stress data. She does not discuss which phonetic correlates give rise to the perceptual notion of word-stress in EA but comments unambiguously that "one of the syllables in a content word is perceived as prominent".

A number of authors note that pitch and word-level ‘stress’ features systematically co-occur in EA. Abdalla (1960:18) appears to be the first instrumental study undertaken of EA. He makes it very clear that F0, duration and ‘stress’ are inseparable:

> “stress, quantity and fundamental frequency... function together in such a way that it is inconvenient to discuss any one of them without reference to the other two”.

By ‘stress’ however, Abdalla appears to mean the percept of a more general notion such as prominence, since he uses a notation system involving primary, secondary and minimal stress, which seem to equate to nuclear, phrasal and word-level stress. He demonstrates that durations of stressed vowels are larger both under higher degrees of stress and in phrase/utterance-final position (ibid. p21).

In their reference grammar of EA, Gary & GamalEldin (1981:125) similarly state that there is a “predictable relation” between high pitch and ‘primary stress’ in EA. Again however, they state this for “polysyllabic word utterances or phrase utterances” so that their generalisation may confound word-level and phrase-level correlates. Similarly in an instrumental study, Rifaat (1991) notes that all stressed syllables in EA are systematically associated with higher F0 than unstressed syllables, but does not mention whether non-tonal correlates also mark this distinction.

In a small instrumental study, Guindy (1988:44-46) found that the stressed syllable in trisyllabic words was marked most often by both the highest F0 peak in the word and also the highest amplitude in the word, but that there were often mismatches, with highest intensity on a different syllable than the one bearing highest F0. In another instrumental study, El Zarka (1997:106-7) observed F0 be the primary correlate of word-level prominence, with intensity as a secondary correlate, and vowel duration also observed in some cases. Another potential non-tonal phonetic correlate of prominence at the word level is noted by Gary & GamalEldin (1981:125) who report consonantal strengthening in stressed syllables vs. unstressed syllables as a cue to word-level prominence: the initial consonant of a stressed syllable is more fortis than the initial consonant of an unstressed syllable.

Turning to other dialects of Arabic, Mitchell suggests that in most dialects “the accented syllable is also marked.. by stress or the expenditure of greater force or energy on the accented syllable" but that these dynamic correlates are "less noticeable" in the dialects of Syria and Kuwait (ibid.). In a later section Mitchell goes further and suggests that in Moroccan Arabic (MA) prominence is based mostly on pitch. There is variation on this topic however in more recent instrumental studies on MA: Boudlal (2001:105ff.) cites two studies which found duration and F0 to be reliable correlates of stress in MA (Hammoumi 1988, Nejmi 1993), whereas his own instrumental study suggests that it is indeed F0 which is the most consistent correlate of word level prominence in MA. It seems therefore that Arabic dialects may vary in whether dynamic cues to word-level prominence are used alongside melodic cues (F0).

A feature of many instrumental studies on Arabic dialects however is difficulty in disambiguating whether increased F0 is a word-level or phrase-level cue to prominence (cf. discussion of this potential confound in: Beckman & Edwards 1994, Vanderslice & Ladefoged 1972).

DeJong & Zawaydeh (1999) investigated duration and vowel formant values as potential correlates of word-level prominence in Ammani Arabic (henceforth Jordanian Arabic, JA). They found that duration was a direct correlate of word stress: syllables were longer if stressed than unstressed[63]. They also found that F0 was higher in stressed syllables than unstressed syllables but were not able to fully disambiguate whether increased F0 was a word-level or phrase-level cue. They opt for the analysis that in JA the observed variation in F0 is a phrase-level effect, on the assumption that pitch accents on stressed syllables are optional in that dialect (ibid. p20).

In another production study of JA (de Jong & Zawaydeh 2002), the same authors compared duration, F1 and F0 in target words either bearing contrastive focus ('lexical focus' in their terms) or falling after a contrastive focus. If JA is like English in conditioning de-accenting of items after a contrastive focus, then this is again a comparison of accented vs. unaccented syllables (as opposed to stressed vs. unstressed). The authors do not report whether or not the post-focal words were associated with pitch movements; however, it is probable that the words were unaccented since the authors do state that the choice to compare words in these two conditions conflates the effects of nucleus placement and focus (ibid. p60).

---

[63] They also observed some effects of word-level prominence on F1 values in [a] vowels which were higher in stressed syllables than in unstressed syllables.

In a study on Lebanese Arabic (LA) Chahal (2001) tackled potential confounding of acoustic correlates at different levels of prominence by classifying her data into different levels before analysis. She elicited broad focus and narrow focus utterances (each containing three target words) and made an auditory transcription which permitted each target word to be classified as either nuclear accented (IP-level prominence), accented (iP-level prominence[64]) or unaccented (word-level prominence). The overwhelming majority of unaccented targets were found in post-focal position, indicating that post-focal 'de-accenting' occurs in LA.

Chahal compared the F0, syllable duration, intensity (RMS) and format values (F1 & F2) of targets across the three prominence levels. She found a significant difference in F0, duration and amplitude between levels, with each of the correlates significantly increased between an IP-level prominence and an iP-level prominence, and, in turn between an iP-level prominence and a word-level prominence[65]. On the basis of these findings Chahal argues that LA is a stress-accent language akin to English in using both tonal and non-tonal cues to prominence. Since Chahal did not directly compare stressed-but-unaccented syllables with unstressed syllables her findings in fact tell us that both tonal and non-tonal correlates are used to mark phrase-level prominences in LA, but the acoustic correlates of word-level prominence itself are not established.

In an instrumental study across Arabic dialects, Al-Ani (1992) compared adjacent syllables in a series of 'construct state' (*iDaafa*) phrases, as pronounced medially in a sequence of repetitions of the same phrase by a group of speakers from 4 different Arabic dialects (speakers from Sudan, Saudi Arabia, Morocco and Iraq). He measured F0 and amplitude at the steady state of the vowel of each syllable in the phrase, as well as the duration of each syllable. He found that, on average across all speakers, stressed syllables had higher amplitude and duration than unstressed syllables, but that F0 did not vary greatly between stressed/unstressed syllables. However, the choice to measure F0 at the steady state of each vowel, rather than pitch maxima or minima, means that differences in pitch accent choice or alignment may have obscured some actual variation between stressed/unstressed syllables, and in addition, the detail of his results suggests that F0 varied from syllable to syllable for some speakers more than others.

[64] This equates to MaP level prominence in the version of the prosodic hierarchy adopted in this thesis; but see Gussenhoven (2004:166-7) for a different view.
[65] The difference in F0 between levels was relatively small in neutral contexts, in part due to pitch accent type employed by some speakers.

Al-Ani analyses these phonetic correlates as being of phrasal stress, which he argues is a reflex of syllable type (heavy/light etc). However, as Chahal (2001:139-40) points out, his results can be re-interpreted in terms of different levels of prominence. Under this view, Al-Ani's study suggests that there are non-tonal correlates of both word-level and phrase-level prominence in the dialects that he studied.

In summary then, the consensus of the descriptive and instrumental literature on EA suggests that both tonal and non-tonal correlates of prominence are found at the word-level, as well as at higher levels of the prosodic hierarchy. This situation is also found, at higher levels of prominence at least, in other Arabic dialects such as JA and LA, and potentially at all levels in other dialects (in Al-Ani's comparative study). There are nonetheless indications that some dialects, such as MA, may not employ non-tonal cues to word-level prominence.

From comparison with other dialects, it cannot be assumed *a priori* that EA uses both tonal and non-tonal correlates, even though the descriptive literature on EA suggests that the correct hypothesis is that both types of correlates will be observed. An instrumental investigation is needed to resolve this issue. The next section (4.1.5) explores methodologies used to investigate correlates of word-level prominence in studies on other languages, in order to identify a suitable methodology for use in EA.

4.1.5 Methods used to investigate stress vs. non-stress accent

As seen in section 4.1.3 above, Beckman (1986) argued persuasively for distinctions among accentual languages according to which of a 'hierarchy' of phonetic correlates are employed in a particular language to mark word-level prominence.

Beckman's claims were based (in part) on a production study in which she compared the acoustic correlates of accented and unaccented syllables within a word, using accentual minimal pairs which differed only in the position or presence of lexical accent. For example in English the ratios of duration, intensity and F0 in stressed vs. unstressed syllables were compared across minimal pairs such as 'pérmit'/'permít'. Likewise in Japanese comparison was made between pairs which differed in the position of lexical accent: 'kamé'/'káme', or in the presence vs. absence of accent: 'ikén'/'iken' (Beckman 1986:146-7). In the English accentual minimal pairs, Beckman found a significant difference in the amplitude and duration of stressed vs. unstressed syllables, as well as in their pitch. In Japanese however only F0 differed significantly across all speakers in accented vs. unaccented condition (there was a small effect of duration for some speakers).

In EA, the position of word-level stress prominence is 100% predictable from the syllabic structure of a word. In addition, as outlined above, since in EA stress is assigned within the PWd which includes all affixes (i.e. 'cyclically'), there are no accentual minimal pairs in EA (cf. example (4.1) above). For this reason Beckman's methodology cannot be reproduced in EA. A study which uses similar methodology to Beckman (1986), comparing stressed vs. unstressed syllables, but in an experimental design that does not rely on the existence of accentual minimal pairs in the language, is Keane (2004).

Keane investigates the acoustic correlates of word-level prominence in Tamil, a language in which, like EA, a pitch movement is observed on every content word in non-phrase-final position. She created stimuli which permitted comparison of segmentally parallel target syllables occurring in initial, medial and final position in trisyllabic words. Since the position of word-level prominence is thought to be fixed in Tamil, in initial position, this provides for comparison of phonetic correlates in stressed vs. unstressed syllables.

Keane extracted a number of dependent variables for comparison including F0, intensity (loudness) and syllable duration. She found that there was no variation in intensity across different positions in the word. In contrast, syllable duration was greater in word-final position, due to proximity to the right edge of the word, but there was no significant difference in syllable duration between initial and medial target syllables. F0 was however significantly higher in initial position. Keane offers two possible analyses of her findings. Firstly, assuming that Tamil has initial word-level prominence, the results suggest that it is marked with tonal correlates only, as in Japanese, with no accompanying non-tonal accentual correlates. Alternatively, the correlates may not indicate initial word-level prominence but rather a word-edge tonal marker at the left edge of the word. It is not clear however in what contexts these two interpretations could be distinguished empirically. Nonetheless, comparison of target syllables in segmentally parallel words can provide reliable evidence regarding the presence or absence of non-tonal correlates of word-level prominence, and this method could therefore be applied successfully to EA.

The primary goal of the experimental studies outlined in this thesis were to establish patterns of pitch accent distribution and establish their interaction with other aspects of EA grammar. The question of whether non-tonal correlates are used in addition to tonal correlates to mark word-level prominence was not anticipated during design of the experiments, in part under the influence of the widely held assumption that 'Arabic' is a stress-accent language like English and that all Arabic dialects will have similar prosody above the level of the word. As we have seen this has been demonstrated to some degree for other dialects such as JA and LA.

Nonetheless EA is sufficiently dissimilar to JA and LA in its pitch accent distribution[66], that the findings for JA and LA regarding correlates of word-level prominence cannot be assumed *a priori* to hold of EA. A small post-hoc analysis of suitable targets, in data collected for other purposes, was carried out in order to provide some preliminary answers to the question of whether or not non-tonal cues mark word-level prominence also in EA. The methodology and results of that small study are described in section 4.3 below.

4.2 Word-level prominence in EA: edge marking or accentual marking?

4.2.1 Rationale

The experiment described in this section explores the possibility that the pitch movement observed on every PWd in EA is not an accentual prominence-related pitch accent, but instead a word-level boundary tone of some sort, marking either the edges(s) of the PWd, or its overall domain. As discussed in section 4.1.2, word-level pitch movements in Korean have been argued to be the reflex of a tonal sequence marking the whole accentual phrase, and associating with its edges.

The specific research question addressed here therefore is whether EA word-level pitch movements are aligned with the word edges or with the accentual head of the word (the stressed syllable). The latter outcome would be consistent with the hypothesis that EA is

[66] In the studies outlined above, both JA and LA showed post-focal de-accenting, which does not occur in EA (see section 3.3.1).

a stress accent language. The question can be explored straightforwardly by examining the alignment properties of pitch movements in stress-medial words, as outlined below.

4.2.2 Methodology

This investigation uses a subset of the 'alignment' corpus collected for use in chapter 7. The full corpus comprises 24 target sentences, each of which contains a target word selected for the segmental and syllabic properties and position of its stressed syllable. The subset of 6 target sentences used here have trisyllabic target words in which a medial closed syllable bears lexical stress (eg [mi'malmil]). The target word occurs as either the second or third word in the carrier sentence. The full set of target words is given in (4.3) below (and the full set of carrier sentences is given in Appendix B B.1).

(4.3) Target words with word-medial stress (stressed syllables are underlined).

id code	*target word*	*gloss*
212119	mi'malmil	nervous
212120	mi'namrad	rebellious
212121	mi'mangih	boastful
212122	mit'manZar	showing-off
212123	mit'namnim	cute/tiny
212124	mu'namnim	cute/tiny

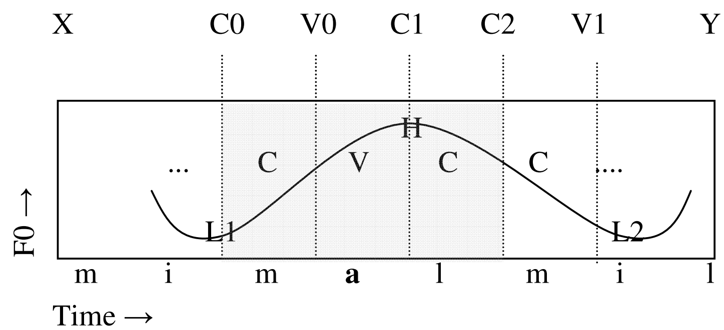
Each of the six target sentences was read three times by 15 speakers of CA (15 speakers x 3 repetitions x 6 targets = 270 tokens). Auditory transcriptions and labelling were carried out by the author with reference to spectrogram and F0 contour extracted using Praat 4.2 (Boersma & Weenink 2004). There were 19 productions which contained a disfluency on or near the target word, and 21 productions in which a phrase boundary was inserted immediately before or after the target word; these 40 were excluded from quantitative analysis since alignment properties could be affected by disfluency, or by proximity to a phrase boundary (Prieto et al 1995, Chahal 2001).

In order to establish the alignment properties of high (H) and low (L) pitch targets in EA pitch movements, following the methodology of Atterer & Ladd (2004), pitch events and segmental landmarks in each target word, as listed in (4.4) below were labelled by hand in each of the 238 tokens (as illustrated in Figure 4.1).

(4.4) Labelling of pitch events and segmental landmarks.

C0	start of first consonant of stressed syllable
V0	start of stressed vowel of stressed syllable
C1	end of stressed vowel of stressed syllable
C2	start of second consonant of intersyllabic cluster
V1	start of vowel of following syllable
L1	valley pitch turning point before peak of test syllable
L2	valley pitch turning point after peak of test syllable
H	peak pitch turning point of test syllable
X	left edge of word
Y	right edge of word

Figure 4.1 Schematised labelling diagram of pitch events and segmental landmarks. (The stressed vowel is marked in **bold** type.)



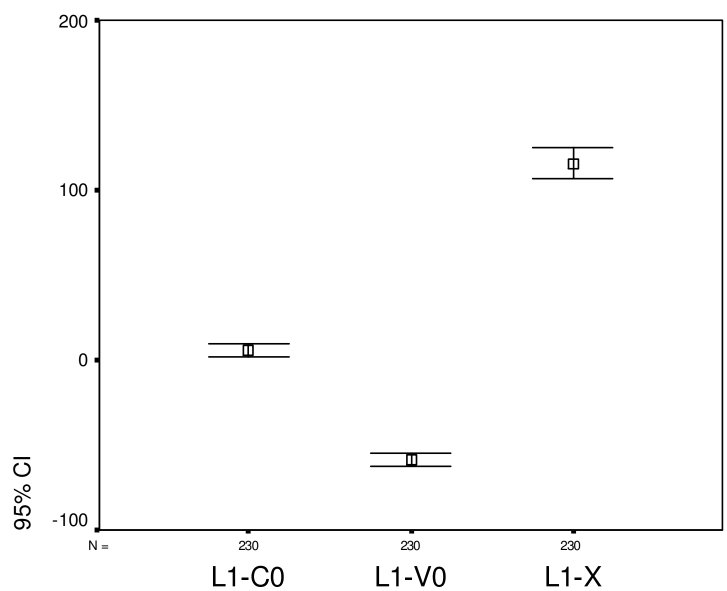
Alignment of the start of the pitch rise is assessed by calculating the position of L1 relative to C0 and V0, that is, by calculating L1-C0, and L1-V0 in milliseconds. Alignment of the pitch peak (H) is assessed by calculating the position of H relative to C1, C2 and V1, that is, H-C1, H-C1 and H-V1. A negative value indicates alignment of a pitch event before the relevant segmental landmark. During labelling it became apparent that the position of L2 (the valley pitch turning point after peak of test syllable) fell consistently in or at the beginning of the following word in the test sentence. This suggests that the pitch movement associated with each PWd in EA is a rising pitch movement (an LH sequence; cf. discussion in 3.4.3), since L2 in fact relates to the following word. Variables related to L2 alignment were therefore not included in the present investigation.

4.2.3 Results

The aim of the investigation of pitch accent alignment in word-medial stressed syllables was to determine whether the individual pitch targets and/or the whole rising pitch movement show alignment to a) the stressed syllable or b) the word edge.

Looking first at the position of L1 (the valley turning point before the pitch peak) there are three relevant variables to consider: L1-C0, the distance from L1 to the onset of initial consonant of the stressed syllable; L1-V0, the distance from L1 to the onset of the stressed vowel; and L1-X, the distance from the L1 turning point to the onset of the initial consonant of the target word (i.e. the left edge of the word). Figure 4.2 shows mean values for each of these variables in milliseconds. A value close to zero indicates that the pitch event is aligned closely to the segmental landmark in question. The graph indicates that in general L1 is aligned slightly closer to C0 than to V0, but that alignment of L1 is very much closer to C0/V0 than to X, and thus that the start of the rise in pitch clearly coincides with the left edge of the stressed syllable, rather than with the left edge of the word.

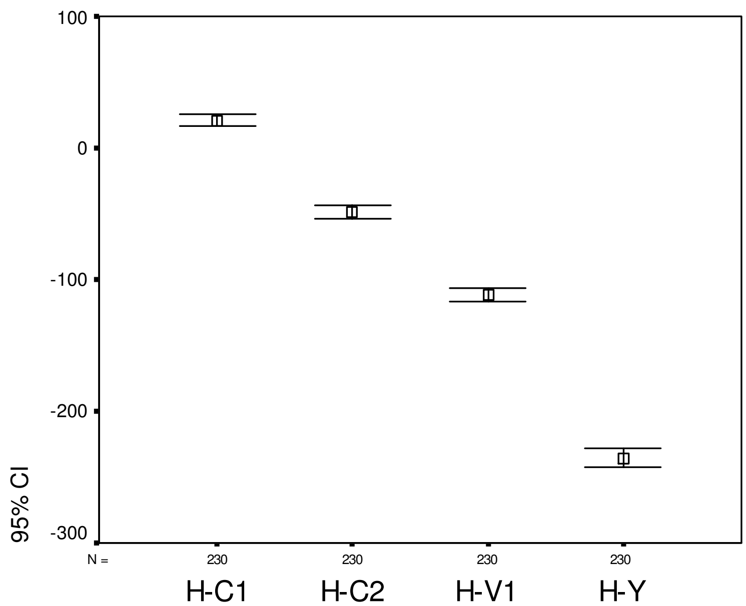
Figure 4.2 95% confidence intervals around mean values of L1 alignment variables.
L1-C0 = L valley to onset consonant of stressed syllable
L1-V0 = L valley to start of stressed vowel
L1-X = L valley to start of word N=230



Turning to the position of the H (peak) target, here there are four variables to consider: H-C1, the distance from H to the onset of the coda consonant; H-C2, the distance from H to the onset of the second consonant in the cluster at the syllable boundary; H-V1, the distance from H to the onset of the vowel of the postaccentual syllable; and H-Y, the distance from H to the offset of the final consonant in the target word (i.e. to the right edge of the word).

Figure 4.3 reports mean values for each of these variables in milliseconds, and again, a value close to zero is an indication that the H pitch event is aligned closely to the segmental landmark in question. It is clear that the pitch peak is very much closer to the segmental landmarks at the edge of the stressed syllable (C1/C2) than to Y, and thus that the end of the rising pitch movement coincides with the end of the stressed syllable, rather than with the right edge of the target word.

Figure 4.3 95% confidence intervals around mean values of H alignment variables.
H-C1 = H peak to end of stressed vowel
H-C2 = H peak to start of second consonant in intersyllabic cluster
H-V1 = H peak to start of vowel of following syllable
H-Y = H peak to end of word N=230



Another way to ascertain whether the rising pitch movement is associated with the stressed syllable rather than with the word is to determine whether there is any correlation between the duration of the pitch rise and the duration of the stressed syllable and/or the duration of the word. The relevant variables are as follows (all calculated in milliseconds): rise duration (H - L); stressed syllable duration (C2 - C0); word duration (Y - X). Comparison amongst these variables reveals that there is a weak correlation between rise duration and stressed syllable duration ($R = 0.295$; $p < 0.01$), but considerably weaker correlation between rise duration and word duration ($R = 0.138$; $p = 0.036$). This is illustrated in Figures 4.4 and 4.5 below, which plot rise duration against syllable duration and word duration respectively; a best fit linear regression line indicates the degree of correlation.

The relationship between the duration of the pitch movement and the duration of the stressed syllable is stronger than the relationship between the duration of the pitch movement and the duration of the word.

Figure 4.4 Scatter plot: rise duration x stressed syllable duration.
The graph indicates the degree of correlation between the duration of the F0 rise and the duration of the *stressed syllable*.

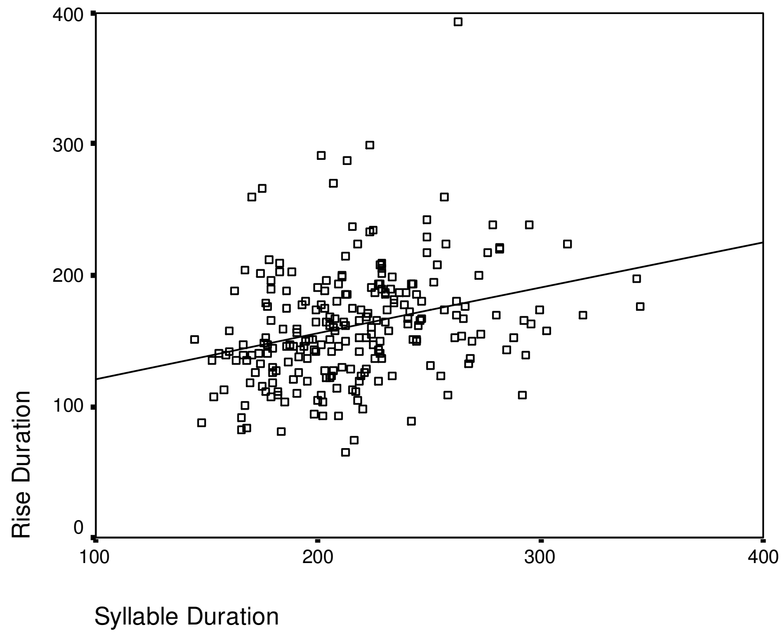
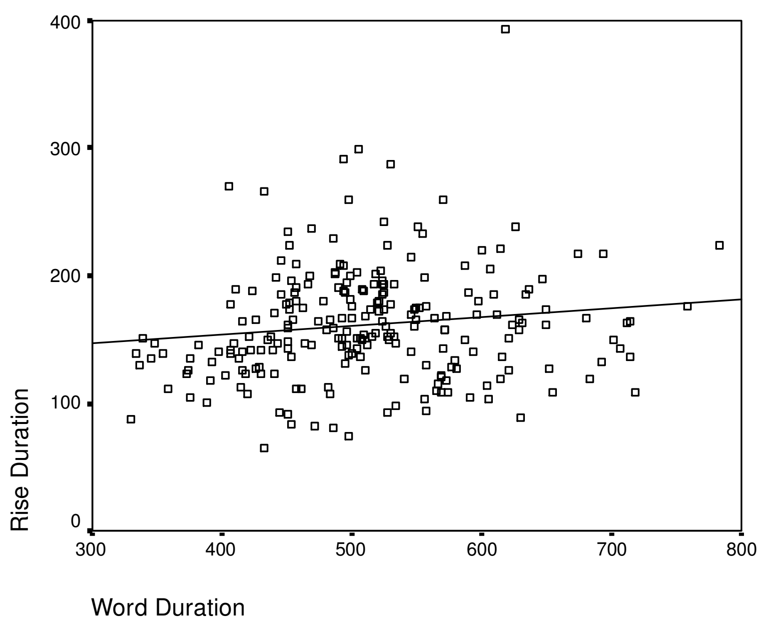


Figure 4.5 Scatter plot: rise duration x word duration.
The graph indicates the degree of correlation between the duration of the F0 rise and the duration of the *word*.



The results of the investigation of pitch accent alignment in word-medial stressed syllables strongly supports the hypothesis that in EA the rising pitch movement in pre-nuclear stressed syllables is phonologically associated with the stressed syllable of the word, rather than with the edges of the word itself. This is true for both the L and the H pitch targets which comprise the rising movement: the L target is much more closely aligned to the left edge of the stressed syllable than to the left edge of the word; similarly, the H target is much more closely aligned to the coda consonant at the right edge of the stressed syllable, than to the right edge of the word.

The hypothesis that EA is an accentual language in which word-level prominence associates to the stressed syllable (prosodic head) of the word is thus supported.

4.3 Investigating the phonetic correlates of word-level prominence in EA

4.3.1 Rationale

Section 4.2 provides evidence to support classification of EA as an accentual language, in which pitch movements at the level of the word are associated with the stressed syllable, or metrical head, of the PWd rather than with the whole of the word or with one of its edges. The next question to resolve is what type of accentual language EA is, within the stress vs. non-stress accent typology suggested by Beckman (1986). As discussed in section 4.1.4 above, the hypothesis, following the assumptions of the descriptive literature, is that EA is a stress-accent language in which both tonal and non-tonal correlates of accentual prominence are used, as opposed to a 'non-stress accent' language, in which only tonal correlates of accentual prominence are found (as in Japanese).

As outlined in section 4.1.5, it is not possible in EA to reproduce Beckman's (1986) methodology identically, since there are no accentual minimal pairs in EA. Nor is it possible to reproduce Chahal's (2001) methodology, which used auditory transcription to identify accented vs. unaccented exemplars of parallel word tokens for comparison from a large corpus; transcription of the large EA corpus studied here found a very small number of accentable PWds which were observed to be unaccented, and these are not in sufficient numbers to permit meaningful comparison. The closest methodology that can be reproduced in EA is that of Keane (2004), which compared segmentally parallel syllables in different positions in words.

Although it would be preferable to design and implement a full study to investigate the non-tonal correlates of EA word-level prominence, this was not possible within the scope of the present study. However after a survey of the speech materials already collected a small set of suitable targets were identified, within the alignment section of the corpus. The properties of these targets and the methods used to obtain measurements of non-tonal acoustic correlates of prominence are described in the next section.

4.3.2 Methodology

In order to directly establish the correlates of word-level prominence, pairs of segmentally parallel syllables are needed, occurring occur in accented vs. unaccented (that is, stressed vs. unstressed) exemplars, within the corpus of speech materials already collected. In addition, since a number of the prior studies on non-tonal correlates of prominence found an overlap of durational effects in word-final syllables, such that syllables were longer if stressed and also if word-final (de Jong & Zawaydeh 1999, Keane 2004), it is important to compare stressed vs. unstressed syllables in non-final position in the word.

Although a survey was made of the whole speech corpus, the most likely candidates for the present purposes all occurred within the alignment corpus, and unfortunately only one such pair was found in which both potential test syllables were non-word-final.

The test syllables chosen for investigation here were word-initial, and occurred in different words within the same test sentence (test syllables underlined; stressed syllables in **bold** type):

(4.5)

bu??	**mu**na	mu**nam**nim	xaaliS	wa	ša9riha	Tawiil
mouth	Muna	tiny/cute	completely	and	hair-her	long

"Muna's mouth is tiny and her hair is long".

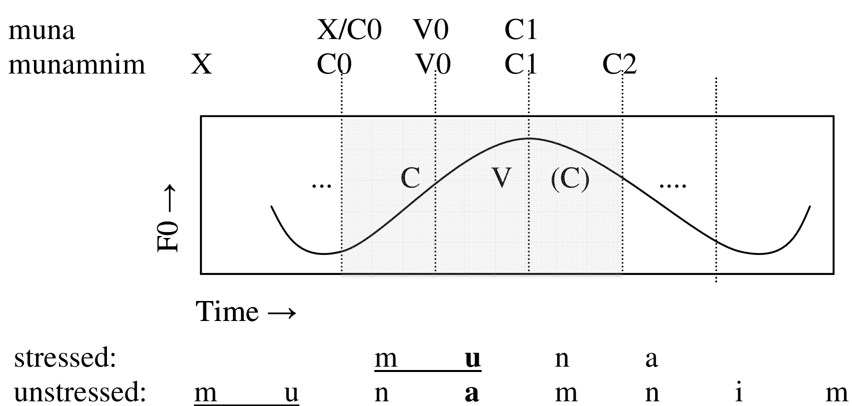
The test syllable was /mu/ in each case, occurring in word-initial position, and either stressed (in ['muna]) or unstressed (in [mu'namnim]). In fact, the initial syllable of /munamnim/ was produced by most speakers with a reduced vowel, [i]: [mi'namnim]. This is perhaps to be expected, since unstressed vowels are routinely reduced in EA, but could also mean that the target syllables are not suitable for direct comparison, if in fact the reduced vowel appears in the underlying lexical form of such words (p.c. El Zarka). In the absence of more suitable targets, investigation of other potential non-tonal correlates of word-level prominence in these targets (besides vowel quality) was nonetheless pursued.

The number of tokens available for analysis was 38 for stressed [mu] in ['muna] and 45 for unstressed [mu] in [mu'namnim] (this was due to insertion of an internal phrase boundary after the word [muna] in some tokens by some speakers, which could have resulted in lengthening of the final syllable of [muna] and/or of the whole test word). The segmental landmarks of the test syllables had already been labelled for other investigations (for section 4.2 and for chapter 7) as outlined in (4.4) above; the relevant labels are reproduced here in (4.6), and are illustrated in Figure 4.6.

(4.6) Labelling of segmental landmarks for investigation of non-tonal cues.

C0	start of first consonant of stressed syllable
V0	start of stressed vowel of stressed syllable
C1	end of stressed vowel of stressed syllable
C2	start of second consonant of intersyllabic cluster (if applicable)
X	left edge of word

Figure 4.6 Schematised labelling diagram of segmental landmarks for investigation of non-tonal cues. (Test syllables underlined; stressed syllables in **bold**.)



The test syllable in each case was the first syllable (S1) of the test word. The duration of S1 (S1DUR) was calculated from the position of labelled segmental landmarks. The mid point of s1 (S1MID) was identified by dividing s1durms by two and adding the result to the value of the landmark variable at the start of S1 in each case (C0 and X respectively). These calculations are shown in (4.7) below.

(4.7)	*test word*	*S1DUR calculation*	*S1MID calculation*
	MUna	S1DUR = C1-C0	(S1DUR/2) + C0
	muNAMnim	S1DUR = C0-X	(S1DUR/2) + X

Using Praat 4.2, measurements of F0 and intensity were extracted at the midpoint of the test syllable in each token, according to the following parameters shown in (4.8). All of the speech recordings were made using a head-mounted microphone, positioned at a uniform distance from the speaker's mouth throughout each recording session, in order to permit extraction of reliable intensity measurements.

(4.8) Parameters used to extract F0 and intensity in test syllables.

variable	*code*	*position:*	*unit:*	*settings (Praat 4.2 defaults):*
F0	S1MIDF0	S1MID	semitones	pitch range 75-600Hz linear interpolation
intensity	S1MIDDB	S1MID	decibels	minimum pitch 100Hz cubic interpolation reference level = auditory threshold pressure

The key variables for comparison between stressed and unstressed conditions in the test syllables are: syllable duration in milliseconds (S1DUR); pitch at the midpoint of the test syllable (S1MIDF0) and intensity at the midpoint of the test syllable (S1MIDDB).

The null hypothesis, that EA is a stress-accent language, predicts the values of all variables to be higher in the stressed condition (S1 in [MUna]) than in the unstressed condition (S1 in [muNAMnim]). Under the alternative hypothesis, that EA is a non-stress-accent language which employs only tonal correlates of accentual prominence, only the values of the melodic variable pitch (S1MIDF0) are predicted to be higher in the stressed condition; the dynamic variables of intensity and duration are predicted not to vary significantly between stressed and unstressed conditions.

4.3.3 Results: acoustic correlates of word level prominence in EA

Descriptive results for the five main test variables, in stressed vs. unstressed condition, are illustrated in Figures 4.7-4.9 below. Bar charts illustrating mean values of each variable per speaker are provided in Appendix B (B.2-B4). For all of the variables the mean value across all speakers is higher in stressed condition than unstressed condition: this is true of duration, F0 and intensity. Paired sample t-tests comparing stressed-unstressed pairs within each token confirm that all of these differences are highly significant at the 0.01 level (see (4.9) below). In summary then this small experimental investigation suggests that EA does employ both tonal and non-tonal cues to word level prominence. This supports the hypothesis, widely assumed in the literature, that EA is a stress accent, rather than non-stress accent, language.

Figure 4.7 Bar chart: mean duration (in milliseconds) in stressed vs. unstressed S1.

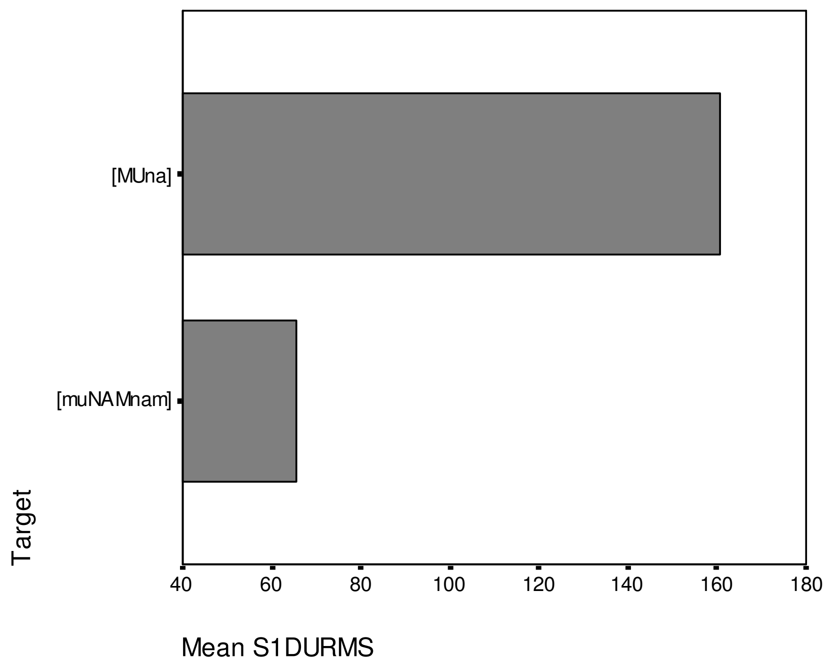


Figure 4.8 Bar chart: mean F0 (in semitones) in stressed vs. unstressed S1.

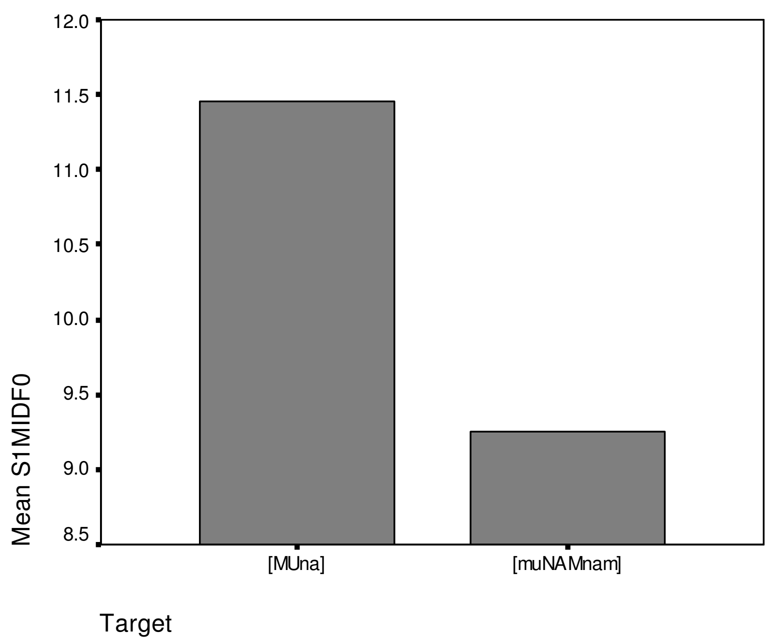
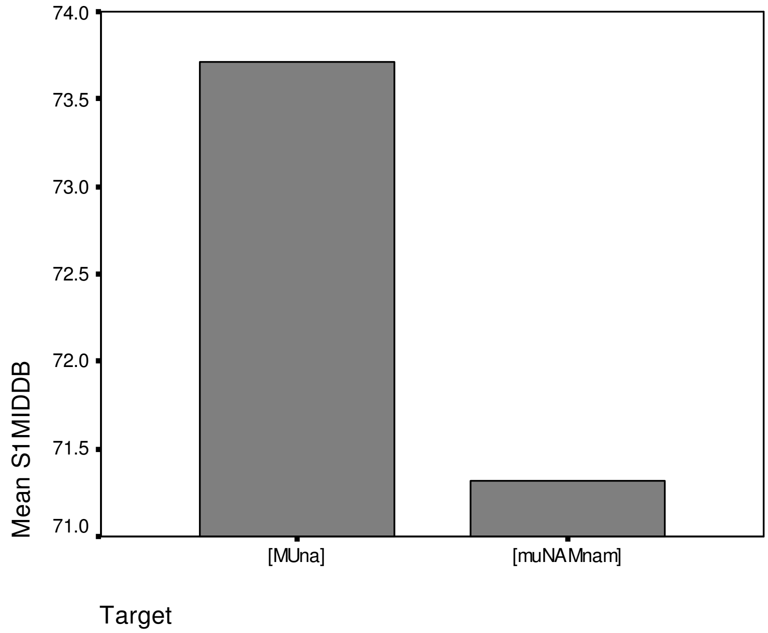


Figure 4.9 Bar chart: mean intensity (in decibels) in stressed vs. unstressed S1.



(4.9) Paired sample t-tests: stressed S1 (SS1) & unstressed S1 (US1) within tokens.

	Paired Differences					t	df	Sig. (2-tailed)
	Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of Difference: Lower	95% Confidence Interval of Difference: Upper			
SS1DURMS US1DURMS	100.948	27.002	4.439	91.945	109.951	22.740	36	.000**
SS1MIDF0 US1MIDF0	1.243	.885	.145	.947	1.538	8.537	36	.000**
SS1MIDDB US1MIDDB	1.879	1.643	.270	1.331	2.427	6.956	36	.000**

** denotes a result which is significant at $\alpha = 0.01$

4.4 Discussion: word-prosodic typology

The results of the experimental investigations described in this chapter, in sections 4.2 and 4.3, confirm the widely assumed hypothesis that EA is a stress-accent language: cues to word-level prominence in EA are closely aligned to the stressed syllable of the word, and these cues are both tonal and non-tonal (non-tonal cues investigated here for EA were stressed syllable duration and intensity).

In terms of the four-way typology proposed by Ladd (1996:155ff.) (see section 4.1.3 above), EA can be classified as a language displaying postlexical stress-accent, and is thus similar to other stress-accent languages like English. However, as discussed in chapter 3, the distribution of pitch accents observed in EA is different from the distribution of pitch accents observed in English.

I would like to suggest that, just as the phonetic parameter of stress vs. non-stress accent is logically independent from the parameter of lexical function of word-level prominence (lexical vs. postlexical), so also the choice of domain within which postlexical pitch is realised is logically independent from these other parameters.

The idea that relative prominence at different levels of the prosodic hierarchy can be marked in different ways is already familiar: we are used to the fact that in English word-level prominence is marked with dynamic correlates (such as duration and intensity) whilst phrase-level prominence is marked with pitch (F0). We also know that use of pitch in English is postlexical. However, the fact that use of pitch in English is postlexical does not necessarily imply that all postlexical use of pitch will be at the phrase-level. EA is precisely what we would expect to see if postlexical use of pitch could also be at the word-level.

To explore this idea further, let us look in turn at two ways of viewing this proposed new parameter of variation. Looking first at what happens at the word-level only, then the difference between English and EA is that: i) in English postlexical tone at the word-level is 'optional' and non-tonal cues are obligatory (only words that are also prominent at the phrase level are marked with pitch); whereas, ii) in EA both types of cues are obligatory at the word-level.

If this is the correct way of thinking, and if obligatory/optional use of pitch at the word level is indeed logically independent from the function of pitch, then we should expect to find languages in which *lexical* use of pitch patterns as it does in English and EA.

A lexical pitch accent language such as Swedish has obligatory lexical tonal cues to word-level prominence, but the new parameter predicts languages in which pitch with lexical function is not always realised at the word level, resulting in restrictions on the surface realisation of lexical tones or accents. One could argue that this is in fact exactly what happens in a language like Japanese in which only one lexical accent per accentual phrase (AP, =MiP) is realised (Gussenhoven 2004)[67]. This way of thinking about the proposed new parameter of variation could be formulated as in (4.10) below[68].

(4.10) Parameters in word prosodic typology.

		Distribution of phonetic correlates	
		stress &/or pitch on every word pitch on some words only	stress &/or pitch on every word pitch on every word
Lexical function	lexical	Japanese	Swedish
	postlexical	English	EA

Turning now to think only about languages in which the function of pitch is postlexical, one could state the proposed new parameter as variation in the *domain* within which postlexical tones are distributed. This involves careful use of the term ‘postlexical’. The term ‘lexical’ means ‘originating in the lexicon’, and since most lexical properties are properties of words, it is common to use the word ‘lexical’ as a synonym for a ‘word-level’ property. My suggestion is that the term ‘postlexical’ should only mean ‘not originating in the lexicon’: EA is what we expect to see if a phenomenon whose origin is not in the lexicon is nonetheless a word-level property.

If this is the right way of thinking, then variation in the domain of distribution of pitch accents should interact freely with the parameter of phonetic typology (stress vs. non-stress correlates). We expect to see languages in which postlexical stress accents may be sparsely distributed, such as English, and richly distributed, such as EA. In addition, we should also expect to see languages in which postlexical *non-stress* accents show differences in surface distribution. This is indeed what we find: in Bengali postlexical *non-stress* accents are sparsely distributed (Hayes & Lahiri 1991), and, as we have seen above, postlexical *non-stress* accents in Tamil are richly distributed (Keane 2004). This way of thinking about the proposed new parameter of variation could be formulated as in (4.11) below.

---

[67] Compare also Shanghai Chinese, in which only one lexical tone per word survives (Yip 2002:111ff., 185ff.).

[68] Note also that Japanese shares with Basque the property of restricting the number of lexical accents that may be realised within a phrase (only one pitch accent may be realised per AP (Beckman & Pierrehumbert 1986). Thus the choice regarding the domain within which pitch is realised is also logically independent of whether word-level prominence is marked with stress or non-stress cues.

(4.11) Parameters in postlexical use of pitch.

		Phonetic typology	
		stress + pitch	*pitch only (non-stress)*
Domain of distribution	*phrase-level*	English	Bengali
	word-level	EA	Tamil

To recap therefore, I propose that it is important to separate out the notion of lexical vs. postlexical use of pitch from the notion of the *domain* within which pitch is distributed, and that this separation yields a new parameter of prosodic variation which makes valid typological predictions.

4.5 Conclusion

This chapter explores the nature of word-level of prominence in EA, in the context of both intonational typology (marking the heads or edges of prosodic constituents) and word-prosodic typology (use of tonal and/or non-tonal correlates of prominence). The widely-held assumptions that EA is a stress-accent language in which pitch marks the stressed syllable of words was borne out by two experimental studies. These confirmed alignment of EA pitch movement with the edges of the stressed syllable rather than with the edges of the words, and use of non-tonal as well as tonal correlates of word-level prominence.

An additional parameter of prosodic variation is proposed whereby language may vary as to which domain is relevant for the realisation of pitch, regardless of the function of pitch in that language. Thus we see variation among languages in which the function of pitch is lexical in the domain within which pitch is realised: in a language like Basque only one lexical accent per phrase is realised. In addition we see variation among languages in which pitch is purely postlexical in the distribution of intonational pitch accents: English has relatively sparse pitch accent distribution but EA has rich pitch accent distribution with a pitch accent observed on every content word.

The next chapter pursues this proposal in its theoretical context, suggesting that this new parameter of variation may be expressed formally as variation across levels of the Prosodic Hierarchy in the domain of pitch accent distribution. Extensive empirical evidence is provided to support the claim that the relevant domain in EA is the Prosodic Word.

5 The domain of pitch accent distribution in Egyptian Arabic

5.0 Outline and aims

Chapters 3 & 4 presented evidence to suggest that Egyptian Arabic (EA) is an intonational stress-accent language with rich pitch accent distribution: a pitch accent is observed on every content word. This chapter explores this empirical fact in its theoretical context and explores possible categories that could be used to express density of pitch accent distribution as a parameter of prosodic variation across languages.

The chapter starts (in section 5.1) by outlining suggestions that have been made in the literature regarding the types of mechanism regulating pitch accent distribution in intonational languages. These fall roughly into two categories, with pitch accents as the direct reflex either of syntactic-semantic structure or of prosodic structure. This chapter pursues a prosodic-structure-based analysis of rich pitch accent distribution in EA. Section 5.2 reviews structure-based analyses of a variety of languages, in which researchers have independently proposed some constituent of the prosodic hierarchy as the domain of pitch accent distribution. Together these are argued to indicate that the domain of pitch accent distribution may vary across languages.

The remainder of the chapter explores empirical evidence from EA (from a pilot study and from the present corpus) in order to identify which of the constituents of the prosodic hierarchy serves as the domain of pitch accent distribution in EA. Section 5.3 examines evidence from prosodic phrasing in complex EA sentences which suggests that MaP boundaries are sparse in EA, and thus that the MaP cannot be the domain of pitch accent distribution. A formal OT analysis of the phrasing facts results in the proposal that the MiP in EA is minimally branching and thus composed of two PWds, both of which are accented; thus the MiP cannot be the domain of pitch accent distribution either.

Section 5.4 reviews the relative accentuation of content and function words in the corpus in detail in order to establish how lexical and function words are mapped to Prosodic Words (PWds). The chapter closes by arguing that the correct generalisation to describe EA rich pitch accent distribution is that every PWd is accented and thus that the domain of pitch accent distribution in EA is the PWd.

5.1 Pitch accent distribution: where do pitch accents come from?

5.1.1 Accentuation and focus

Much of the discussion in the literature regarding pitch accent distribution is inextricably interwoven with arguments about the prosodic reflexes of focus. This is due in part to the fact that in most cases the analyses treat data from Germanic languages (English, German, Dutch) in which there is an obvious and well-established relation between pitch accent distribution and focus context: discourse-given, 'out-of-focus', items fail to be accented (or are 'de-accented' Ladd 1980) in Germanic languages.

In this section I review two competing views of pitch accent distribution, both of which nonetheless view intonational choices as a *grammatical* matter, in which pitch accent placement reflects some aspect of grammatical structure. This contrasts with an inherently *functional* view of intonation, and therefore of the distribution of pitch accents, whereby speakers place accents so as to highlight what they want to on a specific occasion (see, inter alia, Bolinger 1972, Chafe 1974, Halliday 1967). Ladd (1996:160ff.) has dubbed these 'structure-based' and 'highlighting based' views of accentuation respectively. He argues persuasively that systematic cross-linguistic variation in 'default' patterns of accentuation (in questions, of predicates and arguments, and of items of low semantic weight) fatally undermines a purely functional view of intonational choices as the universal highlighting of contextually salient items[69].

In contrast, in the structure-based views described below, the relationship between focus and pitch accent distribution is indirect - there is an intervening structure of some kind that mediates between focus and accent. The position of the focus may well be influenced by universal discourse factors (such as those argued for under the functional view), but once the focussed constituent is identified, the position of accents within that constituent follow regular language-specific rules of pitch accent assignment. Within the structure-based view there are two positions as to what *type* of structure determines accent distribution.

As already noted, in many Germanic languages, focus affects pitch accent distribution: words which are given and therefore not focussed are 'de-accented'. On the basis of evidence from such languages, it has been widely argued that the position and distribution of pitch accents is determined by the *semantic-syntactic structure* of the

[69] The evidence for cross-linguistic variation in accentuation of semantically weak items is set out below.

sentence. The relevant structure has been formulated variously as the position of syntactic focus markers (the theory of 'F(ocus)-marking', Selkirk 1984) or as the distribution of arguments and predicates (the Sentence Accent Assignment Rule, Gussenhoven 1983). In these analyses a mismatch between the focus-determined position of pitch accents and unmarked prosodic phrasing will trigger adjustments in prosodic structure to accommodate the required position of pitch accents, and for this reason they have been termed ‘*accent-first*’ accounts (Selkirk 1984:265ff., cf. Ladd 1996:221ff.).

Other authors maintain a ‘*stress-first*’ structure-based view however, in which the distribution of pitch accents is determined by the distribution of constituents in *prosodic structure*. Under this view focus results in de-accenting indirectly: the focus context may affect phrasing (that is, the distribution of prosodic constituents), and if so it is the changed prosodic structure that results in changes in pitch accent distribution. In a stress-first analysis then, the primary reflex of focus is relative prominence in prosodic structure. The reflex of that prosodic prominence in many intonational languages is of course the presence of a pitch accent, but a stress-first analysis is better equipped to handle languages in which the reflexes of prosodic prominence are found to be marked by other means.

One such example is focus-induced ‘de-phrasing’ in Korean. When an item is ‘given’ by virtue of being repeated, in a Germanic language it will be de-accented, but in Korean such words are ‘de-phrased’. In the example in (5.1) below, reproduced from Ladd (1996), the word [iɾɨmi] ‘name’ is produced within its own phrase when it is ‘new’ information, in the question; but in the response to the question, the same word, now repeated and ‘given’, loses its own phrase status and is incorporated into the same phrase as the preceding word (Ladd 1996:196, citing Jun 1993)[70]:

(5.1)	A:	(satʃ^{h}un-ənni)	(iɾɨmi)	(mwəni)
		(cousin)	(name)	(what)
		“What is cousin’s name?”		
	B:	(satʃ^{h}un-ənni iɾɨmi)		(suni-dʒi)
		(cousin name)		(Suni)

[70] Jun (1993) cites the observed distribution of phrase edge tones as evidence for this phasing analysis.

By definition then, a stress-first account is going to be better placed to account for the facts of a language like EA in which presence or absence of pitch accents does not inherently mark focus. A key piece of evidence in favour of a 'stress-first' account of pitch accent distribution are cases where pitch accents are observed on words which are not focussed.

As an example of 'accent-without-focus' Ladd cites the presence of early (pre-nuclear) accents in broad focus utterances, which have been acknowledged to be problematic for 'accent-first' analyses (e.g.Selkirk 1984:274). For example in a sentence in which the whole noun phrase 'a million dollars' is in broad focus (rather than narrow focus on either 'million' or 'dollars') such as in (5.2) below, both words will be accented by most speakers of English (after Ladd 1996:163,223):

(5.2) I didn't give him my car keys, I gave him [a míllion dóllars].

In English, broad focus is indicated by the presence of a pitch accent on the rightmost element in the phrase. Nonetheless in the above example there is also a pitch accent on the word 'million'. Under an accent-first account additional explanation is required for the extra pitch accent on 'million'[71], whereas under a stress-first account the additional accent is not problematic so long as 'a million dollars' forms a single prosodic phrase and the final word ('dollars') retains relative prominence in the phrase.

The corpus survey described in chapter 3 demonstrated that all content words in EA bear a pitch accent, and this is true regardless of focus context (section 3.3.1). This suggests that even an amended accent-first account would struggle to account for the facts of EA. Languages in which every content word is routinely accented force an analysis in which the relation between pitch accent distribution and focus/information structure is indirect.

Similar in nature to the problem of pre-nuclear accents, another instance of non-focus-marking pitch accents are those in post-nuclear positions. Ladd (1996:160-197)

[71] The accent on 'million' is arguably a different type of pitch accent in most speakers' productions, and Ladd (1996:225-6) suggests ways in which this could be the basis of an explanation, within the various accent-first theories, and which would also accommodate the phonetic facts of the example: if focus were signalled by primary accents only (and the 'extra' pitch accent argued to be a secondary accent).

discusses such examples in detail, as part of the arguments in favour of a grammatical rather than functional account of accentuation.

In most Germanic languages, words which are discourse 'given' in context are routinely 'de-accented', and the consistency of this effect is no doubt the source of the widely held notion that pitch accents are inherently focus-marking. However Ladd demonstrates convincingly that the contexts which in Germanic languages condition de-accenting do not do so in other languages. The non-de-accenting languages that he cites are largely (though not exclusively) of the Romance family. As we might expect given the findings of chapter 3, and as set out below, the same contexts also fail to condition de-accenting in EA.

In a parallel to the Korean example above, in English if a word we might otherwise expect to be accented is *repeated* it fails to be accented, as shown in (5.3) (the following examples use Ladd's notation in which accented words are shown in capital letters 1996:175):

(5.3) A: I found an article for you in a German journal. *English*
B: I don't READ German.

Other languages do not however modify prominence patterns to reflect 'givenness' in this way (Ladd 1996:176):

(5.4) 43% is government OWNED and 57% is privately OWNED. *Hawaiian pidgin*

(5.5) [… o să vedem] ce AVETI si ce nu AVETI *Romanian*
[…we'll see] what you.have and what not you.have
'so let's see what you HAVE and what you don't HAVE'

A similar lack of de-accenting is observed in other Romance languages, and other dialects of English (examples cited in Cruttenden 2006):

(5.6) *Brazilian Portuguese*

Esti	livro	custa	cinco	DOLARES	e esti	aqui	tres	DOLARES.
this	book	cost	five	dollars	and this	here	three	dollars

"This book costs five dollars and this other one three dollars".

(5.7)	a)	¿VIENES	o	no	VIENES?		*Chilean Spanish*
		you(s.)-come	or	NEG	you(s.)-come		
		"Can you come or not?"					
	b)	¿Con LECHE	o	sin	LECHE?		
		With milk	or	without	milk		
		"With or without milk?"					

(5.8) A: You just weren't LISTENING. *Indian English*
B: I was LISTENING.

(5.9) I went to the shop to buy SWEETS, *Singapore English*
but they had totally run out of SWEETS.

The facts of EA suggest that it should be grouped with the 'Romance-type' languages, since it resists de-accenting of textually-given items (see chapter 3; cf. also Hellmuth 2005). The following examples come from a pilot study in which a selection of the data from Ladd (1996 chapter 5) were translated and recorded with two speakers of EA (Hellmuth 2002b). Speakers read both the plain statement B in isolation as in (5.10) and also the A~B statement~response pair in (5.11), in which the word [?almaani] 'German' is repeated and thus 'given'.

(5.10)	B1	ana	maa-b-a?raa-š				?almaani	
		I	NEG-HAB.-read-NEG				German	
		"I don't read German"						
(5.11)	A	ana	l?ayt	ma?aala	liik	fi	gurnaal	bil-?almaani
		I	found	article	for-you	in	journal	in-German
		"I found an article for you in a German journal"						
	B2	ana	maa-b-a?raa-š				?almaani	
		I	NEG-HAB.-read-NEG				German	
		"I don't read German"						

As shown in Figures (5.1-5.2) below however, in both contexts the final word [?almaani] is produced with an accent (albeit in reduced pitch range[72]).

[72] This is probably due to attraction of a focal accent to a negated verb in EA, reported by Mitchell (Mitchell 1993) for other dialects. Chapter 8 investigates pitch range manipulation as a reflex of exhaustive/contrastive focus in EA.

Figure 5.1 B2 sentence in contrast context (*DAGA1EX-FIH3*)

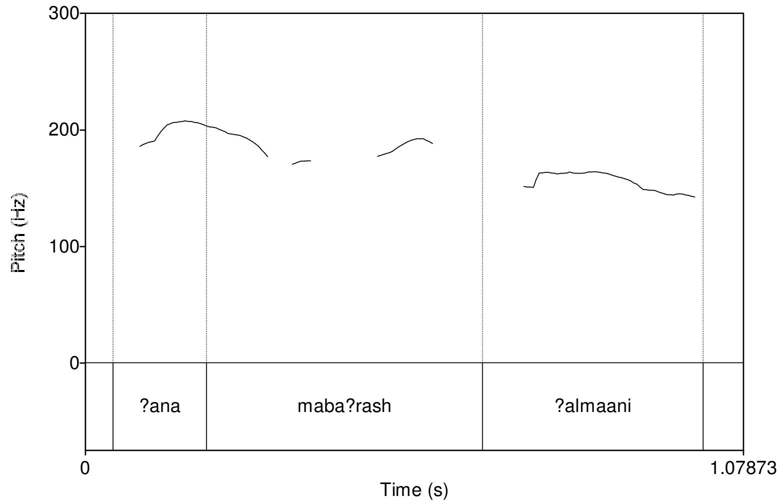
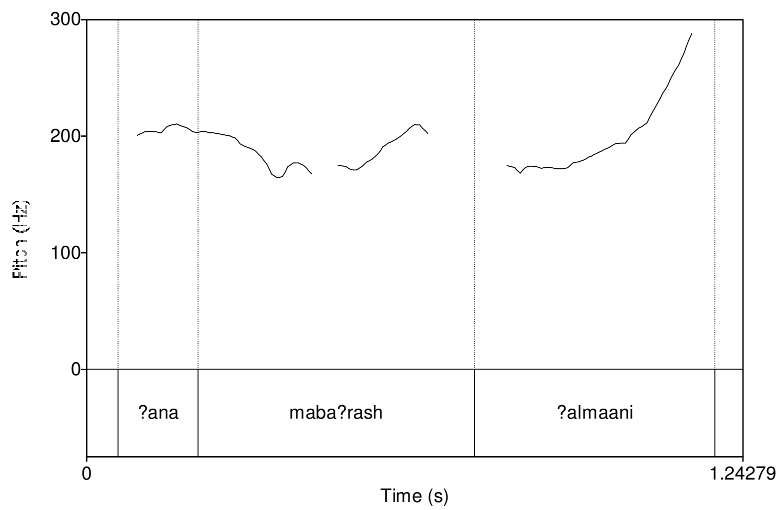


Figure 5.2 B1 sentence in non-contrast context (*DAGA4-AA1*)[73]



A further example below comes from the read narratives section of the thesis corpus during a section of the story devoted to bargaining over prices. There is much repetition of phrases involving the unit of currency [saaG] 'piastre'[74], along the lines of 'you said six piastres, not three piastres'. In a Germanic language the tendency would be to elide the repeated currency unit completely ('you said six piastres not three'), or if the word is overt to de-accent it. In the read narratives, since speakers were reading from a script, the currency unit was always overt and was accented by all speakers. In the retold versions of the narratives, which speakers were asked to produce from memory, there

[73] In this token the speaker ends the phrase with a high rising boundary tone (signalling continuation).
[74] There are one hundred piastres ([saaG] or [?irš]) in one Egyptian pound ([gineeh]).

was variation with some speakers omitting the currency unit and some pronouncing it; again, where overt the word [saaG] was invariably accented.

The sequence of sentences in the bargaining exchange (sentences A-C) are set out in example (5.12) below, together with auditory transcriptions of speakers' productions of the second and third sentences of the exchange (B and C), in which the currency unit [saaG] is repeated from previous sentences in the exchange. Renditions of the sentences in retold versions of the narratives are given in auditory transcription in (5.13). Similarly, in all cases where the currency unit was overt it was accented. Sample pitch tracks of both read and retold renditions of sentences B and C are provided in Figures 5.3-5.7.

Figure 5.3 Accenting of repeated [saaG] in a read rendition of sentence B (*fna2*).

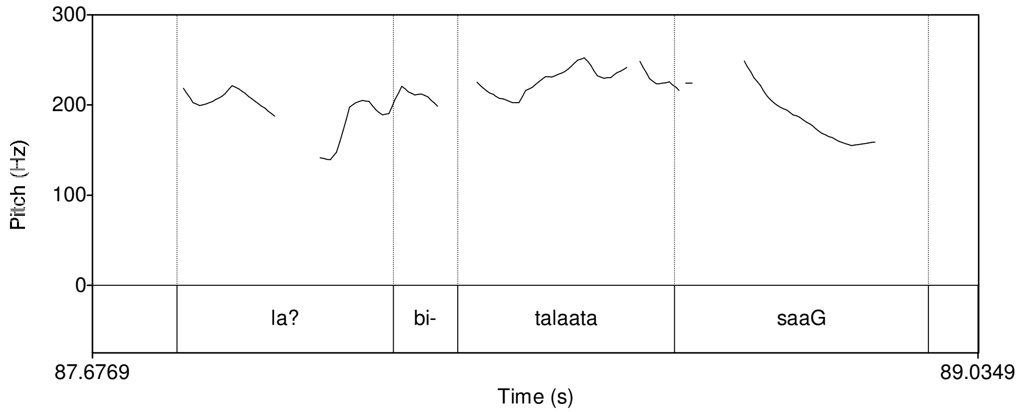
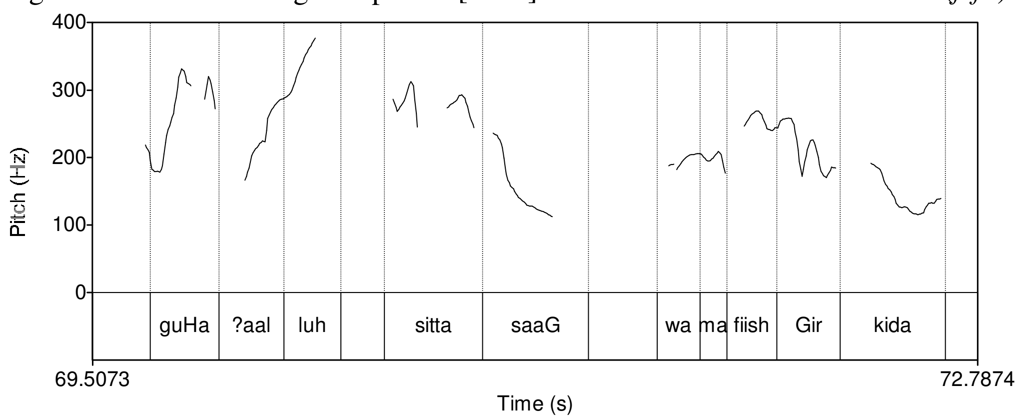


Figure 5.4 Accenting of repeated [saaG] in a read rendition of sentence B (*fsf4*).



(5.12) The bargaining exchange in the Guha narrative.

A													
	ir	-raagil	?aal-luh	?a?ul-lak	?eh	?ana	Habi9-lak	kiilu	bi	sitta	saaG	bass	9ašaan	xaTrak
	the	man	said-him	I-tell-you	what	I	will-sell-you	kilo	for	six	piastres	just	because	your-sake

‘The man said to him. I tell you what. Because it’s you, I’ll sell you a kilo for six piastres.’

B									
	fa	guHa	9ala	Tuul	?aal-luh	la?	bi-	talaata	saaG
	so	Guha	at	once	said-him	no	for	three	piastres
fna2	>	LH*	---	LH*	LH* H-	LH* L-	>	LH*	!LH* L-L%
fsf2	>	LH*	<	LH*	LH* H-	xxx	>	LH*	!LH* L-L%
meh2	>	LH*	---	LH*	LH*	LH* L-	>	LH*	!LH* L-L%
miz2	>	LH*	<	LH*	LH*	LH*	>	LH*	LH* L-
mns2	>	LH*	<	LH* H-	LH* H-	xxx	>	LH*	!LH* L-L%

‘So Guha straight away said to him: “No! Three piastres.”’

C											
	il	bayyaa9	?aal-luh	ya	xuya	?inta	?ult	sitta	saaG	min	da?ii?a
	the	seller	said-him	oh	my-brother	you	said	six	piastres	from	a-minute
fna2	>	LH*	LH* H-	>	LH*	LH*	LH*	LH*	LH*	>	↑LH* L-L%
fsf2	>	LH*	LH* H-	>	LH*	LH*	LH*	LH*	LH*	>	LH* L-L%
meh2	>	LH*	LH* H-	>	LH*	LH*	LH*	LH*	LH*	>	↑LH* L-L%
miz2	>	LH*	LH* H- ~	>	LH*	LH*	LH*	LH*	LH*	>	LH* L-L%
mns2	>	LH*	LH*	>	LH*	LH* H-	LH*	LH*	LH*	>	LH* H-

‘The seller said to him. “Look mate, you said six piastres a minute ago”.’

(5.13) Accenting of overt [saaG] ‘ Renditions of the bargaining exchange in retold narratives.
i) speaker *fna*; ii) speaker *fsf*; iii) speaker *meh*

i)	C	wa	raaH	il	bayyaa9	?aal-luh	manta	lissa	?aayil	min	da?ii?a	waHda	bi	xamsa	SaaG
		and	went	the	seller	said-him	but-you	just	saying	from	minute	one	for	five	piastres
	fna4	>	LH*	<	↑LH*	LH*	LH*	↑LH*	↑LH*	<	↑LH*	↑LH* L-L%	>	↑LH*	↑LH* L-L%

ii)	B	guHa	?aal-luh	sitta	SaaG	wa	ma	fiiš	Gayr	kida				
		Guha	said-him	six	piastres	and	not	there-is	except	that				
	fsf4	LH*	LH* H-	LH*	!LH* L-L%	>	>	LH*	<	LH* L-L*				
	C	il	bayaa9	?aal-luh	9ašaan	xaTrak	?ana	Haddiik	ik-	kiilu	bi	sitta	SaaG	
		the	seller	said-him	for	your-sake	I	will-sell-you	the	kilo	for	six	piastres	
	fsf4	>	LH*	LH* H-H%	LH*	LH* H-	---	LH*	<	LH*	<	LH*	LH* H-L%	

iii)	B	guHa	?aal-luh	la?	bi	sitta	?uruuš[75]				
		Guha	said-him	no	for	six	piastres				
	meh4	LH*	LH*	LH* H-H%	>	LH*	LH* L-L%				
	C	ba9da	kida	?aal-luh	xalaaS	?addiik	kiilu	bi	sitta	?uruuš	
		after	that	said-him	ok	I-give-you	kilo	for	six	piastres	
	meh4	>	LH* L- ==~	LH*	LH* L-L%	LH*	LH*	<	LH*	!LH* L-L%	

75 An alternative word for piastres is [?irš], of which the plural is [?uruuš].

Figure 5.5 Accenting of repeated [saaG] in a read rendition of sentence B (*meh2*).

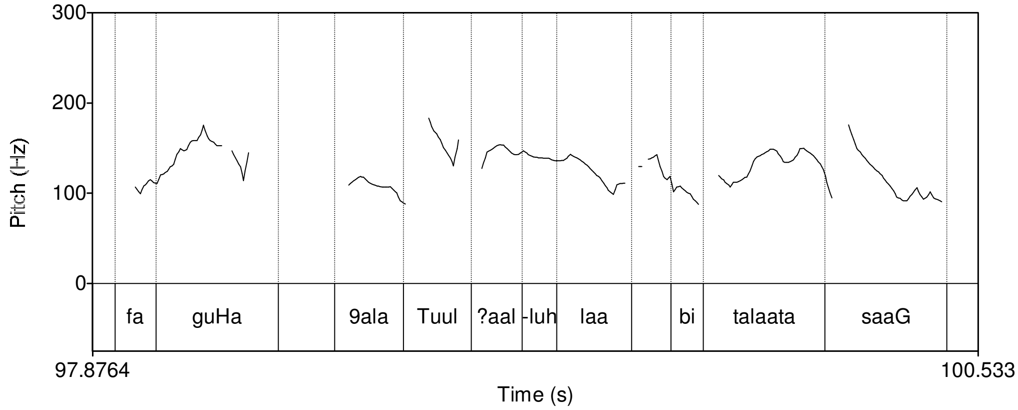


Figure 5.6 Accenting of repeated [saaG] in a read rendition of sentence C (*fsf2*).

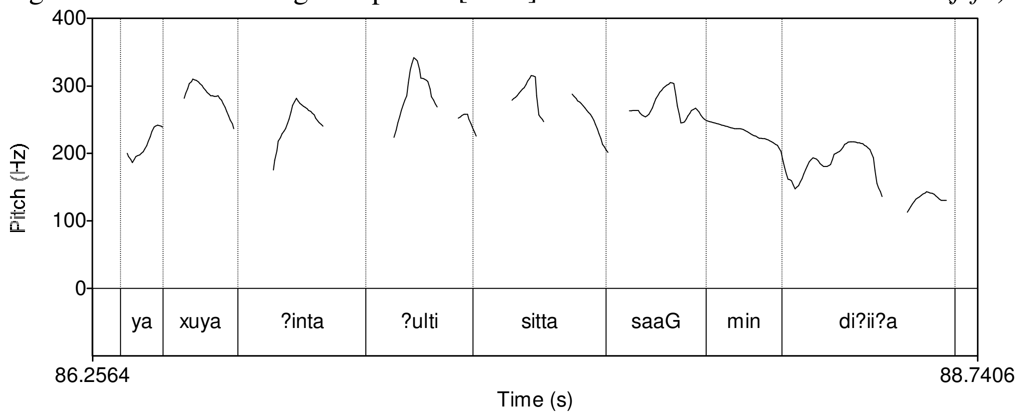
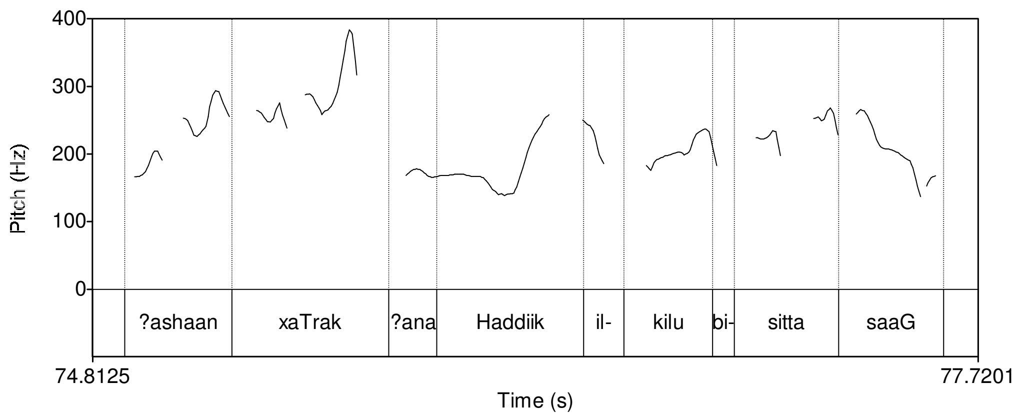


Figure 5.7 Accenting of repeated [saaG] in a retold rendition of sentence C (*fsf4*).



On the basis of Ladd's arguments therefore, and supplemented by evidence from EA as another language in which accentuation proves not to be a direct reflex of semantic-syntactic structure, the analysis of EA that follows is a stress-first structure based account.

5.1.2 The distribution of pitch accents

The claim of this chapter is that the stress-first view is uniquely equipped to account for EA's rich pitch accent distribution, because it directly encodes the relationship between prosodic constituents and accents.

This is illustrated in further examples which Ladd sets out to illustrate how a stress-first account can explain otherwise puzzling cases of changes in pitch accent distribution. The often-quoted sentence in (5.12) has two possible positions for the main accent, which result in two different interpretations of the sentence (Ladd 1996:199,233, citing Halliday 1970):

(5.12) a. DÓGS must be carried.
b. Dógs must be CÁRRIED.

Gussenhoven (1980) has characterised the different interpretations of the sentence as an eventive reading , in (5.12a) 'You must have a dog and carry it', vs. a contingency reading in (5.12b) 'If you have a dog, you must carry it'.

Ladd (1996:233) argues that the distribution of pitch accents in this example is the direct reflex of the phrasing structure: the number of pitch accents indicates the number of phrases into which the sentence is divided. When the sentence is composed of a single phrase, there is a single primary accent: |DÓGS must be carried|. When the sentence is composed of two phrases however there are two primary accents, of which the rightmost receives main relative prominence in the utterance: |Dógs|must be CARRIED|. The definition of primary accent that Ladd gives for these English examples is the main prominence (the Designated Terminal Element (DTE[76])) of the Intermediate Phrase.

[76] This most prominent constituent within X is the *DTE* of X, and can also be described as the *head* of X (see chapter 2 section 2.1.5).

In another similar example, Ladd suggests that a very emphatic rendition of an English sentence, in which every word bears a primary accent, is an indicator that each word is mapped to an individual phrase: again, the number of phrases equates to the number of accents (Ladd 1996:249):

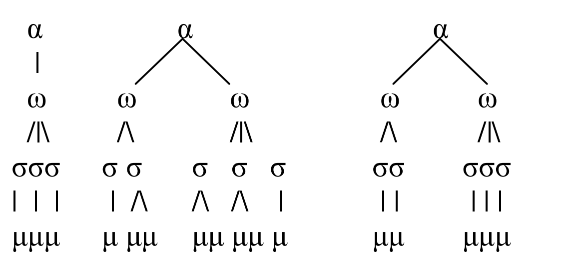
(5.13) A: Everything OK after your operation?
B: Don’t talk to me about it.
H* L H* L H* L H*L-L%
The bútcher chárged me a thóusand búcks!

An important implication of the prosodic-structure-based (stress-first) view of intonation then is the notion that there is a direct correlation between pitch accent distribution and prosodic constituency.

The idea that the distribution of tonal events directly reflects intonational prosodic structure is of course inherent to the original ideas proposed in intonational phonology, regarding the distribution of pitch accents and phrase/boundary tones. Pierrehumbert & Beckman (1986, 1988) suggested that the position of pitch accents and phrase/boundary tones *directly* reflect the position of the heads and edges of a hierarchy of intonationally defined prosodic constituents: Accentual Phrase (AP), Intermediate Phrase and Intonational Phrase.

In the following example in Japanese, the surface tonal contour results from the sequence of tonal events, determined by the position of prosodic phrase edges (the left and right edges of each AP are marked by a H and L tone respectively), and the position of lexically specified falling pitch accents (denoted in small capitals HL) (Pierrehumbert & Beckman 1988:128):

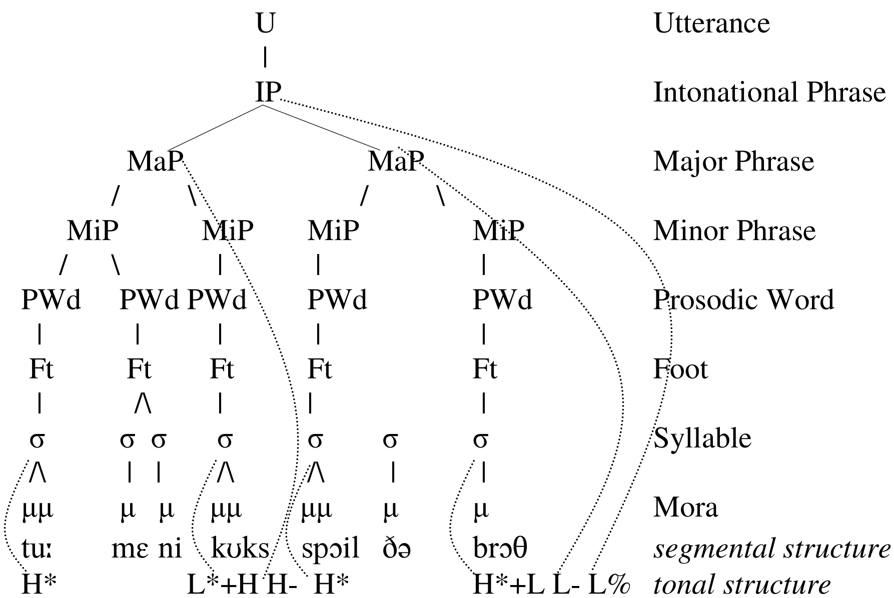
(5.14)



H L H HL L H HL L
ane-no aka’i se’e ta a-wa do’ko desuka’?

The model is designed to be used for other languages: for example in English phrase tones (H-/ L-) arguably mark the right edge of phonological phrases (≈MaP) and boundary tones (H%/L%) mark the right edge of Intonational Phrases (after Gussenhoven 2004:124)[77]:

(5.15)



Pierrehumbert & Beckman do not specify whether or not the hierarchy of prosodic phrases to which they appeal is the same as that which in other theories is constructed as a result of the interface with syntactic structure (that is, the Prosodic Hierarchy assumed here; see chapter 2 section 2.1.5). Many authors have however assumed what Frota (2000) terms an 'integrated view' in which the prosodic hierarchy which results from the interface with syntax is indeed the same prosodic structure to which intonational pitch events are sensitive[78]. This thesis adopts the null hypothesis that there is a single prosodic representation, which is the result of the interface between the syntax and phonology, and which is reflected in all types of prosodic cues[79].

Combining Ladd's ideas and those of Pierrehumbert & Beckman then, the distribution of pitch accents is here argued to reflect the distribution of prosodic constituents. The next section reviews analyses that have been made by a number of researchers which

[77] The representation is here re-interpreted in terms of the conception of the Prosodic Hierarchy assumed in the thesis.
[78] Authors who have set out data in which there is a mismatch between the prosodic representations cued by intonation and by other cues respectively, include Dresher (1994) and Gussenhoven & Rietveld (1992); cf. also (Ladd 1996:237ff.).
[79] Cues to prosodic phrasing in EA are discussed throughout the remainder of this chapter and the next.

appeal directly to this notion that the distribution of pitch accents reflect the distribution of prosodic constituents. Specifically the analyses argue that some language-specific constituent level of the prosodic hierarchy functions as the domain of pitch accent distribution, and constrains both insertion and repositioning of pitch accents.

5.2 Pitch accent distribution as a reflex of prosodic structure

This section reviews analyses of a number of languages in which pitch accent distribution is argued to be determined by prosodic structure, and specifically by some language-specific constituent level of the prosodic hierarchy. The aim of this section is twofold: i) to show that the distribution of pitch accents can provide evidence of the distribution of prosodic phrases; and, ii) that this evidence indicates possible cross-linguistic variation in which level of the hierarchy is the relevant domain of pitch accent distribution. The outcome is that it is not possible to assume *a priori* which level of the prosodic hierarchy will prove to be the domain of pitch accent distribution in EA.

It is worth noting at this point that the analyses outlined below do not share a common view of the formulation of the levels of the Prosodic Hierarchy. The analyses are here grouped according to equivalent levels in the conception of the hierarchy assumed here as set out in chapter 2 (section 2.1.5), reproduced in (5.16) below. The authors' original terminology is used, with equivalents noted clearly in brackets or in a footnote[80].

(5.16) The Prosodic Hierarchy

constituent		*equates to:*	*maps from:*
Utterance	U		
Intonational Phrase	IP		a root sentence or sentence-external clause
Major Phonological Phrase	MaP	phonological phrase intermediate phrase	a maximal projection (XP)
Minor Phonological Phrase	MiP	accentual phrase	a syntactically branching constituent (two PWds)
Prosodic Word	PWd	phonological word	a morphosyntactic word (lexical)
Foot	Ft		
Syllable	σ		
Mora	μ		

[80] The authors also work within varying conceptions of the nature of the syntax-phonology interface and these distinctions will be made clear.

5.2.1 The (major) phonological phrase as the domain of pitch accent distribution

As we have seen, Ladd (1996:249) has suggested that the domain of (primary) accent distribution in English is the Intermediate Phrase, which is here assumed to equate to the MaP[81]. As we shall see below however Selkirk has argued for a different domain as the relevant domain of pitch accent distribution in English (MiP).

Two authors have independently proposed the MaP (in both cases formulated as the Phonological Phrase) as playing a role in pitch accent distribution and repositioning. They are Frota (2000) for European Portuguese (EP) and Post (2000) for French. Both work within a stress-first prosodic structure based view, though they assume slightly different conceptions of the nature of the interface between syntax and phonology: Frota assumes the 'relation-based' mapping from syntax to phonology following Nespor & Vogel (1986), whilst Post assumes an 'end-based' mapping, formalised within Optimality Theory (Selkirk 1986, Selkirk 1996, Selkirk 2000).

5.2.1.1 European Portuguese (EP)

As part of a wide-ranging study of phrasing and focus in EP Frota (2000:186-9) appeals to the relationship between pitch accent distribution and prosodic phrasing, in order to distinguish between the predictions of two of the prosodic theories prevailing at the time: the relation-based and edge-based theories (Nespor & Vogel 1986, Selkirk 1986; see chapter 2 section 2.2.2.2). She notes that this is possible because (Frota 2000:187):

> "within a framework in which the distribution of tonal events is governed by the edges and prominence relations established on the basis of prosodic structure.. the way in which a sequence.. is phrased.. determines which patterns of pitch accent distribution may or may not be implemented".

Frota analyses variant accentuation patterns observed in complex NPs, elicited in both subject and object position of target sentences as shown in (5.17) (Frota 2000:176).

(5.17) [Uma progressiva subida [dos preços]$_{PP}$]$_{NP}$ afectará a economia. *subject position*
a gradual rise in-the prices will-affect the economy

Os jornais prevêem [uma progressiva subida [dos preços]$_{PP}$]$_{NP}$. *object position*
the newspapers anticipate a gradual rise in-the prices

[81] See Gussenhoven (2004:166-7) for a different view however.

Observed accentuation patterns in the complex NPs are as shown in (5.18a & b)[82]. In contrast, the pitch accent distribution in (5.18c) is not observed in any position:

(5.18) Observed pitch accent distribution in complex NPs in EP:

		uma	progressiva	subida	dos preços	
a.		uma progressiva subida dos preços				*subject position*
			PA	PA	PA	80%
			-	PA	PA	20%
b.		uma progressiva subida dos preços				*object position*
			-	-	PA	50%
			-	PA	PA	23%
			PA	PA	PA	27%
c.	*		PA	-	PA	0% *subject/object position*

Frota uses these pitch accent distribution patterns as evidence in favour of the relation-based mapping (Nespor & Vogel 1986) over edge-based mapping (Selkirk 1986). The two mapping algorithms predict different phrasing of the complex NPs, based on analysis of the PP as internal to the noun phrase. Frota argues that the relation-based mapping predicts a Phonological Phrase (≈MaP) break between the head noun and its PP complement as indicated in (5.19a) below. The end-based mapping predicts a single Phonological Phrase boundary at the right edge of the complex NP (which is also the location of the right edge of the embedded PP), as in (5.19b)[83]. Assuming that the rightmost word in each Phonological Phrase must bear an accent, and that all other accents are optional, the possible accent distribution patterns expected under each predicted phrasing are as shown.

(5.19) a.		|uma progressiva subida|$_{PP}$ |dos preços|$_{PP}$				*relation-based*
			PA	PA	PA	
			-	PA	PA	
	*		PA	-	PA	
b.		|uma progressiva subida dos preços|$_{PP}$				*edge-based*
			-	-	PA	
			-	PA	PA	
			PA	-	PA	
			PA	PA	PA	

[82] Frota observes that there are in general more pitch accents in complex NPs when they occur in subject position than in object position. The minimal accentuation pattern in an EP sentence is a final nuclear accent; if there is another pitch accent the most likely location is always sentence-initial; additional prenuclear accents over and above the initial one are optional. Due to its initial position in the sentence then, the subject is always the most likely candidate to bear additional prenuclear pitch accents (ibid.).

[83] In a constraint-based framework these would equate to: a) ALIGNXP >> WRAPXP; and b) WRAPXP >> ALIGNXP (Truckenbrodt 1995, Truckenbrodt 1999, Selkirk 2000). Based on Frota's observations the pitch accent distribution patterns suggests that the correct ranking in EP is: ALIGNXP >> WRAPXP.

The phrasing pattern predicted by the relation-based mapping in (5.19a) is argued to yield exactly those patterns of pitch accent distribution that are observed, and predicts as ungrammatical exactly the accentuation pattern which is never observed.

Crucially, for our present purposes, Frota's argument turns on the assumption that there must be at least one pitch accent in every Phonological Phrase (≈MaP) (Frota 2000:188):

> "within the Phonological Phrase, the rightmost element has priority regarding pitch accent association because it is the most prominent one. Thus, in case there is only one pitch accent in the Phonological Phrase it must go on the phonological-phrase-head".

Frota thus successfully appeals to patterns of pitch accent distribution as evidence of prosodic phrasing, based on the assumption that the head of every Phonological Phrase is obligatorily accented, and that the Phonological Phrase-head in EP is rightmost.

5.2.1.2 French

The domain of pitch accent distribution in French is argued by Post (2000) to be the Phonological Phrase (PP, ≈MaP): "the distribution of pitch accents is conditioned by the Phonological Phrase in French" (Post 2000:81).

The key constraint is again, as in EP, one in which the right edge of a PP (≈MaP) is marked by a pitch accent[84]:

(5.20) RIGHTMOSTPP: Align (PP,R; Pitch Accent, R).

Post suggests that: "the right edge of each PP coincides with the right edge of a lexical head (noun, verb or adjective) .. [except] pre-nominal adjectives)" (ibid.). For the sentence in (5.21a) below then, the phrasing and minimum accentuation is as shown in (5.21b), with pitch accents indicated by asterisks:

[84] This constraint follows the schema of the ALIGN family of constraints (McCarthy & Prince 1993).

(5.21)

a. [[Ces [petits]$_{AdjP}$ enfants [intelligents]$_{AdjP}$]$_{NP}$ [apprennent [à parler [le français]$_{NP}$]$_{VP}$]$_{VP}$]$_{S}$

b.

(Ces petits enfants)	(intelligents)	(apprennent)	(à parler)	(le français)
*	*	*	*	*
These small children	intelligent (pl.)	learn	to speak	the-French.

In fact the most common, ‘default’, pitch accent distribution pattern is as shown in (5.22), with additional optional accents in some phrases (optional accents are marked with smaller asterisks):

(5.22)

(Ces petits enfants)	(intelligents)	(apprennent)	(à parler)	(le français)
* *	* *	*	*	*
These small children	intelligent (pl.)	learn	to speak	the-French.

The appearance of optional pitch accents at other positions within the PP (≈MaP) is regulated in Post’s analysis by constraints requiring pitch accents to align at the left and right edge of each PWd (RightmostPWd; LeftmostPWd)[85]. In an amended version of Post’s analysis, Gussenhoven (2004) replaces LeftmostPWd with a constraint which equates to LeftmostPP and assumes a “high-ranking constraint that only allows an accent on lexical words” (2004:260). In both analyses however, these constraints interact with rhythmic well-formedness constraints banning pitch accents which are too close or too far apart (NoClash/NoLapse)[86]. Post’s pitch accent distribution data come from transcriptions of a corpus of read sentences and spontaneous speech. The data show variation in pitch accent distribution within each PP and this variation is neatly accounted for by limited re-ranking of constraints.

Note that this analysis conflates the two assumptions made by Frota for European Portuguese: instead of stating that the head of every PP (≈MaP) is obligatorily accented and that the PP-head in European Portuguese is rightmost, instead Post proposes a constraint which simply requires a pitch accent at the right edge of every PP.

[85] There is a constraint ensuring that every lexical word maps to a PWd: LexWdPWd: Align (Lex, R/L; PWd, R/L); cf. further discussion of this constraint, and its role in EA, in chapter 6 section 6.2.2.
[86] For further discussion of NoLapse and its potential role in EA see chapter 6 section 6.2.3. In Post’s analysis an additional constraint restricts optional accents from appearing on vowel-initial syllables.

5.2.1.3 Summary: the MaP as the domain of pitch accent distribution

The surface phrasing facts of EP and French are in fact quite similar despite the authors' different conclusions. For our present purposes however these analyses are relevant for two reasons: i) they rely on pitch accent distribution as evidence of prosodic constituency; and ii) in EP and French the domain relevant to pitch accent distribution is argued to be the phonological phrase (here re-named MaP).

5.2.2 The (minor) phonological phrase as the domain of pitch accent distribution

This section again reviews analyses which have made appeal to the distribution of pitch accents (or boundary tones) as evidence of prosodic constituency. Here however, the relevant domain is at a lower level in the Prosodic Hierarchy, namely the MiP (formulated here as the Accentual Phrase by two of the authors). The three analyses surveyed here are Pierrehumbert & Beckman (1986, 1988) for Japanese, Jun (1996, 2005a) for Korean and Selkirk (2000) for English[87].

5.2.2.1 Japanese and Korean

As already noted, Pierrehumbert & Beckman (1986, 1988) developed a theory of intonational phonology in which the position of pitch accents and phrase/boundary tones reflect the position of the heads and edges of a hierarchy of prosodic constituents (Accentual Phrase, Intermediate Phrase and Intonational Phrase). In their 1986 paper they explicitly re-evaluate the nature of English intonational structure in the light of a study of the Tokyo dialect of Japanese, which is a pitch accent language. Although English has a larger inventory of pitch accent types (H*, L+H* etc) than Japanese (H*+L only), the system of association between prosodic structure and tonal events in the two languages is claimed to be the same (Beckman & Pierrehumbert 1986:261 italics mine):

> "the pitch accent in Japanese could.. be treated.. by specifying only starred syllables in the lexicon and inserting a H*+L along with other intonationally specified tones.. This treatment would make the pitch accent in Japanese rather more similar to the pitch accent in English, but it would not eliminate the crucial difference in the *function* of the pitch accent."

[87] The Minor Phrase is a well-motivated constituent level of the Prosodic Hierarchy in Japanese (Poser 1984) and Korean(Jun 1996). Its implementation in English is less-widely accepted, with other authors arguing that the domain relevant to pitch accent distribution in English is the phonological phrase (≈MaP)(Gussenhoven 2004, cf. also Ladd 1996:249 as above).

The Accentual Phrase (AP, ≈MiP) is thus seen as the domain of pitch accent distribution in Japanese, since an AP may contain either 0 or 1 pitch accents, with tone associated to a lexically pre-specified prominent position (the head) in the AP (as in 5.14 above). In addition every AP, even if it contains no lexical pitch accents, is marked with a distinctive LH tone sequence at the left edge (Beckman & Pierrehumbert 1986:262).

Jun (1996, 2005a) proposes that Korean has intonationally defined prosodic structure, which although related to syntactic structure is also heavily influenced by other factors such as speech rate, information structure and phonological weight (constituent length, in syllables). In the Seoul dialect, an AP is marked by a LHLH sequence (which is however only fully realised when there are sufficient syllables available in the AP). These four tones associate in LH pairs to the initial and final syllable pairs of the AP.

Although the PWd is not tonally marked in Korean Jun finds clear phonetic evidence to suggest that the PWd level is present in the language, and can be distinguished from MiP-level constituents. The voice onset time (VOT) of consonants is found to fall into three significantly different groups depending on whether the segment is word-initial and AP-initial, word-initial but AP-medial, or word-medial. This suggests that the PWd boundary is phonetically marked, even though there is no tonal marking[88].

For our present purposes this confirms that the tonal events are a reflex of the AP-level in Korean, which is shown to be frequently co-extensive with, but nonetheless distinct from, the PWd level.

5.2.2.2 English

The analysis of English offered by Selkirk (2000) involves both MiP-level phrases and MaP-level phrases, but relies on MiP rather than MaP as the domain of pitch accent distribution. Selkirk discusses variation in pitch accent distribution patterns according to the detail of focus context, as in a sentence such as (5.23a) below. When the main focus of the sentence is early, the distinction turns on whether or not the post-focal item (‘rollerblades’) is new or given in context. Selkirk suggests that even after a narrow focus a *new* item is accented in English, and resists post-focal de-accenting (main

---

[88] Jun defines the PWd in Korean as the minimal sequence of segments which bear the tonal marking of an AP (MiP) under neutral focus (Jun 1996). Note that although Jun (1996) does not propose an intermediate phrasing level between AP and Intonational Phrase (IP) for Korean, this possibility has been suggested in more recent work(Jun 2004).

sentence prominence is indicated in **bold**; *accented* syllables are marked with an acute accent: [á]) (after Selkirk 2000:251):

(5.23) a. She loaned her rollerblades to Robin.

I heard a rumour that she is selling all her stuff,
but it says here that...
b. → | She **lóaned** $|_{MaP}$ | her róllerblades $|_{MaP}$ | to Róbin $|_{MaP}$

I thought she sold her rollerblades to Robin,
but it says here that...
c. → | She **lóaned** her rollerblades to Robin $|_{MaP}$

Selkirk appeals to two constraints to analyse these surface accentuation facts in English. Working in an 'accent-first' framework in which focus determines the position of accents in an input representation, the first is a faithfulness constraint DEPACCENT, defined formally in (5.24) below, which requires all accents in the output to have a corresponding accent in the interface representation[89]. In essence this is a ban on insertion of any accents not required by the information structure of the sentence. The second constraint that Selkirk proposes, and which is most relevant to our present discussion, is a negative markedness constraint, banning MiPs which do not contain at least one accent, as in (5.25).

(5.24) DEPACCENT: An accent in the output representation must have a corresponding accent in the interface representation.

(5.25) MIPACCENT: Every minor phonological phrase (MiP) must contain at least one accent.

Selkirk assumes that HEADEDNESS is highly ranked[90], and thus that every MaP must contain a MiP. By transitivity therefore, if every MiP contains at least one accent, every MaP must also contain at least one accent. A constraint calling for the right edge of a MaP to align with the right edge of every focussed word (ALIGN$_R$FOCUS) interacts with MIPACCENT and DEPACCENT to yield the correct result as shown in (5.26). MIPACCENT penalises any MaP (by transitivity) that does not contain an accent, whilst DEPACCENT penalises any output accent with no correspondent in the input. The result is a phrasing in a single MaP (to match the presence of a single accent in the input).

[89] That is, in the output from the syntactic representation, where F-marking of syntactic constituents takes place (Selkirk 1984, Selkirk 1995).
[90] HEADEDNESS: Any C^i must dominate a C^{i-1} (that is, a MaP must dominate a MiP; see section 2.2.2).

(5.26) (Selkirk 2000:254)

		[she lóaned$_{V\text{-}FOC}$ [her rollerblades]$_{NP}$ [to Robin]$_{PP}$]$_{VP}$	MIPACCENT	DEP(ACCENT)	ALIGN$_R$FOCUS
☞	a.	|(She **lóaned** her rollerblades to Robin)|$_{MaP}$			*
	b.	|(She **lóaned**)|$_{MaP}$ |(her róllerblades)|$_{MaP}$ |(to Róbin)|$_{MaP}$		**	
	c.	|(She **lóaned**)|$_{MaP}$ |(her róllerblades to Róbin)|$_{MaP}$		**	
	d.	|(She **lóaned**)|$_{MaP}$ |(her róllerblades to Robin)|$_{MaP}$		*	
	e.	|(She **lóaned**)|$_{MaP}$ |(her rollerblades to Róbin)|$_{MaP}$		*	

There are two caveats to note here: firstly, Selkirk is working within an accent-first theory (focus is conveyed by the position of pitch accents in the interface representation), so her analysis essentially matches the number of prosodic constituents to the number of 'underlying' pitch accents[91]. Nonetheless the analysis assigns an important role to the relationship between tone (in the form of pitch accents) and prosodic prominence: in English there is a constraint by which the head of every well-formed MiP bears a pitch accent. Secondly, the analysis uses the distribution of MiPs, by transitivity, as evidence for MaP phrasing (which is the primary concern of the paper); the evidence that it is indeed a MiP-level constituent which must obligatorily contain a pitch accent is therefore only indirect.

5.2.2.3 Summary: the MiP as the domain of pitch accent distribution

Both Japanese and Korean are argued to bear boundary tones, but also to enforce a ban on the maximum number of pitch accents (at most one) which may appear in an Accentual Phrase (≈MiP). English is argued to enforce a ban on the minimum number of pitch accents (at least one) that may appear in an MiP. For our present purposes these analyses are relevant because they again rely on pitch accent distribution as evidence of prosodic constituency, and because in Japanese, Korean and (arguably) English, the domain relevant to pitch accent distribution is the MiP.

[91] The account of English pitch accent distribution in Gussenhoven (2004:274ff.) is also based on underlying pitch accents, though for different reasons, and yet also appeals to a relationship between prosodic constituents and pitch accents in accounting for final surface distribution. In Gussenhoven's account English pitch accents are lexical, but are subject to postlexical rhythmic constraints which operate within the domain of the phonological phrase (=MaP, though see Gussenhoven 2004:166-7).

5.2.3 Summary: pitch accent distribution as a reflex of prosodic structure

This section reviewed analyses which appeal to the distribution of pitch accents as evidence of prosodic constituency: a well-formed constituent at some level (MaP or MiP) is characterised by the occurrence of some tonal event. In addition, the fact that the analyses make appeal to constituents at different levels of the prosodic hierarchy, the possibility arises that the domain of pitch accent distribution may vary.

Whilst in some languages the relevant domain is argued to be MaP or MiP, the claim of this thesis is that in EA the relevant domain is the Prosodic Word (PWd). The empirical evidence in support of this claim is explored in detail in section 5.3 and 5.4 below.

5.3 Evidence from prosodic phrasing in EA

Previous sections of this chapter have reviewed evidence and analyses in the literature which suggest that it is possible to analyse pitch accent distribution as a reflex of the distribution of constituents at some level of the Prosodic Hierarchy. This arises due to a constraint or rule requiring each constituent to contain at least one pitch accent[92]. The choice of constituent level proposed for other languages includes MaP and MiP.

This section explores whether or not it is possible to propose one of these phrase level constituents as the domain of pitch accent distribution in EA. The theoretical basis of such an analysis is discussed briefly in section 5.3.1, which is followed by empirical evidence regarding MaP and MiP in EA from a pilot study in sections 5.3.2 and 5.3.3. These sections show that MaPs and MiPs in EA consistently contain multiple PWds, but that all PWds still bear an accent. Evidence of phrasing in the thesis corpus is also reviewed and discussed in section 5.3.4. On the basis of the mismatch between phrase sizes and pitch accent distribution, the hypothesis that a phrase-level constituent is the domain of pitch accent distribution in EA is rejected.

5.3.1 Pitch accent distribution within a phrase-level constituent in EA?

The hypothesis explored in the following sections is whether or not a constraint on the minimum number of pitch accents per constituent can capture the facts of EA pitch accent distribution. Recall however that in European Portuguese and French it was shown that additional pitch accents besides the one obligatory (head-marking) accent were *optional*. EA pitch accent distribution was demonstrated in chapter 3 to be

[92] In Japanese the generalisation was that an AP may contain at *most* one pitch accent.

extremely consistent, with a pitch accent observed on 97% or more of content words. If some of the pitch accents in EA were 'optional' then we would expect to see greater variety in the distribution of accents. To claim that a phrase-level constituent is the domain of pitch accent distribution in EA involves the assumption that in the overwhelming majority of cases these phrases (whether MaP or MiP) contain a single content word.

The possibility that there may be languages in which phonological-phrase-level constituents are co-extensive with a PWd-level constituent is formally predicted in Truckenbrodt's (1999) constraint-based analysis of the distribution of Phonological Phrases ('P-Phrase', ≈MaP). Truckenbrodt proposes that surface phrasing results from the interaction of constraints of various kinds. These include interface constraints, regulating the alignment of prosodic phrases to maximal projections of lexical categories in syntactic structure (lexical XPs), such as ALIGN-XP,R, which is satisfied when the right edges of XPs and MaPs fall in the same place (and ALIGN-XP,L, which similarly regulates alignment of left-edges). In contrast, a constraint specific to prosodic structure, is *P-PHRASE, a member of the *STRUC family of constraints, is best satisfied by outputs containing the least number of PPs possible[93].

A factorial typology of the interaction of these three constraints yields a number of possible ranking permutations including, as Truckenbrodt (1999:228 fn11) points out, the ranking given in (5.27), whereby the optimal phrasing is one in which both right and left edges of lexical XPs coincide with MaP boundaries.

(5.27) ALIGN,XP,R , ALIGN,XP,L >> *P-PHRASE

Under this ranking, both edges of an XP-complement[94] will be phrased separately from its head, yielding a surface phrasing in which every lexical word is phrased into a PP:

(5.28)	$[X \quad [XP]_{XP}]_{XP}$	$[[XP]_{XP} \quad X]_{XP}$	*syntactic structure*
	$(\quad)_{PP} (\quad)_{PP}$	$(\quad)_{PP} (\quad)_{PP}$	*prosodic structure*

[93] Truckenbrodt also discusses the role of EXHAUSTIVITY, which is argued to be undominated, and is thus not pertinent to the present discussion.
[94] Assuming that the complement is always embedded inside the XP with the head.

Truckenbrodt notes that it is difficult to see what sort of evidence could be used to distinguish PPs and PWds in such a language. In many languages however, as we have seen, evidence from cues to phrasing (phrase tones, lengthening and optional pause) and pitch accent distribution provide evidence regarding PP boundaries, and point independently to boundaries which coincide.

A language like EA is an interesting test case for Truckenbrodt's theoretical prediction, which is explored below in EA by comparing the evidence of cues to phrasing and the evidence of pitch accent distribution. On the basis of the phrasing and pitch accent distribution observed in long, complex sentences in EA, the possibility that in EA PWds are co-extensive with MaPs (or MiPs) is below rejected.

The main empirical phrasing evidence presented here was collected and analysed during a pilot study of EA phrasing read speech data (Hellmuth 2004), but is also supplemented with a survey of phrasing generalisations observed in the thesis corpus.

5.3.2 Evidence that the PWd is not co-extensive with MaP in EA.

This section sets out evidence regarding the size of MaPs in EA from a pilot study, and which suggests that MaPs are rarely co-extensive with the PWd. Data which confirm this finding from the thesis corpus (as discussed in chapter 3) are also presented.

Hellmuth (2004) investigated cues to prosodic phrasing and the resulting phrasing generalizations in a corpus of read speech. This comprised a core set of 38 target subject-verb-object (SVO) sentences in which the syntactic complexity of both subject and object are systematically varied[95].

The pilot study design was a language-specific modification of the Romance Languages Database (RLD) (Elordieta et al 2003), a comparative analysis of phrasing tendencies in Romance languages using a parallel database of SVO sentences in which prosodic weight (number of syllables) and syntactic complexity are systematically varied. Elordieta et al found differing sensibilities to the influence of constituent length/syntactic structure in different languages: in Catalan and Spanish a 'default' |S|VO| phrasing predominates, with a phrase break between subject and verb; in

[95] SVO is the most commonly observed word order in EA, and strongly preferred over VSO with imperfect verbs (Benmamoun 2000), which were used in the pilot study dataset.

contrast, in Standard European Portuguese (SEP, Lisbon variety), the overwhelming majority of tokens are |SVO| and contain no sentence-internal phrase break, except in cases of subjects which are both prosodically long and syntactically branching, or objects containing more than one level of syntactic embedding.

In the EA database, as in the RLD, a simplex subject consisted of a head noun only [N], a branching subject was a noun modified by an adjective phrase [AP] or prepositional phrase [PP], and a 'double-branching' subject contained a noun modified by an AP and a PP:

(5.29)	non-branching	branching	double-branching
	$[\ N\ ]_{NP}$	$[\ N\ [\ AP\ ]\ ]_{NP}$	$[\ N\ [\ AP\ ]\ [\ PP\ ]\ ]_{NP}$

As in the RLD, the constituent length of the subject and object was also varied systematically. However, since in EA long words of 5 syllables or more are relatively uncommon, sequences of nouns in the genitive Construct State (CS) were used to create prosodically heavy targets where necessary eg [bint ʕamm] (daughter-aunt) "cousin (f.)". CS sequences have been shown to function as a single syntactic word (Borer 1996) hence these can be used to increase prosodic weight without increasing syntactic complexity. This resulted in a database where increases in prosodic weight correspond with increases in number of prosodic words (unlike the RLD in which number of prosodic words varies with syntactic complexity and increases in prosodic weight are in terms of syllable count only).

Segmental phonological processes have been shown to exhibit 'sandhi' properties and to be sensitive to prosodic phrase boundaries[96]. In order to investigate (non-)application of a rule of epenthesis as a potential cue to phrasing in EA, epenthesis contexts were placed across all potential phrase boundaries. Epenthesis applies systematically in EA to break up sequences of three consecutive consonants, by insertion of an epenthetic vowel between C_2 and C_3: eg /bint gami:la/ → [binti gamiila] 'beautiful girl', and has been reported to apply across word boundaries within a domain larger than the phonological phrase (MaP) in EA and within the phonological phrase (MaP) in other dialects of Arabic (Watson 2002:64). Creation of segmentally parallel epenthesis contexts (C_2=[m]; C_3=[b]) greatly limited the choice of lexical items and it was impossible to

[96] As discussed in chapter 2 section 2.1.5.

vary the number of syllables in the verb since all suitable CC-final verbs in EA are bisyllabic.

In addition to the core dataset, some additional sentences were recorded in which parenthetical expressions (such as [bin-nisba li-l-9amm] 'according to-the-uncle') were inserted into otherwise 'non-branching' targets. These were used to help decide what level of phrasing was being cued, since it has been observed that parenthetical expressions induce a full intonational phrase (IP) boundary at their right edge (Nespor & Vogel 1986, Frota 2000).

An excerpt from the database, showing variation in subject position, in both syntactic complexity & prosodic length, is given in (5.30) below.

(5.30) Sample data from the phrasing study (Hellmuth 2004).
Non-branching object condition, subject condition varied

a	il-film			biyGumm	bint 9amm-i
	the-film			upsets	cousin-my
b	nihaayit-l-film			upsets	cousin-my
	end-the-film			upsets	cousin-my
c	siyaasi	muhimm		biyxumm	balad-na
	politician	important		cheats	country-our
d	il-muhandis-l-mi9maari	l-muhimm		cheats	country-our
	the-architect	the-important		cheats	country-our
e	il-mumassil	l-muhimm	fi-l-film	biyGumm	bint 9amm-i
	the-actor	the-important	in-the-film	upsets	cousin-my
f	il-miGannawaati	l-muhimm	fi-nihaayit-l-film	biyGumm	bint 9amm-i
	the-singer	the-important	in-end-the-film	upsets	cousin-my

The full database was recorded with two female EA speakers (NY and MF). Target sentences were presented typed in Arabic script with EA-specific lexical items and spelling conventions used to encourage speakers to produce dialectal renditions, and reduce potential higher register interference from use of written prompts. Speakers read the sentences three times each at varying speech rates: the first at normal pace, the second slower, and the final repetition the slowest of the three. A total of 38 targets x 2 speakers x 3 repetitions yielded 228 tokens, which together with 3 x 2 additional repetitions of double-branching subject + object targets yielded a total of 234 tokens. Recordings were made in a sound-proof room using ProTools 6.0 on MBox directly to digital format at 44100Hz 16bit, then re-sampled at 22050Hz 16bit. Auditory transcriptions were made of all tokens with reference to an F0 contour and spectrogram using Praat 4.2 (Boersma & Weenink 2004).

Targets produced in a single phrase showed a typical EA declarative intonation contour, as described in chapter 3 (section 3.4.1), in which a pre-nuclear rising pitch accent is associated with the stressed syllable of each prosodic word in the sentence (see Figure 5.8a below). Each successive peak is lower than the previous one, displaying a cross-linguistically common pattern of declination. The peak of the final (nuclear) pitch accent, associated with the last word in each sentence, is usually produced considerably lower than might be expected from the effects of declination alone however. The last word is also lengthened due to its pre-boundary position (cf. Prieto et al 1995, Frota 2000, Chahal 2001).

Where a phrase boundary was inserted, the two speakers used slightly different clusters of cues to mark boundaries. A range of possible cues to phrasing were transcribed: local pitch range reset, local pitch accent lowering (ie a non-sentence-final pitch accent produced lower than might be expected from the effects of downstep), lengthening (of a word to the left of a boundary), failure of epenthesis (no epenthetic vowel inserted), insertion of a pause and a high (continuation) or low phrase tone. Speaker NY used the first four of these and speaker MF the second four:

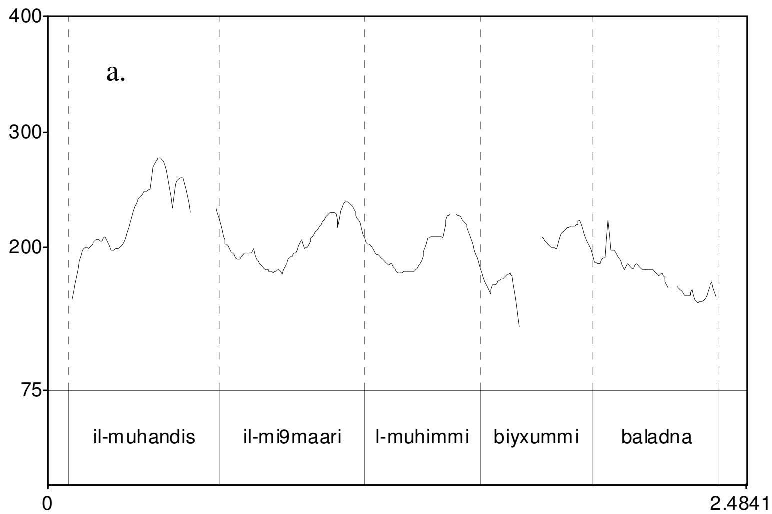
(5.31) Cues to phrasing used in phrasing pilot study (Hellmuth 2004)

Cue	Speaker
local pitch reset	speaker NY
phrase-final pitch accent lowering	speaker NY
pre-boundary lengthening	speaker NY, speaker MF
failure of epenthesis	speaker NY, speaker MF
pause	speaker MF
phrase tone (H-/L-)	speaker MF

As a working hypothesis during transcription, whenever two or more of these cues were observed at a single point this was marked as a phrase boundary. A sample pair of utterances from each speaker, with and without phrase boundaries, are provided in Figures 5.8 and 5.9 below.

Figure 5.8 Sample pair of utterances: speaker NY
Long branching subject + short non-branching object

	il-muhandis	l-mi9maari	l-muhimm	biyxumm	balad-na
	the-architect		the-important	cheats	our-country
a.	(				)
b.	(		)	(	)



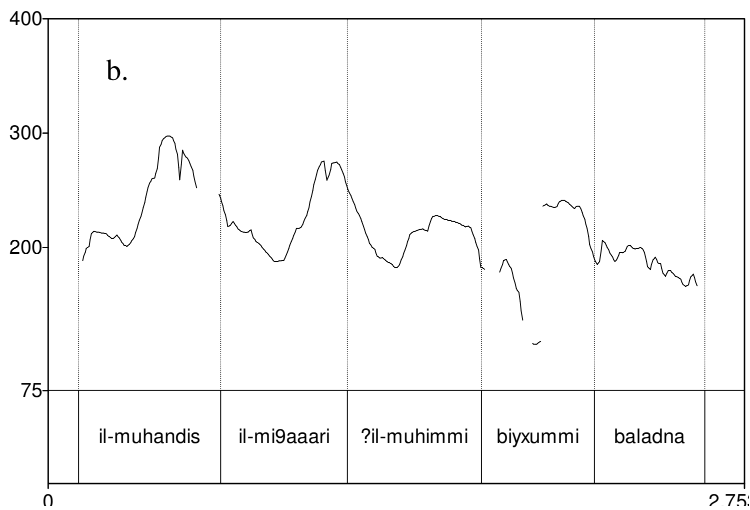
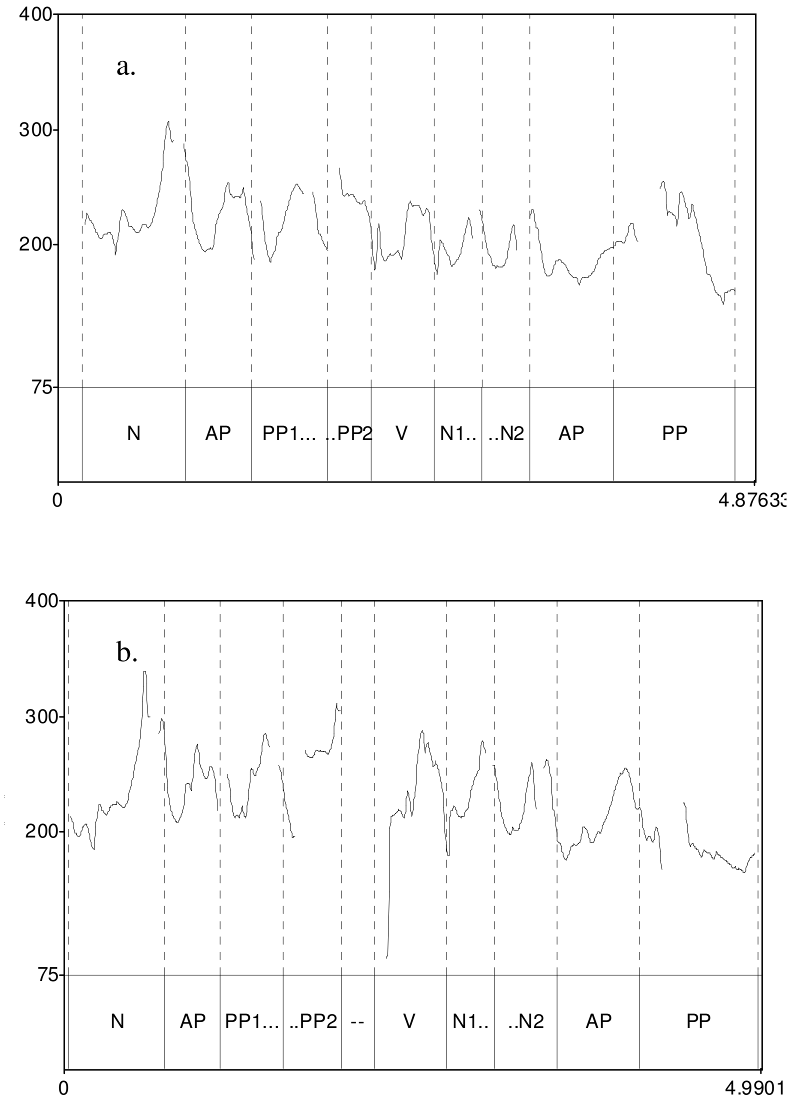


Figure 5.9 Sample pair of utterances: speaker MF
Long double-branching subject + long double-branching object

il-miGannawaati	l-muhimm	fi-nihaayit-il-film	biyGumm…
the-singer	the-important	in-end-the-film	upsets…
… banaat 9ammit-i	l-9ayaniin	min-l-?iskandariyya	
…cousins(f)-my	the-ill	from-Alexandria	

	N	AP	PP	V	N	AP	PP
a.	(						)
b.	(		)	(			)



For the purpose of determining the typical size of MaPs in EA, the crucial question is which level of phrasing within the prosodic hierarchy is being cued here, however infrequently, by the two speakers.

Empirical evidence to support the view that these boundaries mark the edges of MaP level phrases comes from the additional recordings made which contained parenthetical expressions, and from sensitivity of epenthesis to phrasal boundaries. For speaker MF, epenthesis failed across all of her (few) boundaries (3 out of 120), whereas for speaker NY this cue was most frequently observed at the right edge of a parenthetical expression. For NY at least then, failure of epenthesis seems to mark the intonational phrase (IP). Those NY boundaries across which epenthesis applies, but which are nonetheless marked by a cluster of tonal cues, are therefore judged to be MaP boundaries[97].

There is also a theoretical argument that the observed phrase breaks are MaP-level boundaries. On the rare occasions when a phrase break is observed, the break invariably falls between the subject and verb (at the right edge of the subject NP). Treating these as MaP breaks is consistent with Truckenbrodt's (1999) assertion that cross-linguistically the level of phrasing sensitive to syntactic maximal projections is an instantiation of the phonological phrase (notated here as MaP)[98].

Taking clusters of tonal cues to indicate MaP level phrases then, a summary of the observed actual phrasings as shown in (5.32) for normal speech rates and in (5.33) for fast speech rates, indicates that the majority of target SVO sentences are produced in a single, large MaP: 91% of tokens are produced as |SVO|. Those sentence-internal boundaries that do occur fall between the subject and the verb: |S|VO|, and a break in any other position, particularly between verb and object, is never observed: *|SV|O|.

These results indicate clearly that MaPs in EA are large, and certainly not co-extensive with the PWd.

---

[97] The patterns of epenthesis in EA in the pilot study are consistent with those observed in studies of EA word segmentation, in which epenthesis applies across whole sentences (p.c. Rajaa Aquil).

[98] Two unusual NY phrasings in which epenthesis failed after a one-word subject were analysed as topicalisations of the subject, with an IP boundary inserted at the right edge of the subject.

(5.32) MaP phrasing patterns observed at normal speech rates

<table>
<tr><td>fast
subject:</td><td>#PWds:</td><td>non-branching
object</td><td>branching
object</td><td>double-branching
object</td></tr>
<tr><td rowspan="2">non-branching</td><td>1</td><td>|SVO|</td><td>|SVO|</td><td>|SVO|</td></tr>
<tr><td>2</td><td>|SVO|</td><td>|SVO|</td><td>|SVO|</td></tr>
<tr><td rowspan="2">branching</td><td>2</td><td>|SVO|</td><td>|SVO|</td><td>--</td></tr>
<tr><td>3</td><td>|SVO|</td><td>|SVO|</td><td></td></tr>
<tr><td rowspan="2">dbl-branching</td><td>3</td><td>|SVO|</td><td>--</td><td>|SVO|</td></tr>
<tr><td>4</td><td>|SVO|</td><td></td><td>|S|VO|~|SVO|</td></tr>
</table>

(5.33) MaP phrasing patterns observed at slow speech rates

<table>
<tr><td>slow
subject:</td><td>#PWds:</td><td>non-branching
object</td><td>branching
object</td><td>double-branching
object</td></tr>
<tr><td rowspan="2">non-branching</td><td>1</td><td>|SVO|</td><td>|SVO|</td><td>|SVO|</td></tr>
<tr><td>2</td><td>|SVO|</td><td>|SVO|</td><td>|SVO|</td></tr>
<tr><td rowspan="2">branching</td><td>2</td><td>|SVO|</td><td>|SVO|</td><td>--</td></tr>
<tr><td>3</td><td>|S|VO|~|SVO|</td><td>|S|VO|~|SVO|</td><td></td></tr>
<tr><td rowspan="2">dbl-branching</td><td>3</td><td>|S|VO|~|SVO|</td><td>--</td><td>|S|VO|~|SVO|</td></tr>
<tr><td>4</td><td>|S|VO|~|SVO|</td><td></td><td>|S|VO|~|SVO|</td></tr>
</table>

The pilot study used specially constructed sentences, of relatively unnatural complexity, to elicit information about the relative sensitivity of EA MaPs to syntactic and prosodic complexity. The corpus examined in the present thesis (as outlined in chapter 3) did not contain such unnaturally complex sentences; however, a phrasing survey of sentences from the read and re-told narratives collected for chapter 3 reveals that similar cues to phrasing were used by speakers and that the phrasing generalisations observed in the narratives are consistent with the findings of the pilot study. These are reviewed and analysed in section 5.3.4. Before that, the next section (5.3.3) sets out the evidence from the pilot study regarding the MiP phrase level in EA.

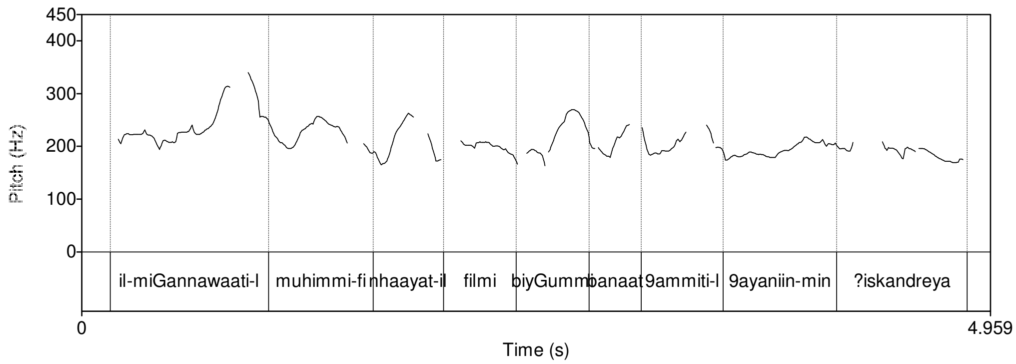
5.3.3 Evidence that the PWd is not co-extensive with MiP in EA.

This section reviews two types of evidence which suggest that MiPs in EA are not co-extensive with the PWd.

Empirical evidence in support of proposing a MiP level of phrasing in EA comes primarily from instances where an MiP boundary appears to be optionally tonally marked. In these cases the pitch accent at the right edge of the MiP shows local final lowering, and is followed by a local pitch reset at the start of the new MiP (to the pitch level of the start of the previous MiP, rather than to the pitch level of the start of the previous MaP). This resembles the ‘rhythmic boost’ pitch peak enlargement observed at the left edge of two-PWd MiPs in Japanese (Kubozono 1993, Shinya et al 2004).

An example of rhythmic boost in EA is illustrated in Figure 5.10. The first four PWds are phrased together into a MaP, marked by final lowering of the pitch accent on MaP-final [film] ‘film’, and local reset of pitch on the next word [biyGumm] ‘upsets’. Within the initial MaP the four PWds are grouped into two pairs of PWds, with a slight local reset of pitch at the start of the first PWd of the second pair ([nihaayit] ‘end’). This local reset is the ‘rhythmic boost’.

Figure 5.10 Sample utterance illustrating ‘rhythmic boost’ effect (*1905 ny3*)



N	AP	PP	
il-miGannawaati	l-muhimm	fi-nihaayit	il-film
the-singer	the-important	in-end-	the-film
V	NP	AP	PP
biyGumm…	… banaat 9ammit-i	l-9ayaniin	min-l-?iskandariyya
upsets…	…cousins(f)-my	the-ill	from-Alexandria

‘The important singer at the end of the film upsets my sick cousins from Alexandria’

N AP PP V N AP PP
|( )$_{MiP}$ ( )$_{MiP}$|$_{MaP}$ |( )$_{MiP}$ ( )$_{MiP}$|$_{MaP}$

There is also theoretical evidence for MiP based on its role as a ‘counting device’ to explain the MaP phrasing patterns discussed in the previous section. The claim is that the sparse MaP phrasing facts can be ascribed to a preference in EA for prosodically branching phrases at both MiP and MaP, and this claim is framed in a constraint-based analysis in the remainder of this section (following Hellmuth 2004).

The most striking aspect of the MaP phrasing generalizations reported in section 5.3.2 above is the lack of phrasing breaks observed in EA, which cannot be explained by reference to sensitivity to morpho-syntactic structure alone: all of the main syntax-phonology interface theories predict a phrasing break between subject and verb in SVO sequences which have an overt object (Nespor & Vogel 1986, Selkirk 1986, Inkelas & Zec 1995, Selkirk 2000, Truckenbrodt 1999).

Importantly however, any explanation of the predominance of |SVO| in EA must be analysed as *interacting* with syntax-phonology interface conditions, since whenever speakers do insert a MaP phrase break it is invariably at the right edge of the subject. As outlined in chapter 2 (section 2.2.2), interface conditions on prosodic structure have been shown to be sensitive to the edges of syntactic maximal projections (Selkirk 1986, Selkirk 2000, Truckenbrodt 1999). The presence of MaP boundaries at the right edge of the subject NP yielding |S|VO|, and the systematic absence of such boundaries at the left edge of the object NP, *|SV|O|, suggest that right-edge sensitivity is at work in EA. This can be expressed as an alignment constraint ALIGNXP,R requiring an MaP for each maximal projection (XP), such that the right edge of the XP coincides with the right edge of the MaP (Selkirk 2000), as shown in (5.34):

(5.34) [[S]$_{NP}$ [V [O]$_{NP}$]$_{VP}$]
( )$_{MaP}$ ( )$_{MaP}$

Nonetheless, |SVO| phrasings predominate, and I suggest that this is due to a phonological well-formedness constraint which is sensitive to the internal structure of lower level prosodic constituents, with a preference for branching (binary) structure. Specifically, in EA, the constraint BINMAP$_{MIP}$ (“A Major Phrase consists of two Minor Phrases”, see example (2.19) in chapter 2) outranks ALIGN XP,R so that MaP phrase breaks fall at the right edge of XPs, but only when all of the resulting phrases are of sufficient prosodic weight. By analogy with observations made for Japanese that MiP can be defined as a node that branches into two words (Kubozono 1993), I propose that the MiP in EA at normal (fast) speech rates is a constituent formed of two prosodic words (PWd): BINMIP: “A minor Phrase is composed of two Prosodic Words”. A MaP that meets the minimal binarity requirement of BINMAP$_{MIP}$ will contain at least two MiPs each of which in turn contains two Prosodic Words; thus a well-formed MaP must contain at least four PWds. This formulation of the phonological well-formedness of MaPs in EA correctly predicts the contexts where |SVO| phrasings are observed at normal rates.

To illustrate the analysis, let us examine sentences with subjects of increasing numbers of prosodic words, produced at normal (fast) speech rates. A subject composed of one, two or three PWds is not 'heavy' enough to form an independent MaP; only four PWds is enough (recall (5.32) above). The tableau in (5.36) below evaluates possible output phrasing options for a sentence with a 2PWd non-branching subject, and motivates the ranking between well-formedness and interface constraints (note that for the sake of space in the tableaux 'PWd' is notated as 'ω')[99]:

(5.35) BinMaP$_{MiP}$ >> Align XP,R

(5.36) Non-branching long S (2PWds) + non-branching long O (2PWds)

[[N]$_{XP}$ [V [N]$_{XP}$]$_{XP}$] ω ω ω ω ω	BinMaP$_{MiP}$	Align XP,R
☞ a. ( ) |S VO|		*
b. ( )( ) |S|VO|	**!	

Sentences with double-branching subjects, as in the tableau in (5.37), show that *all* potential MaPs must be phonologically well-formed: even though the long double-branching subject could be mapped to a single phonologically well-formed MaP, since it contains four PWds, the VP complex is too 'light' since it contains insufficient PWds to form two MiPs:

(5.37) Double-branching long S (4PWds) + non-branching long O (2PWds)

[[N[AP]$_{XP}$ [PP]$_{XP}$]$_{XP}$ [V [N]$_{XP}$]$_{XP}$] ω ω ω ω ω ω ω	BinMaP$_{MiP}$	Align XP,R
☞ a. ( ) |SVO|		**
b. ( )( ) |S|VO|	*	*

In cases with long and/or complex objects, two prosodically well-formed MaPs could be mapped from the sentence *if* the phrase break were between the verb and object. In these cases however an |SVO| phrasing is nonetheless observed, due to the effects of WrapXP which requires each syntactic XP to be contained within a MaP (Truckenbrodt 1995, Truckenbrodt 1999). In (5.38) below, candidates a. and b. are both equally phonologically well-formed with respect to BinMaP$_{MiP}$, and both incur equal alignment

[99] The workings of the evaluation metric in OT are usually illustrated by means of a tableau in which the input form can be compared to various candidate output forms (listed in the lefthand column). Each candidate is evaluated against constraints listed from left to right in columns, in the order of their ranking: highest to lowest. Asterisks in a column denote violation of that constraint by the candidate in question. A '!' symbol indicates a 'fatal' violation, that is, the candidate has fared 'worst' on that constraint, and the fatal violation eliminates it from evaluation against lower ranked constraints.

violations (both have XP right edges without corresponding MaP right edges). Here WRAPXP acts as a tie-breaker, favouring the |SVO| production in which the VP is not divided prosodically but 'wrapped' with the subject in a single MaP:

(5.38) Double-branching short S (3Pwds) + double-branching short O (4PWds)[100]

[[N[AP]$_{XP}$[PP]$_{XP}$]$_{XP}$ [V [N [AP]$_{XP}$[PP]$_{XP}$]$_{XP}$]$_{XP}$] ω ω ω ω ωω ω ω	BINMAP$_{MiP}$	ALIGN XP,R	WRAP XP
☞ a. ( ) |SVO|		**	*
b. ( )( ) |SV|O|		**	***!
c. ( )( ) |S|VO|	*!	*	

If both the subject and the verbal complex are of sufficient prosodic weight to meet the well-formedness condition then the surface (winning) candidate is predicted to be the one that least violates the next most highly-ranked constraint, which is the interface ALIGNXP,R. In (5.39) below, candidate b. has the least alignment violations and is predicted to be the winner. The analysis thus predicts that when all potential MaPs are sufficiently heavy (when BINMAP$_{MiP}$ is unviolated) the only phrasing that will be observed is |S|VO|.

(5.39) Double-branching long S (4Pwds) + double-branching long O (4PWds)

[[N[AP]$_{XP}$[PP]$_{XP}$]$_{XP}$ [V [N [AP]$_{XP}$ [PP]$_{XP}$]$_{XP}$]$_{XP}$] ω ω ω ω ω ωω ω ω	BINMAP$_{MiP}$	ALIGN XP,R
a. ( ) |SVO|		**
☞ b. ( )( ) |S|VO|	*	*

In fact however the production results showed *variation* between |S|VO| and |SVO| in this context (recall 5.32 above). Further investigation is needed with a larger number of speakers to clarify whether the prediction of the analysis is confirmed in the form of a tendency to |S|VO| or whether there is in fact free variation[101]. Crucially however, for our present purposes, in a sentence with a long/complex subject and object, which must by this analysis contain MiPs composed of more than one PWd, even in these cases every PWd is accented (as illustrated in Figure 5.10 above). The MiP thus cannot be the domain of pitch accent distribution in EA.

---

[100] From presently available data there is no way to infer how WRAPXP is ranked with respect to the other constraints (hence it is illustrated separately here at the side of the tableau). Establishing a ranking for WRAP with respect to BINMAP/ALIGN would require targets of considerable, and probably implausible, length (such as a complex NP composed of a 4Pwd N + 4PWd AP or PP, or a 4PWd N-AP + 4PWd PP).

[101] The hypothesis that MiP consists of two PWds has to be revised to account for phrasing patterns at slower rates, where MiP appears to be sensitive to number of *syllables*. See Hellmuth (2004) for discussion, and Jun (2003) for similar sensitivity observed in Korean APs.

5.3.4 Phrasing evidence from the narratives corpus

The last sections showed that in EA prosodic phrases at the MaP level are consistently large, and argued that a preference for binary MiPs can account for sparse MaP phrasing. Neither MaP nor MiP are routinely co-extensive with the PWd in EA. In this section we explore whether these findings are reproduced in the thesis corpus data.

Much of the corpus (described in chapter 3 section 3.2.1) consists of read sentences which are neither prosodically long nor syntactically complex. A survey was thus made of selected portions of the narratives database, chosen because the narrative contains sequences of text without punctuation marks in the written text, to which speakers are at liberty to assign whatever prosodic phrasing they choose. Each speaker read the narrative three times, as well as being asked to re-tell the story later from memory[102]. The second reading of the narrative by each speaker, together with their retold semi-spontaneous rendition, was prosodically transcribed by the author, with reference to F0 track and spectrogram extracted using Praat 4.2 for the corpus survey described in chapter 3. The most common cues to phrasing in this subset of the narratives dataset are boundary tones, such as H- or L-, and pauses, and as discussed above, these are assumed to mark the edges of MaP level constituents.

The phrasing facts observed in this corpus subset are consistent with the findings of section 5.3.2 above. There are examples of moderately long/complex monoclausal sentences which are phrased into a single MaP. Example (5.40) below shows the auditory transcription and phrasing of a 5PWd monoclausal sentence[103].

(5.40) Speakers' read speech phrasings of a 5PWd monoclausal sentence.

speaker	guHa	kaan	Tuul	9umr	-uh	9aayiš	fi	-l-	?ariyaaf
fna	LH*	---	LH*	LH*	<	LH*	<	<	!LH*
fsf	LH*	<	<	LH*	<	LH*	<	<	!LH* L-L%
meh	LH*	<	(LH*)	LH*	<	LH*	<	<	LH* H-
miz	LH*	<	LH*	LH*	<	LH*	<	<	LH* H-
mns	LH*	<	<	LH*	<	LH*	<	<	LH* L-

guHa	kaan	Tuul	9umr	-uh	9aayiš	fi	-l-	?ariyaaf
Guha	was	all	life-	his	living	in	the	villages
[[NP]	AUX	[AdvP		]	[V	[PP		]]]s

'Guha had lived all his life in the countryside.'

[102] The full text of the narrative is provided in Appendix A.13

[103] In examples 5.40-5.53 MaP boundaries are shown enclosed within vertical lines: |MaP|, and a gloss and syntactic analysis is provided; the transcription conventions used are those set out in chapter 3 section 3.2.2).

There is evidence from the phrasing in the narratives corpus to confirm that the right edges of VP-internal XP boundaries do not trigger MaP phrase boundaries. This confirms the finding in Hellmuth (2004) that the effects of the interface constraint ALIGNXP are obscured in EA. For example, (5.42) below shows speakers' read narrative phrasings of a 4PWd monoclausal sentence with VP-internal XPs[104].

(5.41) Read speech phrasings of a complex 4PWd monoclausal sentence.

speaker	?ana	9awz	?awzin	lak	kiilu	bi	balaaš
fna	>	↑LH*	LH*	<	LH*	<	!LH* L-L%
fsf	>	LH*	LH*	<	LH*	<	!LH* L-L%
meh	>	LH*	LH*	<	LH*	<	!LH* L-L%
miz	>	LH*	LH*	<	LH* =	<	LH* L-L%
mns	>	LH*	LH*	<	LH*	<	!LH* L-L%

?ana	9awz	?awzin	lak	kiilu	bi	balaaš	
I	want	I-weigh	for-you	a-kilo	for-	free	
[[NP]	[V	[V	[PP]	[NP]	[PP		$]]_{VP}]_{VP}]_{S}$

'I will weigh you out a kilo for free!'

This treatment of complex XPs is maintained in speakers' spontaneous (retold) narrative productions. For example, (5.43) and (5.44) below show the same sentence, in spontaneous retellings by speakers *fsf* and *meh*, which are phrased into a single MaP.

(5.42) Single MaP phrasing (retold by speaker *fsf*) of a complex clause.

?ana	mumkin	?akuun	baddiik	kiilu	bi	balaaš	
>	LH*	LH*	LH*	H*	<	!LH* L-L%	
I	maybe	I-could	I-give-you	a-kilo	for	free	
[[NP]	ADV	AUX	[V	[NP]	[PP		$]]_{VP}]_{S}$

'I could maybe give you a kilo for free.'

(5.43) Single MaP phrasing (retold by speaker *meh*) of a complex clause.

?eh	ra?y-ak		?addiik	kiilu	mooz	bi	balaaš	
LH*	LH*		→LH*	→LH*	LH*	<	!LH* L-L%	
what	your-opinion		I-give-you	a-kilo	bananas	for	free	
[NP	NP t	[C	[V	[NP	]	[PP		$]]_{VP}]_{CP}]_{S}$

'How about if I give you a kilo for free?'

---

[104] Note that speaker *miz* lengthens the word [kiilu] 'kilo' slightly, which might suggest phrase-final lengthening. However in accordance with the methodology of Hellmuth (2004), a phrase boundary was transcribed only when at least two acoustic or tonal cues to phrasing were observed (see 5.3.2 above).

The results from the narratives corpus are also consistent with the idea that it is a preference for branching prosodic phrases (BINMAP$_{MiP}$) which outweighs the effects of the interface constraint (ALIGNXP). For example, a 5PWd monoclausal sentence, such as in (5.44a) below (reproduced from (5.40) above), cannot be subdivided at the MaP level because the resulting MaPs would contain less than two MiPs (assuming that MiPs must themselves be minimally branching at normal speech rates). The sentence is thus rendered by all speakers in a single MaP[105].

(5.44) Phrasing analysis: two observed renditions of a 5PWd monoclausal sentence

a. |([gúHa]) (ken [Túul] [9úmr-uh]) ([9áayiš] [f-il-?ariyáaf]$_{PWd}$)$_{MiP}$ |$_{MaP}$
b. |([gúHa] ken Tul [9úmr-uh]) ([9áayiš] [f-il-?ariyáaf]$_{PWd}$)$_{MiP}$ |$_{MaP}$
Guha was all life-his living in-the-country
'Guha had lived all his life in the countryside.'

In (5.44) there is slight variation in pitch accent distribution between different speakers: *fna* and *miz* accent the word [Tuul] 'all', whereas *fsf* and *mns* leave it unaccented (there is a possible accent on the word in speaker *meh*'s rendition)[106]. In both cases there is insufficient prosodic material to form more than one well-formed MaP.

The narratives also contain multiclausal sentences, which show that short clauses may be phrased independently. The example in (5.45) below involves a short introduction to a fragment of reported speech , phrased into two MaPs of fewer than 4 PWds each by three out of five speakers, in violation of the constraint BINMAP$_{MiP}$.

(5.45) Example of a sequence of very short clauses (introducing reported speech). ('xxx' indicates that a word was omitted.)

	guHa	?aal-luh	(laa?)	bi-	talaata	SaaG
fna2	LH*	LH* H-	xxx	>	LH*	LH* L-L%
fsf2	LH*	LH* H-	xxx	>	LH*	LH* L-L%
meh2	LH*	LH*	LH* L-	>	LH*	LH* L-L%
miz2	LH*	LH*	LH*	<	LH*	!LH* L-L%
mns2	LH*	LH*	xxx	>	LH*	!LH* L-L%

guHa	?aal-luh	laa?	bi-	talaata	SaaG
Guha	said-him	no	for-	three	piastres
[[NP]	[V	[NP]]$_{VP}$]$_{S}$	[[PP		]]

'Guha said to him: "No, three piastres!"'

---

[105] Pitch accents are indicated with an acute accent mark eg [Túul]; PWd boundaries are indicated by square brackets [PWd], MiP boundaries by round brackets (MiP), and MaP boundaries by vertical lines |MaP|. I assume that PWds align to the right edge of an MiP. This is explored further in chapter 6 (6.2.4).
[106] The relative accentability of modifiers such as /Tuul/ is discussed in chapter 3.3.2.

The option of phrasing such short fragments into individual MaPs is, I suggest, due to the influence of an interface constraint regarding mapping of clauses to Intonational Phrases (IPs). Any string which maps to an IP, must also map to at least one MaP, by virtue of the constraint HEADEDNESS, which requires a constituent at one level of the prosodic hierarchy to dominate at least one constituent of the next level down, and is assumed by some authors to be unviolated (e.g. Selkirk 1996). In (5.45) above then, the mid-utterance boundary should be analysed as an IP (as well as MaP) boundary.

As discussed above in section 5.2.3, the distinction MaP and IP boundaries was established on the basis of evidence from non-application of epenthesis across IP boundaries. Chahal (2001) distinguished between two phrase levels in Lebanese Arabic (LA) on the basis of gradient, though statistically discernible phonetic correlates to phrasing. These took the form of increased final lengthening and earlier peak delay at higher boundaries, according to the level of constituent boundary marked; final lengthening was greater in IP-final words than in MaP final words, for example. It is beyond the scope of the present investigation to establish this gradient distinction empirically in EA, and thus, although the auditory transcription identified these inter-clause boundaries as MaP boundaries they are now re-analysed on theoretical grounds as IP boundaries, as in (5.46) below[107].

(5.46) Phrasing analysis: observed renditions of a sequence of short clauses.

a)	$_{\mathbf{IP}}(\ _{\mathrm{MaP}}\vert\ _{\mathrm{MiP}}(\ _{\mathrm{PWd}}$[guHa]	[?aal-luh])	(bi-	[talaata]	[SaaG]$_{\mathrm{PWd}}\ )_{\mathrm{MiP}}\ \vert_{\mathrm{MaP}}\ )_{\mathbf{IP}}$
b)	$_{\mathbf{IP}}(\ _{\mathrm{MaP}}\vert\ _{\mathrm{MiP}}(\ _{\mathrm{PWd}}$[guHa]	[?aal-luh]$)\vert)_{\mathbf{IP}}$	$_{\mathbf{IP}}(\vert($bi-	[talaata]	[SaaG]$_{\mathrm{PWd}}\ )_{\mathrm{MiP}}\ \vert_{\mathrm{MaP}}\ )_{\mathbf{IP}}$
	Guha	said-to-him	for-	three	piastres

'Guha said to him: "No, three piastres!"'

This section has shown that the claims set out regarding EA phrasing in sections 5.2.2 and 5.2.3 hold in the thesis corpus: prosodic well-formedness constraints outweigh interface alignment constraints (BINMAP$_{\mathrm{MiP}}$ >> ALIGN XP,R) and MaP-binarity involves sequences of 4 PWds or more. Crucially, these facts are consistent with rejection of MaP and MiP as potential domains of pitch accent distribution in EA, since even in the largest MaPs in the narratives every PWd bears a pitch accent. The next section (5.4) explores accentuation of function and content words in order to determine the relationship of PWd distribution to morphosyntactic categories.

---

[107] This re-analysis is consistent with the notion that parentheticals and other non-root sentences form an independent IP (Nespor & Vogel 1986, Selkirk 2005a).

5.4 Pitch accent distribution within the Prosodic Word in EA?

The results of the corpus survey in chapter 3 revealed that there was a pitch accent on (almost) every content word in EA. This section seeks to establish on empirical grounds whether that generalisation - ‘an accent on every content word’ - can be reformulated as ‘an accent on every Prosodic Word’.

The previous section set out empirical and theoretical evidence to suggest that the domain of pitch accent distribution in EA is not a phrase-level constituent (neither MaP nor MiP), since these constituents are shown consistently to be composed of more than one PWd. The most likely remaining candidate is therefore the next level down in the Prosodic Hierarchy: the Prosodic Word (PWd).

This section therefore explores the treatment of *function* words in EA in order to determine whether the distribution of pitch accents in EA is defined over a morphosyntactic category (content words) or a prosodic category (Prosodic Word). The null hypothesis, following the hypothesis that phonological processes may refer directly only to prosodic categories[108], is that the correct generalisation in EA is indeed definable in terms of a prosodic category, that is, the Prosodic Word. The way in which function words are treated should help clarify if this is the correct generalisation for EA.

5.4.1 Treatment of function words in the corpus.

Function words in the dataset fall into three groups: i) function words which are always unaccented, ii) function words which are usually unaccented, and, iii) function words which are usually accented. A summary list is provided in the table in (5.47) below.

[108] This assumption has been termed the Indirect Reference Hypothesis (Selkirk 1986, Inkelas & Zec 1990, Truckenbrodt 1999).

(5.47) Treatment of function words in the corpus

a. always unaccented		b. usually unaccented (varies)		c. usually accented	
[il]	the	[min]	from	[?inn-uh]	that-it (comp.)
		[9ala]	on/at	[?inn-ak]	that-you (comp.)
[fi]	in				
[bi]	with/by	[fii]	there is	[9alašaan]	in order to
[ma]	[...] as	[miš]	not	[lissa]	just/soon
		[wala]	or		
[ya]	vocative	[bass]	only	[9alayya]	on/at me
[maa]	negative			[9alayk]	on/at you
		[illi]	that (rel. pron.)	[9alayh]	on/at him
[law]	if	[?inna]	that (comp.)		
[?aw]	or	[da]	that (deictic) m.		
[wa]	and	[di]	that (deictic) f.		
		[la?]	no		
		[?awya]	yes		
		[Gayr]	except		
		[9ašaan]	in order to		
		[bayn]	between		
		[taHt]	under		
		[ba9d]	after		
		[bitaa9]	belonging to (s.)		
		[bitu9]	belonging to (pl.)		
		[?ana]	I		
		[?inta]	you (m)		
		[?iHna]	we		
		[huwwa]	he		
		[hiyya]	she		
		[kaan]	he was (aux. vb.)		
		[kaanit]	she was (aux. vb.)		
		[kunt]	I was (aux. vb.)		

5.4.1.1 Function words which are always unaccented.

Function words which are never accented include the definite article [il] 'the', the prepositions [bi] 'with/by' and [fi] 'in', and the conjunctive particle [ma] 'as' which renders a preposition into an adverbial conjunction, as in the phrase [ba9d ma X] 'after X' (where X is a clause)[109]. Also unaccented are the vocative particle [ya] (eg [ya guHa] 'hey Guha!') and the negative particle [ma] 'not' (the first part of the discontinuous negative circumfix [ma + š]), and also connectives [wa] 'and', [law] 'if' and [?aw] 'or'.

[109] Salib (1981:311)

Some of the function words which are always unaccented are indeed listed by Watson (2002:93) as ‘unstressable morphemes’, including [il] ‘the’ and [wa] ‘and’. Mitchell (1990:127-8) similarly notes that there are particles which are not themselves stressed but which are usually treated as part of the following noun or verb without affecting accent, and that these include prepositions [bi] ‘with/by’, [fi] ‘in’ and the clausal complementizer [?an] ‘to’.

There is however a property that all of the function words which are never accented share, which is that they are prosodically subminimal. The minimal word in EA is obligatorily bimoraic, in order to form a stress foot which in EA is a moraic trochee (Broselow 1976, Hayes 1995, Watson 2002, for word minimality in general see McCarthy & Prince 1995, Downing 2006). In addition in EA consonant extrametricality applies, so that in a monosyllable a singleton coda does not render the syllable heavy (Hayes 1995, Watson 2002).

There are no surface violations of minimality whatsoever in EA[110], such that subminimal cognate words in other dialects, as also subminimal loanwords, are invariably prosodically enhanced in EA, via gemination, vowel lengthening or epenthesis. For example, whereas other spoken dialects tolerate subminimal words of Classical Arabic origin such as /?ab/ ‘father’ and /?ax/ ‘brother’, in EA these words are expanded when pronounced in isolation by gemination: [?abb], [?axx]. Similar repair processes apply to commonly used subminimal function words, so that /kam/ ‘how many?’ emerges in EA as [kaam], /man/ ‘who?’ as [mi:n], and /ma9/ ‘with’ as [ma9a] (Watson 2002:88-9)[111].

All of the function words in the corpus which are invariably unaccented are subminimal words, and so arguably are not PWds.

5.4.1.2 Function words which are usually unaccented.

Function words which are unaccented in most cases but not all and which are also subminimal include the preposition [min] ‘from’, the deictics ([da]/[di] ‘that’ m./f.) and

[110] Hayes (1995:87) termed this an absolute ban on ‘degenerate feet’.

[111] In theory such repairs processes could apply to some of the function words listed in (5.47a) above ([ya] could lengthen to [yaa] for example), however no such examples were observed in the present corpus.

the negation particle [miš] "not" (the stand-alone merged continuous form of the discontinuous negative circumfix [ma + š]).

There are no obvious properties to link the contexts in which these words appear accented. In some cases the context may be slightly more emphatic, particularly in the case of the negation particle and the deictics. It seems that the 'default' status of these words is that they are unaccented, but they can on occasion be 'promoted' to accent status, if the speaker wishes. In the case of the preposition [min] 'from', this is only accented when it forms a stress foot by cliticisation to adjacent segmental material, as for example when there is a pitch accent on [min] when followed by the definite article in certain cases; for example, in the sentence in (5.48) (from the align corpus), the sequence [min-is] 'from the'[112], was accented in 4 out of 18 tokens.

(5.48)

HaSalit	9ala-minHa	min-is	sifaara	9ala-šaan	tiruuH	tidris	fi	?amriika
obtained	at-grant	from-the	embassy	in-order	she-goes	she-studies	in	America

'She got a grant from the embassy to go and study in America.'

Function words which are prosodically of sufficient size to form a stress foot and thus are inherently stressable, but are nonetheless usually unaccented, include the 'pseudo-verb' [fii] 'there is', the complementizer [?inna] 'that', the relative pronoun [?illi] 'that', all forms of the auxiliary verb [kaan] 'to be' and all pronouns. Similarly prepositions which are prosodically large enough and yet are usually unaccented include: [Gayr] 'except', [bayn] 'between', [9ašaan] 'in order to', [taHt] 'under' and [9ala] 'on/at'.

5.4.1.3 Function words which are usually accented.

Function words which are usually accented include inflected forms of complementizers and prepositions such as [?inn-uh] 'that-he' and [?inn-ak] 'that-you', as well as [9alayya] 'at me', [9alayk] 'at you' and [9alayh] 'at him'. These are all of sufficient prosodic size to form fully stressable words. The additional property that these words share however is that they also incorporate pronominalised arguments to the verb, which may make them more prone to bearing an accent. This matches the observation by Mitchell (1990:127-8) that inflected particles are regularly accented.

[112] The [l] of the definite article [il] assimilates completely to a following word-initial coronal consonant in EA (Watson 2002).

Similarly the prepositional phrase [9ala šaan] ‘in order to’ is routinely accented on the word [šaan] ‘cause/condition’, in contrast to its lexically equivalent abbreviated counterpart [9ašaan] ‘in order to’, which is usually (though not always) unaccented. This suggests that [9ala šaan] is indeed analysed by speakers as a full prepositional phrase comprising the preposition [9ala] plus the lexical word [šaan], whereas the foreshortened version [9ašaan] is analysed by speakers as a function word.

5.4.1.4 Summary: accentuation of function words in the corpus

In summary therefore, function words in EA which are never accented are all prosodically subminimal, but even potentially stressable polymoraic function words are usually unaccented in neutral contexts. Those function words which are usually accented are not only polymoraic but also inflected and thus arguably incorporate a lexical head.

(5.49) Accentuation of function words in EA: summary

	accented if subminimal	accented if bimoraic
uninflected function words	✗	✗ optionally ‘promotable’
inflected function words	✗	✓
lexical (content) words	✓	✓

5.4.2 The prosodic realisation of function words

Selkirk (1996) offers an analysis of the prosodic realisation of weak monosyllabic function words in English, which alternate between strong forms, containing full stressed vowels, and weak forms, containing reduced unstressed vowels. She argues on empirical grounds that the correct analysis of English non-phrase-final weak monosyllabic function words in English is as free clitics and offers a constraint-based analysis.

A key claim that Selkirk wants to make is that interface constraints between morphosyntactic and prosodic structure make no reference at all to functional categories in morphosyntax, but only to lexical categories (Selkirk 1996:191). The relevant alignment constraints at the word level are thus of two kinds only, governing alignment of lexical word edges to PWd edges, and of PWd edges to lexical word edges[113]:

[113] The interface constraints take the form of alignment constraints (McCarthy & Prince 1993).

(5.50) word alignment: ALIGN(LEX,L; PWD,L) or ALIGN(LEX,R; PWD,R)
PWd alignment: ALIGN(PWD,L; LEX,L) or ALIGN(PWD,R; LEX,R)

If both ALIGN (PWD,R; LEX,R) and ALIGN(PWD,L; LEX,L) are respected no function word will ever be mapped to a PWd in output representation. As Selkirk points out: "the [PWD:LEXWD] constraints form part of the explanation for the fact that function words typically do not have the status of PWd" (Selkirk 1996:192). For practical purposes in the current discussion each left/right edge sensitive pair of constraints (ALIGN(LEX,L; PWD,L) and ALIGN(LEX,R; PWD,R)) will be conflated into a single constraint:

(5.51) LEXWD:PWD A lexical word maps to a PWd[114].
PWD:LEXWD A PWd maps to a lexical word.

Surface counterexamples to either of these constraints are an indication of minimal violation of the constraint. Selkirk proposes four possible prosodic structures for function words, each of which minimally violates either one or more interface constraints or one of the prosodic domination constraints that make up the Strict Layer Hypothesis (for formal definitions see chapter 2 section 2.2.2). These structures are illustrated in (5.52) below.

A structure in which a function word maps to a full PWd violates HEADEDNESS since an unstressed function word (which in English must be analysed as an unfooted syllable) is dominated directly by a PWd. The structure for function words as free clitics violates EXHAUSTIVITY at the MaP level (the MaP dominates a prosodic category not of the level immediately below it in the hierarchy; hence $\text{EXHAUSTIVITY}_{\text{MAP}}$).

The structure proposed for affixal clitics violates NONRECURSIVITY at the PWd level (the recursive structure involves domination of a PWd by a PWd; hence $\text{NONRECURSIVITY}_{\text{PWD}}$). Finally, the structure proposed for internal clitics violates one half of the pair of LEXWD:PWD and PWD:LEXWD constraints (since only one edge of the lexical words aligns fully to the edge of the PWd, and only one edge of the PWd aligns fully to the edge of the lexical word).

[114] This constraint is equivalent to MCAT=PCAT in McCarthy & Prince (1993).

(5.52)

i) MaP

PWd PWd

((*fnc*)$_{\text{PWd}}$ (*lex*)$_{\text{PWd}}$)$_{\text{MaP}}$

'PWd' structure violates HEADEDNESS (a PWd dominates a σ)

ii) MaP

PWd

(*fnc* (*lex*)$_{\text{PWd}}$)$_{\text{MaP}}$

'free clitic' structure violates EXHAUSTIVITY$_{\text{MaP}}$ (a MaP directly dominates a σ)[115]

iii) MaP

PWd

((*fnc* *lex*)$_{\text{PWd}}$)$_{\text{MaP}}$

'internal clitic' structure violates LEXWD:PWD(L) and PWD:LEXWD(L)

iv) MaP

PWd

PWd

((*fnc* (*lex*)$_{\text{PWd}}$)$_{\text{PWd}}$)$_{\text{MaP}}$

'affixal clitic' violates NONRECURSIVITY$_{\text{PWD}}$ (a PWd dominates a PWd)

If in English the free clitic structure is tolerated, it follows that the constraint which it violates (EXHAUSTIVITY) must be outranked in English by some other constraint. The various structures proposed by Selkirk (1996) for English yield the following ranking:

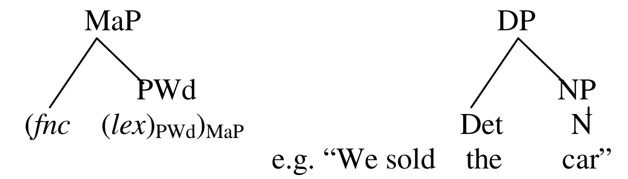
(5.53) LEXWD:PWD, NONRECURSIVITY >> PWD:LEXWD, EXHAUSTIVITY

In keeping with Seklirk's claim that interface constraints "make no reference to functional categories at all" (Selkirk 1996:191) then, the prosodic structures proposed for English closely mirror the relevant morphosyntactic structure. For example the free clitic analysis proposed for the majority of function words is consistent with a syntactic analysis involving functional projections.

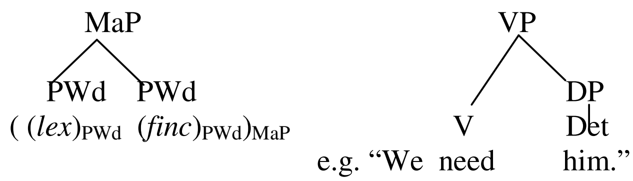
[115] This is just one potential configuration of the position of the free clitic within the MaP; the issue of the direction of cliticisation of function words in EA (pro- vs. en-clisis) is discussed in section 5.4.4 below.

(5.54) Prosodic analysis of English weak vs strong pronouns (Selkirk 1996:206)

i) weak function words [Fnc] = 'free clitic'



ii) phrase-final function words [Fnc] = PWd



In Selkirk's view then, prosodic structure mirrors morphosyntactic structure. An alternative view of the status of function words has been proposed by Zec (2002), on the basis of empirical evidence from (Standard) Serbian and other NeoŠtokavian dialects[116].

Zec argues that there are function words in Serbian which pattern identically in their morphosyntactic distribution, but have different prosodic realisations (bound vs. free function words). The Serbian data is relevant to an analysis of EA function words for two reasons: firstly, because the prosodic realisation of function words in Serbian is sensitive to prosodic minimality conditions, and secondly, because Serbian is a lexical pitch accent language in which every PWd obligatorily bears a pitch accent (Zec 2002).

Zec's diagnostic for PWd status in Serbian is whether or not a word bears a pitch accent. The basic facts of Serbian are that disyllabic function words are usually accented whereas monosyllabic function words are not; compare unaccented monosyllabic [naš] 'our (masc.)' with accented disyllabic [naše] 'our (neuter)' as in (5.55) below (Zec 2002:6):

[116] Standard Serbian equates to former Eastern Serbo-Croat, which is referred to as NeoŠtokavian dialect 2 (NŠ2) in Selkirk (1996), whilst the dialect that Zec refers to as Herzogovian equates to Selkirk's NŠ1.

(5.55) This blue building is …

i) [[na ▪ še]$_{PWd}$ [pózorište]$_{PWd}$]$_{PPhr}$ ‘our theatre’
our(neut) theatre

ii) [naš [stûdio]$_{PWd}$]$_{PPhr}$ ‘our studio’
our(masc) studio

Zec analyses Serbian monosyllabic free function words, such as [naš] ‘our (masc.)’, as ‘free clitics’ using the structure that Selkirk proposes for all English weak function words; the function word falls outside the PWd of the following lexical word, but within the same higher level phrasal constituent (which Zec notates as PPhrase)[117].

In order to account for the fact that monosyllabic free function words are not accented Zec appeals to a PWd minimality constraint:

(5.56) PWDSIZE: A PWd is minimally disyllabic. Zec (2002)

In Serbian however all free function words of sufficient prosodic size are accented (and thus have PWd status, under Zec’s assumption that the diagnostic for PWd status is accentuation), so Zec proposes a constraint whereby *all* morphosyntactic words are mapped to PWds:

(5.57) MWD:PWD A morphological word maps to a PWd[118].

In contrast to function words, Serbian lexical words of any prosodic size are accented, and therefore have PWd status, even if subminimal, as illustrated in (5.58): the lexical word [nov] is accented regardless of the fact that it is monosyllabic.

(5.58) This blue building is …

i) [[no ▪ vo]$_{PWd}$ [pózorište]$_{PWd}$]$_{PPhr}$ ‘(a) new theatre’
new(neut) theatre

ii) [[no ▪ v]$_{PWd}$ [stûdio]$_{PWd}$]$_{PPhr}$ ‘(a) new studio’
new(masc) studio (Zec 2002:7)

---

[117] Zec also explores the fact that even when accented a polysyllabic function word is ineligible to bear phrasal stress, and that function words of any prosodic size may bear a pitch accent if focussed; these topics are however peripheral to our present purposes and are not pursued here.

[118] Zec in fact proposes a pair of constraints requiring alignment of the left and right edges respectively of the MWord to the PWord: Align (MWd L/R, PWd L/R). This detail is not however relevant for the present purposes and thus I have conflated the two constraints for ease of exposition.

To explain this fact Zec additionally appeals to the same LEXWD:PWD constraint as Selkirk, and assumes it to be undominated in Serbian. By ranking PWDSIZE between LEXWD:PWD and MWD:PWD the basic facts of Serbian are captured[119]:

(5.59) LEXWD:PWD >> PWDSIZE >> MWD:PWD
(monosyllabic function words do not attain PWd status)
(monosyllabic lexical words are invariably granted PWd status)

5.4.3 Analysis of the accentuation of function words in EA.

Recall that in EA the basic facts are as follows: function words in EA which are never accented are all prosodically subminimal, and even potentially stressable polymoraic function words are usually unaccented in neutral contexts. Those function words which are usually accented are not only polymoraic but also inflected and thus arguably incorporate a lexical head. These facts are summarised in the table in (5.60) below (repeated from 5.49 above).

(5.60) Accentuation of function words in EA: summary

	accented if subminimal	accented if bimoraic
uninflected function words	✗	✗ optionally ‘promoted’
inflected function words	✗	✓
lexical (content) words	✓	✓

There is obviously a role in EA for a prosodic minimality constraint, since accentuation of function words shows sensitivity to prosodic size. However the facts are slightly different from those observed in Serbian, in which all function words are accented, provided they are of sufficient prosodic size (disyllabic) whilst all content words are accented, even if subminimal (hence: Zec’s proposes LEXWD:PWD >> PWDSIZE >> MWD:PWD). Since in EA word minimality equates straightforwardly to foot bimoraicity[120] I adopt the constraint FTBIN (McCarthy & Prince 1993, Yip 2002):

(5.61) FTBIN Feet must be binary under syllabic or moraic analysis.

Since, as already mentioned, there are no surface violations of minimality in EA whatsoever (Watson 2002:88-9), FTBIN is assumed to be undominated in EA.

[119] Zec goes on to treat dialectal variation between NŠ dialects and to reject Selkirk’s claim that prosodic structure always directly reflects morphosyntactic structure. These matters are not however pursued further here since they are not directly related to the goal of the chapter.
[120] See Downing (2006) for a different view.

In EA it is only lexically inflected function words that are regularly accented; uninflected function words, even if bimoraic, are usually unaccented. In EA then, there appears to be no role for MWD:PWD[121]. Assuming that inflection of function words involves syntactic incorporation to the functional head of a pronominalised lexical head, then, following the thread of Zec's analysis, there is a different ranking in EA than in Serbian:

(5.62)	EA	FTBIN >> LEXWD:PWD (there are no surface violations of word minimality)
	Serbian	LEXWD:PWD >> PWDSIZE (monosyllabic lexical words are accented)

Another key difference between Serbian and EA is that in EA function words are accented only if promoted to PWd status, even if they are prosodically large enough to form a well-formed PWd. In contrast, in Serbian all viable function words are accented. This suggests that there is a role in EA for Selkirk's constraint PWD:LEXWD, which penalises instances of PWds mapped from non-lexical morphosyntactic categories.

Crucially, the fact that a function word is only accented if promoted to PWd status suggests that the correct generalisation for EA is that there is an accent on every PWd.

5.4.4 The prosodic realisation of function words in EA.

This section argues for a 'free clitic' analysis of unaccented function words in EA, based largely on empirical evidence to exclude the other structures proposed by Selkirk (1996) for function words: I will show that unaccented function words do not attain independent PWd status, nor are they incorporated into the PWd with a lexical word as an internal or affixal clitic.

What is the diagnostic for PWd status in EA? Setting aside the claim of this thesis, that in fact accentuation is a diagnostic of PWd status, another clear diagnostic is available, since as in English, the domain of stress assignment in EA is the PWd (Watson 2002). Indeed there appear to be no other phonological processes which apply in EA within the PWd domain; most apply instead either within a MaP type phrase-level domain (Watson 2002), or within the whole utterance (El Zarka 1997, Hellmuth 2004).

[121] Either this constraint is outranked in EA (perhaps by a *STRUC type constraint such as *PWD) or its role in Serbian should be re-evaluated.

Much of the empirical evidence in favour of a free clitic analysis for EA function words (when unaccented) thus comes from the facts of word stress. In particular, unaccented words are shown not to have the status of PWd by virtue of the fact that they undergo a process of unstressed vowel shortening (USVS) (Watson 2002:226-7).

As an example, compare accented [Túul] 'all' vs. unaccented [Tul] in (5.63a) vs (5.63b) below (reproduced from 5.44 above). Pitch tracks and spectrograms for a sample of each type of production of this sentence are provided in Figures 5.11 and 5.12 below.

(5.63) Phrasing analysis: two observed renditions of a 5PWd monoclausal sentence

a) |([gúHa]) (ken [Túul] [9úmr-uh]) ([9áayiš] [f-il-?ariyáaf]$_{PWd}$)$_{MiP}$ |$_{MaP}$
b) |([gúHa]) (ken Tul [9úmr-uh]) ([9áayiš] [f-il-?ariyáaf]$_{PWd}$)$_{MiP}$ |$_{MaP}$
Goha was all life-his living in-the-country

The fact that an unaccented word does not bear word stress in EA excludes the possibility of analysing it as an independent PWd.

Since function words are almost always unaccented there are no minimal pairs to demonstrate that unaccented function words undergo USVS. However, Watson (Watson 2002:226-7) notes that when the long vowel in question is a mid vowel [ee] or [oo][122], after USVS the resulting short vowel is raised: [ee]> [i]; [oo] > [u]. In instances of the function word /Gayr/ 'except', which is almost invariably unaccented, the mid vowel [Geer] is both shortened and raised resulting in [Gir]. An example of this is illustrated in Figure 5.13 below.

[122] Words containing [ee] and [oo] in EA are cognate with Classical Arabic words containing diphthongs [ay] and [aw] respectively

Figure 5.11 Accented /Tuul/ is produced with a long vowel: [Tuul] (*fna2*)

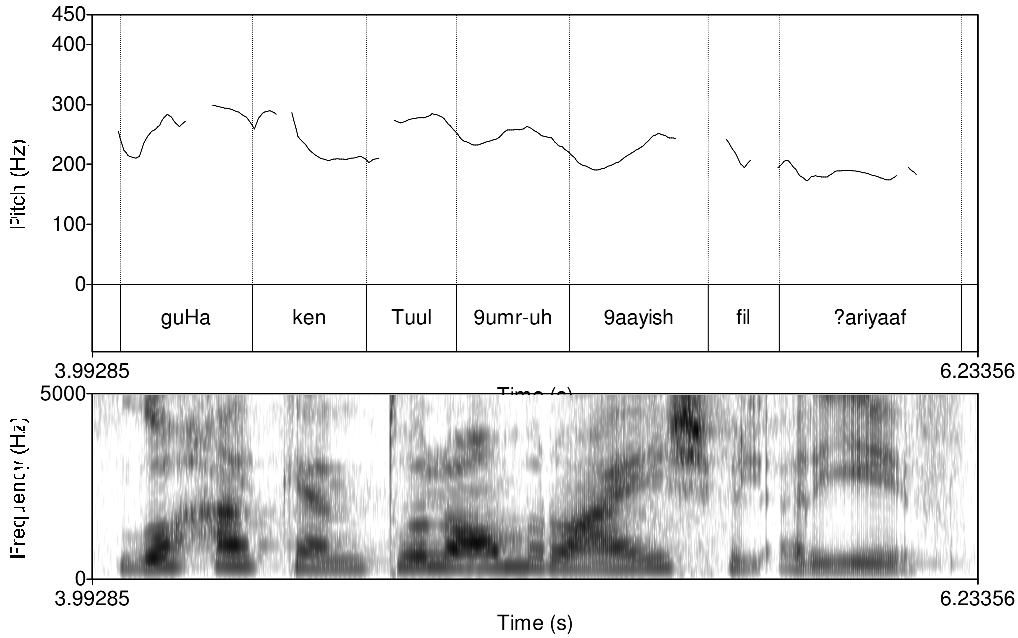


Figure 5.12 Unaccented /Tuul/ shows vowel shortening: [Tul] (*fsf2*)

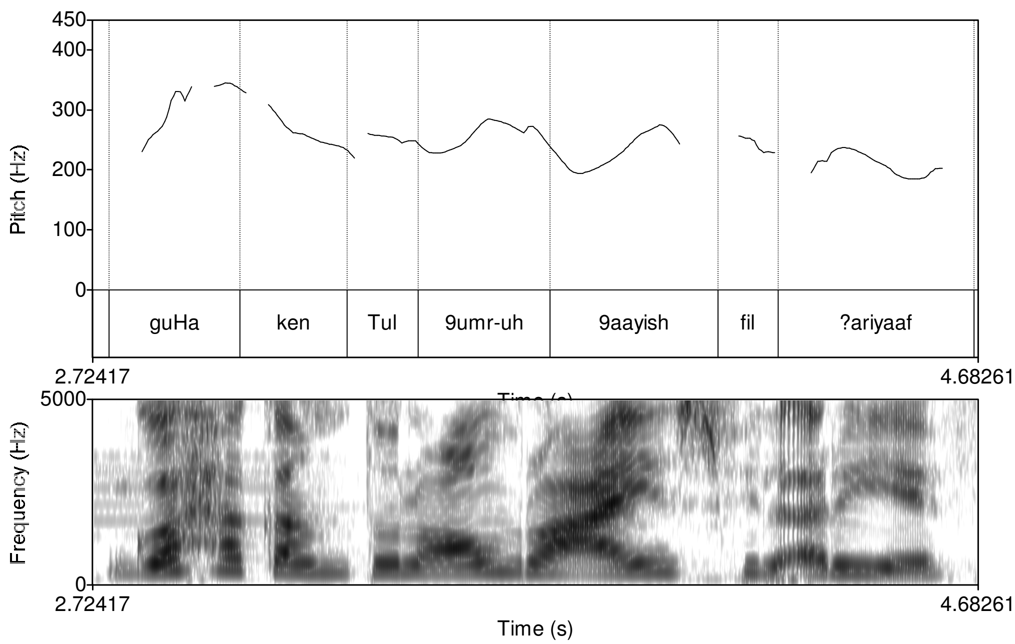
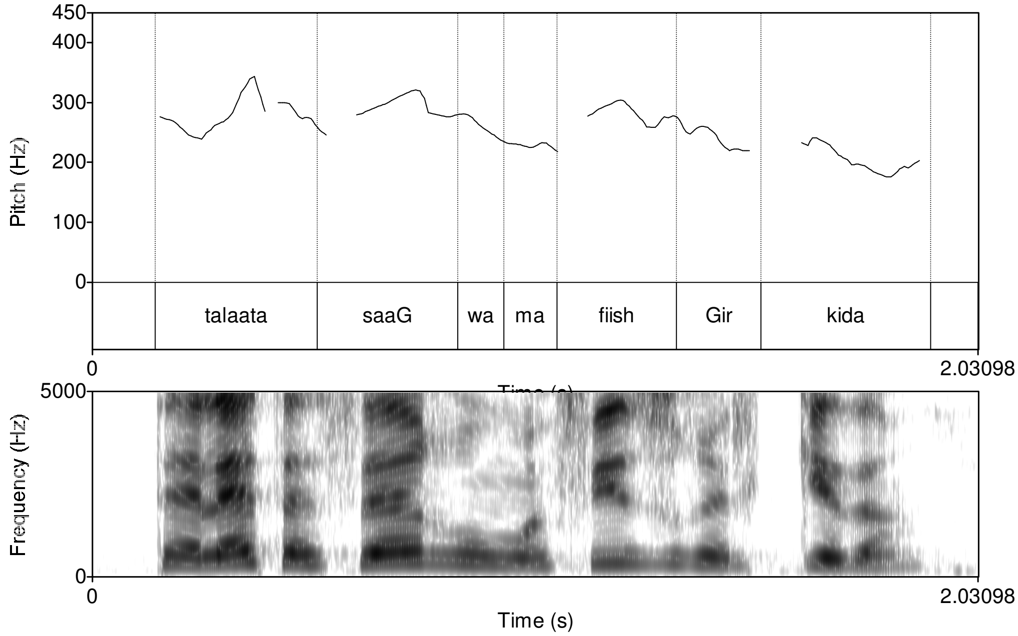


Figure 5.13 Unaccented /Gayr/ shows vowel shortening and raising: [Gir] (*fsf2*)



talaata	SaaG	wa	ma	fiiš	Gayr	kida
three	piastres	and	NEG	there-is	except	that

‘Three piastres, and not a penny more’

In order to exclude the possibility that unaccented function words are incorporated into the PWd with a lexical word (either as an ‘internal’ or ‘affixal’ clitic) the evidence is again from stress assignment. Affixes which are fully incorporated into the PWd induce stress-shift in EA. An example is pronominal suffixes: [bágara] ‘cow’ ~ [bagárt-i] ‘my cow’ (compare non-stress-shifting affixes in Palestinian Arabic in example 4.1 in chapter 4). In contrast unaccented function words do not induce stress-shift in adjacent lexical words in EA, and thus cannot be analysed as incorporated into a preceding or following PWd; that is, they are neither affixal nor internal clitics. The examples in (5.64) show that unaccented function words do not induce rightward stress-shift; those in (5.65) show that although a trisyllabic word composed of three open syllables (CVCVCV) in EA is stressed on the first syllable, stress does not shift leftwards onto a CV monosyllabic function word pro-cliticised to a CVCV disyllable, as in (5.65b) taken from the narratives corpus.

(5.64) a. taláata SáaG wa ma fííš Geer kída (three piastres and no more)
b. * taláata SáGwa...
* taláata SaGwáma...

(5.65) a.

9ágala		Gánama		málika
bicycle		sheep		queen

b.

fa	+	gúHa	→	fagúHa	*fáguHa
so		Guha		so Guha	

I suggest therefore that unaccented function words in EA should be analysed as free (pro-)clitics to an adjacent (accented) lexical word with PWd status. This implies that a non-exhaustive prosodic structure is tolerated, and thus the following ranking (as argued by Selkirk for English and by Zec for Serbian):

(5.66) PWDSIZE >> EXHAUSTIVITY
monosyllables are free clitics within a higher constituent with the lexical PWd

As we have seen, Selkirk's analysis of such free clitics places them within a higher phrase-level constituent. Proclisis in English is argued to arise due to a constraint which prefers structures with a PWd edge right-aligned to the phonological phrase edge:

(5.67) ALIGN(MAP, R; PWD,R): For any MaP in the representation, align its right edge with the right edge of some PWd.

What evidence is there in EA that unaccented function words procliticise to the following lexical word (as opposed to enclisis to a previous lexical word)? The consensus in the literature is that Arabic function words form a prosodic unit with following rather than preceding material (see inter alia Al-Ani 1992, Rifaat 2004, Watson 2002), and this is a key argument in favour of a proclitic analysis of unaccented EA function words.

The surface phonetic realisation observed on function words in the corpus survey yields little additional evidence. The pitch accent itself takes the form of a rising pitch movement which (as shown in chapter 4) is closely tied to the stressed syllable. In most cases, after the rising pitch accent pitch simply falls gradually towards the start of the next stressed syllable, across whatever unstressed syllables intervene, regardless of which word they belong to. The pitch contour thus reveals very little information about the exact position of the edges of PWds, nor, as a result, the prosodic affiliation of unstressed syllables. There is therefore no way to reliably judge whether intervening function words are prosodically joined with the preceding or following word[123].

Given the assumption then that the correct direction of cliticisation of function words in EA is rightwards, the next question to resolve is what phrase-level constituent the function word is incorporated into along with the following lexical word. I suggest that the null hypothesis is that the function word cliticises to the lexical word (which has PWd status) within a constituent of the 'next level up' in the Prosodic Hierarchy. This is based on the assumption that HEADEDNESS is undominated. If so then the correct alignment constraint to capture EA pro-cliticisation is as follows:

(5.68) ALIGN(MIP, R; PWD,R): For any MiP in the representation, align its right edge with the right edge of some PWd.

5.4.5 Summary

This section has presented empirical and theoretical evidence to suggest that the correct generalisation to account for the distribution of pitch accents in EA is that they associate to every PWd. This is indicated by the fact that function words can be 'promoted' to PWd status and thus be accented, so long as the resulting PWd is of sufficient prosodic size (bimoraic). In contrast, unaccented function words are neither themselves stressed nor induce stress-shift in an adjacent lexical word. A free clitic structure is therefore proposed for unaccented function words in EA.

Clarifying the treatment of function words in this way enables us to refine the generalisation regarding pitch accent distribution in EA: the domain of pitch accent distribution is the PWd, and pitch accent distribution in EA is thus a phonological rather than a lexical phenomenon.

[123] Recall that an attempt was made to transcribe direction of cliticisation during auditory transcription of the thesis corpus described in chapter 3, though without great success for the reasons set out here.

5.5 Discussion

The question arises as to why a language might mark every PWd with pitch? As noted above, all other phonological processes in EA apply within domains larger than the PWd: only stress assignment (and pitch accent distribution) apply within the PWd.

A process as pervasive as syllabification applies across word edges within a phrase level domain, usually described as the Phonological Phrase (≈MaP). This is apparent from the application across word boundaries of syllable repair processes such as vowel syncope (restricted to high vowels) and closed syllable shortening, as illustrated in (5.69) below. In (5.69a) /wi fi gawáab-na/ 'and in book-our' is syllabified as [wifgawábna], with shortening of the long vowel in /gawaab/, which falls within a closed syllable after affixation of the pronominal suffix /-na/ 'our'. In the parallel example in (5.69b) closed syllable shortening also applies but in addition the high vowel in the fist syllable of /kitaab/ 'book' is vulnerable to syncope, resulting in syllabification of /wi fi kitáab-na/ as [wfiktábna]. Note that procliticisation of the function words does not induce stress shift.

(5.69) Across-phrase syllabification examples (Kenstowicz 1980:48)

a.	/fi gawáab-hum	Gálta	wi fi gawáab-na	maa-fii-š	Gálta/
	in letter-their	mistake	and in letter-our	NEG-there-is-NEG	mistake
	[figawáb-hum	Gálta	wifgawáb-na	mafiiš	Gálta]
b.	/fi kitáab-hum	Gálta	wi fi kitáab-na	maa-fii-š	Gálta/
	in book-their	mistake	and in book-our	NEG-there-is-NEG	mistake
	[fiktáb-hum	Gálta	wfiktáb-na	mafiiš	Gálta]

Other prosodic repair processes that apply within domains larger than the word include vowel-vowel sequences repairs such as glottal-stop-epenthesis and glide formation (Watson 2002:228ff.). In addition many assimilatory processes are argued to also apply across word-boundaries within the phonological phrase (≈MaP), including coronal sonorant assimilation, voicing assimilation and palatalisation (Watson 2002:235ff.)[124]. Indeed as also discussed, epenthesis has been shown to apply across MaP boundaries, within the utterance (as can be seen in Figure 5.10 above) (Hellmuth 2004).

The only perceptual information to indicate the distribution of PWds for listeners in EA could be argued to be from the correlates used to mark prominence at the PWd level,

[124] El Zarka (1997:145ff.) argues that rhythmic redistribution of secondary stresses is sensitive to phonological phrase boundaries.

most notably the pitch accent. This parallels the suggestions made by Jun (2005b) in her discussion of the rich pitch accent distribution she observes in Spanish and Greek, that pitch accents in such languages may serve a word segmentation function (cf. chapter 3 section 3.5):

> "where pitch accent occurs at a regular interval (i.e. on almost every content word) with a similar type of pitch accent, each of the accents would provide a cue for a word boundary, functioning similarly to the Word boundary tone in Serbo-Croatian or the Accentual Phrase boundary tone in Korean. ... [with] the perceptual equivalence of word segmentation, whether marked by the head tone or by the edge tone of the unit.." (Jun 2005b:447)

Note that Jun draws a parallel between languages which mark the edge of every PWd and those which mark the head of every PWd. These could fulfil the demarcative and culminative prominence functions familiar in word-stress typology (Hayes 1981, Hayes 1995), and could contribute to effective perceptual marking of the PWd constituent[125].

5.6 Summary

This chapter has explored in some detail the theoretical mechanisms which have been proposed to account for density of pitch accent distribution, both in general and in specific languages, and then presented empirical evidence from EA which combines to suggest that the relevant domain of pitch accent distribution in EA is the Prosodic Word (PWd).

The key empirical evidence is from prosodic phrasing in complex EA sentences, which suggests that MaP boundaries are sparse in EA, and thus that the MaP cannot be the domain of pitch accent distribution. The role of the Minor Phrase (MiP) in EA was discussed and argued to be minimally branching and thus composed of two PWds, both of which are accented, so that the MiP cannot be the domain of pitch accent distribution either.

[125] Phonetic cues to prosodic constituency have been argued to play an important role in first language acquisition, so that children are able to infer syntactic constituency from prosodic cues by a process of 'prosodic bootstrapping' (see papers in Morgan & Demuth 1996). Under this hypothesis it is plausible to expect there to be some kind of phonetic correlate of every level of prosodic constituency which maps from a morphosyntactic category, and thus that there is some correlate of PWd constituency in every language.

Evidence from accentuation of content and function words in the corpus reveals that the correct generalisation to describe EA rich pitch accent distribution is that every PWd is accented and thus that the domain of pitch accent distribution in EA is the PWd.

The next chapter seeks a formal analysis to encode this generalisation, which will additionally capture the fact that EA has a small pitch accent inventory, as observed also in other languages with rich pitch accent distribution.

6 EA pitch accent distribution as a tone-prominence relation

6.0 Outline and aims

The previous chapter presented empirical evidence to suggest that the domain of pitch accent distribution in Egyptian Arabic (EA) is the Prosodic Word (PWd). This chapter now explores a formal analysis to encode density of pitch accent distribution as a parameter of prosodic variation across languages. Specifically, a formalism is adopted in which the phenomenon of a pitch accent on every PWd arises as a result of the relative ranking in EA of constraints governing the relationship between phonological tone and prosodic prominence.

This chapter starts in section 6.1 by outlining suggestions that have been made in the literature regarding the types of mechanism which might regulate the relationship between phonological tone and prosodic prominence. A particular conception of tone-prominence relations is then set out, in which surface relations between tone and prosodic prominence result from the interaction of a pair of inherently-ranked fixed hierarchies of markedness constraints which regulate association of tone to prosodic prominence, and of prosodic prominence to tone, respectively[126]. The relationship is conceived of as two-way, and hence is dubbed here a theory of tone↔prominence relations. In essence this is simply an extension of existing structure-based notions of pitch accent distribution (which were described in chapter 5). In EA the phonology requires every constituent at some level of the prosodic hierarchy to be associated with phonological tone, just as it does in English; however, the relevant level varies cross-linguistically: in English it is (arguably) the MiP, but in EA it is the PWd.

Section 6.2 offers a formal analysis of key data from the EA thesis corpus, and demonstrates how a tone↔prominence account can explain aspects of the surface EA facts, including for example variation in accentuation of function words. Section 6.3 explores implications and possible cross-linguistic applications of a theory of tone↔prominence relations. The chapter concludes by outlining two specific typological implications of the theory as applied to EA, as a preface to further experimental investigations presented in the remainder of the thesis (chapters 7 & 8).

[126] This conception of tone-prominence relations is inspired by ideas in Selkirk (2004b, 2005b)

6.1 Tone-prominence relations

As discussed in the previous chapter, it appears that the domain of pitch accent distribution in intonational languages may vary. Whilst in some languages the relevant domain is MaP or MiP, the claim of this thesis is that EA is a language in which the relevant domain is the PWd.

The idea that individual languages display a privileged relationship between pitch (or phonological tone) and some level of the prosodic hierarchy, and that the particular level may vary cross-linguistically, is not new. In this section we review formal mechanisms which various authors have suggested may underlie cross-linguistic variation in association between tone and constituents at different levels of the prosodic hierarchy.

Two notions are fundamental to these analyses. Firstly, the idea that the attraction between tone and a prosodic constituent is a *two-way relationship*, with association both of tone to the prosodic constituent and of the prosodic constituent to tone. This springs directly from the insights of Autosegmental Phonology, proposed by Goldsmith (1976) as an explanation of the possibility of both many-to-one and one-to-many relationships between tone and prosodic structure. The second conceptually important notion is that, if tone associates with 'some level of the prosodic hierarchy', then cross-linguistic variation can be *derived* from the intrinsically hierarchical nature of prosodic representation[127]. These two concepts are examined in turn below.

6.1.1 Tone-prominence: a two-way relationship

In his seminal work on Autosegmental Phonology Goldsmith (1976) argued that the mobility and multiple affiliation of tones is an indication of their autosegmental status. He proposed a representation in which tones operate on a separate tier in the phonological representation, and associate with elements in the prosodic structure according to well-formedness conditions:

[127] In chapter 5 there is evidence both of association of tone to heads of constituents and alignment of tone to the edges of constituents. Since the primary goal of this investigation is to account for the distribution of EA pitch accents, which associate to a head rather than align to an edge (as shown in chapter 4), the discussion here concentrates on association of tone to prosodic heads and leaves alignment of tones to edges to future research.

(6.1) Goldsmith (1976) Well-Formedness Conditions

1. Every TBU must have a tone.
2. Every tone must be associated to some TBU.
3. Association proceeds one-to-one, left-to-right.
4. Association lines must not cross.

The first two conditions in particular capture the fact that there may be both many-to-one and one-to-many relations between tones and prosodic structure, resulting in multiple tones on one syllable (contour tones) or tone-spreading across multiple syllables.

The two-way relationship between tones and syllables has been formalised within Optimality Theory in different ways by different authors. Myers (1997) suggests that the association relation between tones and TBUs expressed in Goldsmith's (1976) Well-Formedness Conditions can be formulated as a correspondence relation (that is, a specific type of relation) between the tonal and prosodic representations. This suggestion picks up on a comment in McCarthy & Prince (1995:266) that "the phenomena comprehended by the theory of autosegmental association are.. a special case of correspondence"[128]. Myers (1997 section 2) proposes the following constraints:

(6.2)	SPECIFY(T)	A syllable must be associated with a tone.
	*FLOAT	A tone must be associated with a syllable.

These constraints fall within the MAX and DEP families of constraints, respectively: SPECIFY(T) requires an element in the prosodic (tone-bearing) representation to have a correspondent in the tonal representation (the 'tonal tier' in Myers' terms); conversely, *FLOAT requires an element in the tonal representation to have a correspondent in the prosodic (tone-bearing) representation.

Anttila & Bodomo (2000) adopt this notion of correspondence, or faithfulness, between representations, and also Myers' constraints, and also appeal to two further constraints:

(6.3)	*CONTOUR	A syllable must be associated with at most one tone.
	*SPREAD	A tone must be associated with at most one syllable.

[128] Correspondence Theory (McCarthy & Prince 1995) provides for *faithfulness* constraints which require an element in one subrepresentation (such as the tonal tier) to have a correspondent in another subrepresentation (such as the prosodic representation). There are three basic types of faithfulness constraints regulating correspondence relations: MAX (do not delete a correspondent), DEP (do not insert a correspondent) and IDENT (correspondents must be identical in both representations).

Anttila & Bodomo note that these four of constraints together, if satisfied, capture the fact that the optimal correspondence between a tone and a tone-bearing unit (TBU) is one-to-one, and that they could be expressed as single (positively formulated) constraints regulating the relationship between tonal and prosodic structure:

(6.4) T:TBU Every tone is associated to exactly one TBU.
(does the work of *FLOAT and *SPREAD)

TBU:T Every TBU is associated to exactly one tone.
(does the work of SPECIFY(T)[129] and *CONTOUR)

In contrast to 'tone-TBU' as a correspondence relation, Yip (2002:83ff.) proposes an apparently identical set of constraints, which are however conceived of as *markedness* constraints on the well-formedness of tone~prosodic structure relations in output representation (rather than as faithfulness constraints between subrepresentations):

(6.5)

	*FLOAT	A tone must be associated with a TBU.
	SPECIFYT	A TBU must be associated with a tone.
	NOCONTOUR	A TBU may be associated with at most one tone.
	NOLONGT	A tone may be associated with at most one TBU.

By positing tone- and/or structure-specific markedness constraints one could argue that it is no longer necessary to refer to correspondence between subrepresentations in the formulation of constraints. In view of this, this chapter pursues an analysis of tone-TBU relations by means of constraints which are conceived of as markedness constraints on the properties of tone-TBU relations in output representation.

A variety of tonal phenomena have been analysed using these constraints, in their different forms. Myers (1997 section 2) appeals to SPECIFYT to account for spreading of lexical tone to underlyingly toneless syllables in Shona, and to *FLOAT to account for complete deletion of a tone in sequences undergoing Meussen's Rule (which deletes the second in a sequence of two H tones; de-linking of the second H would result in a floating tone). Anttila & Bodomo (2000:128ff.) use the constraint *TONELESS (that is, SPECIFYT) to account for H tone insertion and spreading to underlyingly toneless stems in Dagaaré (NW Ghana). Similarly, Yip (2002:162ff.) uses SPECIFYT to analyse L tone-spreading to underlyingly toneless syllables in Igbo.

[129] Antilla & Bodomo call this constraint *TONELESS rather than SPECIFY(T), but with the same definition.

Interestingly, especially for our present purposes, Gussenhoven (2000) appeals to the tone-TBU ideas of Anttila & Bodomo (2000) in his analysis of the lexical pitch accent language Roermond Dutch which is a mixed language with both lexical (pitch accent) and postlexical (intonational) tone. Gussenhoven makes a distinction between phrase tones which *align* to the edges of constituents (without being associated to a TBU), and lexical pitch accents which *associate* to TBUs[130]. Certain L- phrase tones in Roermond display both alignment to a phrase edge and secondary association to a (non-phrase-final) stressed syllable, resulting in a stretch of low level pitch between the last stressed syllable and end of the phrase, as in (6.6) below (Gussenhoven 2000 example 49)[131].

(6.6)

[Miene VOOT[1] zit aan miene bein[2]] *Roermond Dutch*
%L H* L%H
my foot sit on my leg
'My foot is attached to my leg'

Gussenhoven suggests that a relation of *association* may be created in order to satisfy a tone-TBU constraint which he formulates as follows (Gussenhoven 2000 example 45):

(6.7) Tone Bearing Unit (TBU) [+son] (μ)'σ "tones must be associated with a sonorant mora in a stressed syllable"

He notes that this formulation collapses into one constraint the two-way relationship which could be expressed by means of two constraints:

(6.8) TBU→T TBUs are associated with tone.
T→TBU Tones are associated with TBUs.[132]

Gussenhoven suggests that it is the constraint TBU→T which in Roermond penalises any stressed mora which bears no tone, and as a result attracts secondary association (and thus surface leftward spreading) of the L- phrase tone.

[130] The TBU in Roermond is the mora (Gussenhoven 2000, Gussenhoven 2004, Peters 2005).
[131] Gussenhoven argues that the boundary tone is realised before the lexical Accent 2 on [bein] 'leg'.
[132] Gussenhoven (2004:149) notates these constraints "TBU←T" and "T→TBU" respectively; for the sake of clarity I have instead used a notation in which only the linear order of the elements "T" and "TBU" indicates the direction of the relation captured.

The apparent conflict over whether the relationship between tone and prosodic structure is best captured by faithfulness or markedness constraints, is to some extent conceptually similar to arguments explored by DeLacy (1999:17ff., Appendix A1) regarding the use of positive or negative markedness constraints. DeLacy argues that the attraction between prosodically prominent positions (stressed syllables) and high tone (conceived of as tone-driven stress, and observed in languages as diverse as Ayutla Mixtec, Standard Serbian, Tibetan and Vedic Sanskrit) is better captured by means of negatively formulated markedness constraints (such as *HD/L "No low tone on stressed syllables.") than by positive formulated markedness constraints (such as HD:H "Stressed syllables have high tone.").

DeLacy argues that although either negative or positive constraints can account for Ayutla Mixtec, only an analysis using negative markedness constraints can express the fact that some languages 'conflate' categories which are distinguished by others. For example, one language may treat a long vowel and diphthong differently for the purposes of tone assignment, whilst another language may group these categories together and treat them identically. In addition, he suggests that an analysis using 'negative' markedness constraints has the advantage of avoiding overgeneration, since it will always prefer structures in which 'less is better'[133].

Yip (2000, 2002:98-99) has however pointed out that there are tonal phenomena which do not seem to yield to a negatively formulated markedness analysis. In Zhuang, for example the initial syllable in a bisyllable will bear a H tone regardless of the quality of the following tone, and Yip suggests this is because the initial syllable is the head and that there is a constaint requiring associated of a H tone to a Head: Head = H. It is not enough to appeal to DeLacy's *Hd/L ("No low tone on heads.") because L tones do survive into the output, on other syllables. In Mandarin Chinese a contrastive stress cannot be realised on a L-toned syllable, whereas it can be realised on any of the following[134]: H, MH or HL. There is no way to capture this particular 'non-contiguous' scale of tones other than by appealing to a positively formulated markedness constraint, which requires a contrastively stressed syllable to bear H tone[135].

[133] It is also consistent with McCarthy's hypothesis that all markedness constraints are negatively formulated (McCarthy 2003:78).
[134] 'M' indicates a 'Mid' tone, so 'MH' is a high rising tone.
[135] Yip also points out cases involving tone which require positional faithfulness constraints and sequential markedness constraints (see Yip 2000). Selkirk has also offered an account of Ayutla Mixtec

In view of this, the analysis below will use positive markedness constraints to capture tone↔prominence relations.

Finally, as well as T-TBU constraints Anttila & Bodomo (2000) also employed constraints encoding a more standard kind of correspondence relation between input and output tones:

(6.9) $T{:}T_i$ Every output tone is linked to exactly one input tone.
$T_i{:}T$ Every input tone is linked to exactly one output tone.

These equate fairly straightforwardly to widely accepted tone-specific faithfulness constraints: DEP$_{TONE}$ (Every tone in the output has a correspondent tone in the input), MAX$_{TONE}$ (Every tone in the input has a correspondent tone in the output) and IDENT$_{TONE}$ (Correspondent segments in input and output are identical in tone features.).

We turn in the next section to theories and analyses of tone~prosodic structure relations which have made appeal to the notion of variation in the target of tonal association.

6.1.2 Tone-prominence: variation in the target of tonal association

As is well-known, among tone languages the prosodic constituent which functions as tone-bearing-unit (TBU) may vary cross-linguistically, being usually either the mora or the syllable (see summary in Yip 2002:73-76).

As an example, in Dagaaré underlying HL or LH tones associate left-to-right to syllables; since Dagaaré permits at most one tone per syllable, the underlying contour spreads over two syllables (Yip 2002:141-2):

(6.10)	Dagaaré:	LH	bààlá	'sick-person (sg.)'
		HL	núórì	'mouth (sg.)'

If the TBU in Dagaaré were the mora then the tones would associate to left-to-right to moras, resulting in a contour tone on a long vowel or diphthong (*[bàálá]; ([á] denotes a high tone; [à] denotes a low tone).

using positive markedness constraints (Selkirk 2005b), made possible by appeal to a different analysis of the stress system, using END-RULE-L, instead of ALL-FT-L (cf. McCarthy 2003).

In contrast, in Kunama (Eritrea) underlying tones associate left-to-right to moras, resulting in word-medial contour tones on heavy syllables ([ā] denotes a mid tone) (Yip 2002:141-2):

(6.11) Kunama: MHM mōódā 'quarrel'

It has also been suggested that the TBU may vary in lexical pitch accent languages. For example Peters (2005) has argued that the TBU varies across different Central Franconian dialects of Dutch, of which Roermond Dutch is an example. All of these dialects have mixed pitch accent/intonational systems and feature an Accent 1/Accent 2 lexical contrast. In the Venlo dialect the TBU is the mora, whereas in the Tongeren dialect the TBU is the syllable. This distinction is in part observed in the phonetic realisation of bitonal lexical pitch accents. For example, in the Venlo dialect, an Accent 2 H*H pitch accent is realised as a short plateau within a bimoraic syllable, whereas an Accent 1 plain H* leaves the second mora of the stressed syllable available as a target for secondary association of a following L% phrase tone, resulting in a sharp fall in pitch within the syllable. Additional evidence comes from neutralisation of the tonal contrast between Accent 1 and Accent 2 in certain contexts. In Tongeren the contrast between Accent 1 and Accent 2 can be fully realised in monomoraic syllables resulting in near minimal pairs such as: /kas^1/ 'cupboard' vs. /kas^2/ 'candle'; in the same context in Venlo the contrast is neutralised (due to the lack of targets for association of tones).

The target of tonal association in languages with *only* postlexical tones, that is, in purely intonational languages, is a subject of much debate (explored in detail in chapter 7 section 7.1). A common assumption in the literature is that the surface alignment of pitch targets is a reflex of their phonological association[136], and further that the target of that association is almost invariably the stressed syllable of an accented word. Similarly, in an early formulation of the target of association of intonational pitch accents in English, Pierrehumbert & Beckman (1988:159) describe 'central' association of tone to target in English as follows:

> "English permits at most one pitch accent per metrical foot, and the accent is located on the stressed syllable. This might be described by saying that accent is a foot-level property that is attracted to the head syllable."

[136] See Xu &Liu (2005) however for a different view.

The restriction 'at most one pitch accent per metrical foot' sounds very like the kind of arguments that are made to support proposal of the mora or the syllable as the TBU in tone languages: in the example in (6.10) above from Dagaaré, the language allows at most one tone per syllable, hence an underlying HL contour does not surface on a bimoraic syllable. Gussenhoven uses a similar argument that the target of pitch accent association in English might be the foot, based on the fact that stress shift, which he analyses as *accent* shift, is blocked in words containing only one metrical foot (these examples are from Gussenhoven 2004:142, for a summary see Hayes 1995):

(6.12)	a.	Chinése	b.	obése	
		a Chínese book		an obése person	* an óbese person

Emerging quantitative evidence also suggests that English pitch accents may display phonological association to a domain larger than the syllable. Ladd has reported that alignment of the peak in English phrase-final rising-falling nuclear accents is best described by a measure of peak delay as a proportion of the duration of the PWd; in contrast peak delay as a proportion of the duration of the accented syllable did not yield a consistent result (Ladd 2005). The phrase-final words tested were monosyllables and disyllables (such as 'mine' and 'miner'), and in both of these cases the PWd is co-extensive with the (single) metrical foot in the word. These quantitative results might therefore equally be an indication of peak alignment relative to the *foot* as TBU - the domain to which tones display association - in English.

It seems then that, as well as differences among intonational languages in the domain of pitch accent distribution across levels of the Prosodic Hierarchy (as demonstrated in chapter 5), there may also be variation across the hierarchy in the target of pitch accent association (the TBU), and also whether the local target at that level is the head or edge of the constituent.

Gussenhoven (2004:148ff.) has expressed this possibility in the form of local expansions of "T↔TBU" constraints. For example he suggests that each of the T↔TBU constraints that he proposes (in example 6.8 above) represents an inherently ranked family of constraints, ranging from the general to the specific. The most general TBU in a language might be, say, the mora (according to the language in question) in which case the most specific TBU would be the accented mora, hence: $T \rightarrow {}'\mu \gg T \rightarrow \mu$. Thus, "if a language associates tones with moras it will also associate them with accented moras" (Gussenhoven 2004:149). He proposes a general markedness constraint banning association of tones to prosodic structure located at some point in this fixed hierarchy, which enables us to capture languages in which tones associate only with stressed syllables (rather than with all syllables) (Gussenhoven 2004:149):

(6.18) NOASSOC TBUS are not associated with tones.

T→ 'σ >> NOASSOC >> T→σ

For Gussenhoven, the sister family of constraints, T→TBU, also splits into a fixed hierarchy of constraints ranging across types of tones (such as H, L, H* and L*). The ranking of *CROWD (which mitigates against multiple association of tones to TBUs) relative to the T→TBU constraints determines which tones are associated to TBUs and which are left unassociated (and are thus realised as a leading or trailing tone, for example). Gussenhoven suggests that the notion of the starred tone ('*') notation in AM theory is the reflex of the constraint ranking: H*→TBU >> L→TBU[137].

Goldsmith (1987) noted a generalisation that can be made about *all* tone-TBU relations, which he formalised as the Tone-Accent Attraction Condition (TAAC):

(6.19) The Tone-Accent Attraction Condition (TAAC)

> "A tone-to-grid structure is well-formed iff there is no tone-bearing syllable which has a lower level of accent than a toneless syllable. (Thus if a syllable S has tone, all syllables with greater level of accent than S must also bear tone.)" (Goldsmith 1987)

In a metrical analysis then, under the assumption that accent equates to relative prominence at a particular level, having a 'lower level of accent' equates very simply to 'non-head' (i.e. 'non-DTE') status at that level. The metrical representation is arranged in such a way that relative prominence at any particular level implies relative prominence at all lower levels. In turn, lack of relative prominence at a particular level excludes the possibility of relative prominence at all higher levels (cf. Hayes 1995, the Continuous Column Constraint, and also chapter 2 section 2.1.5 on the properties of the prosodic hierarchy).

[137] Gussenhoven appeals to L→TBU to account for the lack of trailing L in French nuclear accents (Gussenhoven 2004: ch13).

Goldsmith's TAAC implies attraction of tone to prosodic heads/relative metrical prominence as a general property of well-formed tone-prosodic structure relations. In a more abstract sense then Goldsmith's condition is about 'tone-prominence attraction'. If we adopt this view, that attraction of tone to a prosodic constituent is an indication that the constituent is metrically prominent (it is the DTE of a constituent at the next level up in the hierarchy), then analysis of attraction of tone to prosodic constituents can be analysed by exploiting the pre-existing asymmetry between heads and non-heads within a prosodic constituent.

There are two ways of expressing hierarchical relations in Optimality Theory(McCarthy 2002): by harmonic alignment of natural prominence scales, or by encoding *stringency relations* among linguistic forms into constraints[138]. The inherently hierarchical nature of metrical/prosodic representation, as discussed in chapter 2 section 2.1.5, yields linguistic forms (that is, constituents at different levels of the hierarchy) between which stringency relations inherently exist[139].

Goldsmith's condition certainly holds of instances in which tone and prominence (or accent) interact, but it is more controversial to hypothesise that the condition reflects a property of well-formed tone-prosodic structure relations in general. Whilst many tone languages display tone-prominence interaction, there is also plenty of evidence from tone languages in which tonal distribution is entirely independent of metrical prominence (which is instead expressed by other means[140]). Indeed as we have seen, there are also intonational languages in which, as Beckman (2004) and Jun (2005b) point out, phonological tones are independent of metrical prominence (and are anchored instead at the edges of prosodic constituents)[141]. Since EA appears to be a clear-cut case of attraction of tone to metrical prominence the focus of our attention will be on testing a theory of tone↔prominence relations against the facts of EA. Potential application of the theory to languages in which tones are attracted to the edges of constituents is reserved for the discussion section at the end of the chapter.

---

[138] If two constraints A and B stand in a stringency relation then the violations of A will always be a proper subset of the violations of B: A "imposes a more stringent test" than B does (McCarthy 2002:20).
[139] Assuming LAYEREDNESS and HEADEDNESS to be undominated as per Selkirk (1996).
[140] These include such as prosodic lengthening, reset of downstep, restricted segmental distribution in metrically non-prominent positions (Downing 2004).
[141] Again, in such languages prominence is expressed by other means including reset of downstep and prosodic lengthening (Jun 1996, Beckman & Pierrehumbert 1986, Pierrehumbert & Beckman 1988).

In the next section (6.1.3) I set out a theory of tone↔prominence relations which exploits the inherently hierarchical stringency relations that hold between levels of the prosodic hierarchy.

6.1.3 A theory of tone↔prominence relations

This section sets out the properties of the main constraints argued to be responsible for patterns of tone-prominence relations and the constraints are used to analyse the facts of EA in section 6.1.4 below.

In line with the basic notion that the relation between tone and prosodic structure is two-way (discussed in 6.1.1), following Selkirk (2004b), I propose a set of T↔P constraints which are positively formulated markedness constraints on output representations[142], of the following form:

(6.20)

a.	T→μ(P)	In the output representation, every tone (T) is associated with (the mora that is) the head of a prosodic constituent of level P.
b.	μ(P)→T	In the output representation, every (mora that is) head of a prosodic constituent of level P is associated with some tone (T).

The constraints vary across constituent levels of the prosodic hierarchy in two fixed hierarchies; note the reverse direction of ranking in the two families of constraints:

(6.21)
a. Tone-to-Prominence constraints [T→P]
T→μ >> T→μ(σ) >> T→μ(Ft) >> T→μ(PWd) >> T→μ(MiP) >> T→μ(MaP) >> T→μ(IP)

b. Prominence-to-Tone Constraints [P→ T]
μ(IP)→T >> μ(MaP)→T >> μ(MiP)→T >> μ(PWd)→T >> μ(Ft)→T >> μ(σ)→T >> μ→T

For ease of presentation, from now on these constraints will be referred to simply by the relevant prosodic domain whose head attracts/requires tone. Thus T→Ft stands for T→μ(Ft), and PWd→T stands for μ(PWd)→T, and so forth.

[142] Selkirk (2005b) suggests that the presence of tone-specific constraints in the grammar eliminates the need for a tone-specific subrepresentation separate from the melodic representation for segments (i.e. the tonal and segmental tiers can be ‘conflated’). This implication is beyond the scope of the present study and will not be pursued further.

Crucially, the T→P constraints permit the notion of 'TBU' itself to be encoded as 'the prosodic head of some level of prosodic constituency' (cf. Yip 2002:141). Thus, a constraint "T→TBU", in a language in which the TBU is the syllable, can be interpreted as follows: "T→μ(σ): A tone is required to be associated to (the mora that is) the head of a syllable".

Selkirk suggests that there are three types of tone (Selkirk 2005b):

i) phonemic tones: part of the underlying representation of words;
ii) morphemic tones: floating tones, present in the morphosyntactic input;
iii) epenthetic tones: inserted in the output representation.

An 'epenthetic tone' is defined as one inserted in order to satisfy phonological output constraints, rather than in an effort to remain faithful to underlyingly present tones (whether present in the lexical entry, phonemic tone, or in morphosyntactic representation, morphemic tone).

These three types of tone are indistinguishable in the output representation and will thus be treated identically in phonetic interpretation. In addition they are all equally subject to the influence of T↔P markedness constraints. In contrast, only those tones which appear in input representation will be affected by the following faithfulness constraints.

(6.22) MAX$_{\text{IO TONE}}$ Every tone in the input has a correspondent in the output.
DEP$_{\text{IO TONE}}$. Every tone in the output has a correspondent in the input.[143]

The T→P constraints interact with the constraint against tonal deletion ("MAX$_{\text{TONE}}$"). The language-specific ranking of MAX$_{\text{TONE}}$ relative to the fixed hierarchy of T→P constraints determines which of them is most obviously active in a particular language. In contrast, P→T constraints interact with the constraint against tonal insertion ("DEP$_{\text{TONE}}$"). The language-specific ranking of DEP$_{\text{TONE}}$ relative to the fixed hierarchy of P→T constraints determines which of them is active in a particular language. An alternative constraint whose interaction with the constraint hierarchies may be relevant would be the markedness constraint *T (cf. Gussenhoven 2004:257). The distinction between the effects of the above faithfulness constraints (MAX$_{\text{TONE}}$ & DEP$_{\text{TONE}}$) and the

[143] Henceforth these two input-output faithfulness constraints will be called simply MAX$_{\text{TONE}}$ and DEP$_{\text{TONE}}$.

markedness constraint *T is discussed in the context of their relevance for analysis of EA in section 6.1.4 below.

What determines the direction of the inherent ranking of these fixed rank hierarchies? Selkirk suggests that in the case of T→P the highest ranked constraint is the one requiring tone to be associated with some segment that is head of a mora. For P→T constraints, the highest ranked constraint is the one involving the highest level of the prosodic hierarchy, on the basis that it is more economical to mark prominence at higher levels of the hierarchy than at lower levels. In the unmarked case then tones will be associated to moras, and (minimally) the head mora of every IP will bear tone. In the case of lexical tones, the unmarked case will look like a tone language (with every lexical tone associated to a mora); in the absence of lexical tones, the unmarked case will be an intonational stress accent language (with tone inserted to mark the head of every IP). All other possible language types (such as pitch accent languages and mixed accentual intonational systems) will lie somewhere along the continuum between the two extremes of these 'pure' language types, from a pure tone language such as Vietnamese to a pure intonation language such as English.

The mirror-image ordering of the two constraint families can also be supported by the plausible assumption that archetypal 'T→P languages' have phonemic tone, that is, it is underlying tone that surfaces. Such languages are tonal (tone and pitch accent languages), and in these the unmarked TBU is indeed at the lower end of the prosodic hierarchy (mora or syllable). Similarly, archetypal 'P→T languages' could be argued to have epenthetic tone, in which tone largely functions to highlight prosodic prominence. These are intonational languages, and in these the domains whose prominence are tonally marked are indeed towards the upper end of the prosodic hierarchy (MaP or IP). We turn now to the types of effects caused by T↔P constraints. Focussing on P→T constraints, whose effects are most relevant to EA, these call for the head of every prosodic constituent at some level of the hierarchy to bear tone. As mentioned in section 6.1.2 above, Yip (2002:162ff.) uses SPECIFYT to analyse L tone-spreading to underlyingly toneless syllables in Igbo. In a T→P account one might say that the active constraint in Igbo is σ→T ('the head of every syllable must bear tone'), with the result that default, epenthetic tone is inserted onto underlyingly toneless syllables. Similarly, in chapter 5 section 5.2.2, it was suggested that English is a language in which tones are inserted associated to the heads of MiP level constituents, due to the effects of

MIPACCENT ("Every minor phonological phrase (MiP) must contain at least one accent"). In a T→P account one might say that the active constraint in English is MiP→T ('The head of every MiP must bear tone'), then the distribution of pitch accents in English as described in Selkirk (2000) can be captured[144]. The most relevant constraint for our present purposes is PWd→T, which Selkirk (2005b) suggests could be responsible for rich pitch accent distribution observed in Italian (Grice et al 2005):

(6.23) PWd→T A mora that is head of a PWd is required to be associated to tone.

It is this P→T constraint which appears to hold the key to understanding the distribution of pitch accents in EA. The main purpose of the remainder of this chapter is to determine whether a tone↔prominence theory of this sort can really handle the facts of a language like EA in which every PWd bears a pitch accent, that is, in which the head of every PWd is marked with tone. It is beyond the scope of this study to test the implications of the theory empirically on languages other than EA, so the chapter focuses on how T↔P theory might account for EA.

A key advantage of T↔P theory, if it can capture the EA data, is that it formalises the notion that pitch accents in EA are purely epenthetic, and thus could be said to predict that the language will have a small pitch accent inventory. In segmental phonology we are used to the notion that an epenthetic segment, inserted into the phonological representation to fill some gap, is usually a 'default' segment, such as a centralised vowel [ə] or an unmarked stop ([t]). If all EA pitch accents are 'default' pitch accents it is perhaps to be expected that they are all of one type. Note that for the time being I assume (non-trivially) that the T↔P constraints are blind to the quality of tone inserted, and thus that the inserted 'default' tone could be a complex tone (such as L+H*) rather than a simplex tone (such as H*)[145].

The next section previews the remainder of the chapter, setting out in more detail the hypotheses explored for EA within the tone↔prominence theory conception of the relationship between phonological tone and prosodic structure.

---

144 Recall however that Selkirk (2000) relied on an accent-first conception of pitch accent insertion.
145 This assumption is revisited briefly in section 6.3.3 below.

6.1.4 Testing the tone~prominence theory against EA

Chapter 5 set out in detail empirical evidence in support of the claim that in EA the domain of pitch accent distribution is the Prosodic Word (PWd). In particular, the claim is that phonological tone (pitch) functions in EA to mark prosodic prominence at the level of the PWd. Nonetheless EA is uncontroversially a purely intonational language, in which tone plays no part in the lexical realisation of any morphemes (cf. Hyman 2001:1367). This contrasts strongly with other intonational languages such as English in which tone similarly plays no lexical role, but is arguably used to mark prosodic prominence at a different (higher) level of the prosodic hierarchy (such as MiP, Selkirk 2000).

The hypothesis explored here for EA is twofold. Firstly, that the constraint driving rich pitch accent distribution in EA is PWd:T, which outranks constraints mitigating against insertion (DEP_{TONE}). The analysis is worked out in section 6.2 below, testing the hypothesis that the following ranking results in rich pitch accent distribution in EA:

(6.24)

IP→T >> MaP→T >> MiP→T >> PWd→T >> $\boxed{\text{DEP}_{\text{TONE}}}$ >> Ft→T >> σ→T >> μ→T

Note that DEP_{TONE} ('Don't insert tones') and *T ('Avoid tones altogether') have exactly the same effect in EA, in which there are no tones in input representation. For the present I therefore analyse EA with DEP_{TONE} only. Instances in which T* and DEP_{TONE} might have different effects in a language are discussed briefly in section 6.3.2.

Secondly, in a slight departure from Selkirk's conception of the T↔P constraints, I would like to explore what role T→P plays in EA. Specifically I would like to suggest that the T→P markedness constraints regulate all tones and not only tones of lexical origin. This claim is a direct result of the decision to formulate the T→P hierarchies as markedness constraints, which therefore only 'see' output representation. However I believe that this conception of tone↔prominence relations (as markedness constraints) is the only conception of the T↔P relationship which is consistent with the notion of 'the unity of pitch phonology' (see discussion in chapter 2 section 2.1, and Ladd 1996:147ff.).

Specifically, this hypothesis means that although the ranking of MAX_{TONE} relative to the T→P hierarchy is effectively irrelevant in EA (since there are no tones present in input representation[146]) the relative ranking of DEP_{TONE} relative to the T→P hierarchy is nonetheless relevant. What would be the effects of T→P in a language without lexical tone? I suggest that the ranking of DEP_{TONE} relative to the T→P constraints can derive the 'TBU' of intonational languages also. The notion of TBU is not widely used in the analysis of intonational languages. Nonetheless as we have seen (6.1.2 above) cross-dialectal variation in the TBU for mixed pitch accent/intonational languages has been reported (the Central Franconian dialects, Peters 2005), and indeed entertained for intonational languages in which tone is fully postlexical (Ladd 2005).

This second hypothesis is explored with respect to EA in chapter 7.

6.1.5 Summary

This section reviewed mechanisms that have been proposed in the literature to express the two-way relationship between tone and metrical structure/prominence, as well as methods of encoding cross-linguistic variation in the target of tonal association across constituents in the inherently hierarchical prosodic structure. The basic facts of a theory of tone↔prominence relations suggested by Selkirk were outlined. Finally the specific hypotheses that the theory permits us to predict for EA were explored: firstly, that PWd→T outranks DEP_{TONE} and results in insertion of epenthetic tone to each PWd in the prosodic representation; and secondly, that T→P constraints can be used to derive the TBU in an intonational language such as EA (explored in chapter 7).

The first of these hypotheses is tested in the remainder of this chapter. Specifically, the next section (6.2) offers a formal analysis of EA pitch accent distribution using P→T constraints, in interaction with interface constraints on the mapping between morphosyntactic and prosodic structure.

[146] There could in principle be morphemic tones in EA, which are inserted in the morphosyntax (e.g. related to focus or topic status) and which would therefore be present in the input to the phonological component of the grammar. In such cases the relative ranking of MAX_{TONE} would become relevant. I am not aware of any tonal phenomena in EA which merit positing such tones, but assume that MAX_{TONE} would be ranked in the same position as DEP_{TONE} relative to the T:P hierarchy.

6.2 Analysing EA pitch accent distribution: formal analysis

In this section I present an analysis of speech data from the thesis corpus (examples highlighted in chapter 5) to establish the relative ranking in EA of P→T constraints and relevant interface constraints mapping between morphosyntactic structure and prosodic structure. Section 6.2.1 explores the distribution of phonological tone in EA, reflected in the ranking among PWD→T, DEPTONE, and FT→T. Then section 6.2.2 explores interaction of these with constraints on the mapping between lexical morphosyntactic words and PWds (LEXWD:PWD), to capture the generalisation that it is usually better in EA *not* to map a lexical word to a PWd, than to leave a PWd unaccented. Section 6.2.3 treats variation in the accentuation of function words, exploring the role of FTBIN, PWD:LEXWD and NOLAPSE. Finally in section 6.2.4 the question of the direction of cliticisation of unaccented function words in EA is discussed.

6.2.1 The distribution of phonological tone in EA

This section sets out the part of the grammar which is responsible for the distribution of phonological tone in EA, analysing data in which every Prosodic Word is accented. To illustrate the relative ranking among PWD→T, DEP$_{TONE}$ and FT→T in EA let us observe how some sentences are treated in speakers' actual productions (these sentences are taken from the narratives corpus, as discussed in section 5.3.4 above).

In an example from the focus section of the corpus, the sentence /maama bitit9allim yunaani bil-layl/ ('Mum learns Greek in-the-evenings') was treated uniformly by all speakers in all productions, with a pitch accent on all four content words, as in (6.25).

(6.25)

maama	bitit9allim	yunaani	bi-	-l-	layl	
ǀ LH*	LH*	LH*	<	<	LH*	L-L% ǀ
mum	learns	Greek	in-	-the-	night	
[[NP]$_{NP}$	[V	[NP]$_{NP}$	[PP		[NP]$_{NP}$	]$_{PP}$]$_{VP}$]$_{S}$

The preference for accentuation of PWds over fewer inserted tones indicates that the constraint PWD→T outranks the constraint militating against tone insertion, DEP$_{TONE}$, as illustrated in the tableau in (6.26)[147]. In candidate (a.) every PWd is accented, by insertion of four pitch accents, each of which is penalised in the form of a single categorical violation of DEP$_{TONE}$; in candidate (b.) no accent is inserted on [maama], so

[147] In the tableaux lexical categories in the input form are marked with a subscript: "lex". Vowels bearing a pitch accent in output forms are indicated with an acute accent mark on the accented vowel: eg "á" or "ú". The edges of prosodic constituents are marked at **[**PWd**]**, **(**MiP**)** and **ǀ**MaP**ǀ** level.

(b.) incurs fewer violations of DEP$_{\text{TONE}}$ but at the cost of violating PWD→T. The winning candidate is (a.) indicating that PWD→T outranks DEP$_{\text{TONE}}$.

(6.26) PWD→T >> DEP$_{\text{TONE}}$

	/maama$_{lex}$ bitit9allim$_{lex}$ yunaani$_{lex}$ bi-l-layl$_{lex}$ /	PWD→T	DEP$_{\text{TONE}}$
☞	a. |([máama] [bitit9állim])([yunáani] bi-l-[láyl])|		****
	b. |([maama] [bitit9állim])([yunáani] bi-l-[láyl])|	*!	***

The next task is to determine the ranking between PWD→T and FT→T, which will confirm whether P→T constraints really are in a stringency relation as claimed in section 6.1.2 above. In EA the foot is the moraic trochee, composed of either a CVC or CVV heavy syllable, or a sequence of two light syllables: (CVCV); 'degenerate' feet of any kind, formed of less than two moras, are not tolerated in any position[148]. The foot structure of our example sentence is thus as follows (— denotes a heavy syllable; • denotes a light syllable; <x> denotes an extrametrical consonant; feet are underlined):

(6.27) — •   • — — •   • — •   — —
thus: maa.ma   bi.tit.9al.li<m>   yu.naa.ni   bil.lay<l>
Ft   Ft Ft   Ft   Ft Ft

An output candidate such as candidate (c.) in the tableau in (6.28) below, in which every *foot* bears a pitch accent, is disfavoured by DEP$_{\text{TONE}}$, which must thus outrank FT→T:

(6.28) Ft→T >> DEP$_{\text{TONE}}$

	/maama$_{lex}$ bitit9allim$_{lex}$ yunaani$_{lex}$ bi-l-layl$_{lex}$ /	DEP$_{\text{TONE}}$	FT→T
☞	a. |([máa.ma] [bi.tit.9ál.li<m>])([yu.náa.ni] bil-[láy<l>])|	****	**
	c. |([máa.ma] [bi.tít.9ál.li<m>])([yu.náa.ni] bíl-[láy<l>])|	******!	

By transitivity therefore we can say that the following ranking holds in EA:

(6.29) PWD→T >> DEP$_{\text{TONE}}$ >> Ft→T

	/maama$_{lex}$ bitit9allim$_{lex}$ yunaani$_{lex}$ bi-l-layl$_{lex}$ /	PWD→T	DEP$_{\text{TONE}}$	FT→T
☞	a. |([máama] [bitit 9állim])([yunáani] bil-[láyl])|		****	***
	b. |([maama] [bitit 9állim])([yunáani] bil-[láyl])|	*!	***	****
	c. |([máama] [bitít 9állim])([yunáani] bíl-[láyl])|		*******!	

[148] See chapter 2 section 2.3.1 for a summary of EA stress assignment.

In the tableau in (6.29) candidate (a.) fully satisfies PWD→T, and candidate (c.) fully satisfies FT→T[149]. The violations of DEP_{TONE} incurred by candidate (a.) are a subset of those incurred by candidate (c.). We can thus say that candidate (a.), the PWD→T obeying candidate, 'harmonically bounds' candidate (c.), the FT→T obeying candidate, and thus that the candidate satisfying FT→T can never win under any ranking.[150]

Indeed we can demonstrate schematically that a candidate satisfying a P→T markedness constraint at the top of the prosodic hierarchy will *always* harmonically bound candidates satisfying a 'lower' P→T constraint: given undominated HEADEDNESS and LAYEREDNESS, a high level P→T constraint will always incur a subset of the DEP_{TONE} violations incurred by lower level T→P constraints. This is illustrated in the table in (6.30) below, and confirms the claim made in section 6.1.2 that stringency relations hold inherently between constituents of the prosodic hierarchy. Having established this, FT→T is excluded from the analysis from now on.

(6.30) Schematic comparison of the violations of DEPTONE incurred by candidates satisfying P→T constraints

	Candidates	*Outcome*
a	satisfies IP→T (every IP has tone)	violates DEPTONE x 1 per IP
b	satisfies MiP→T (every MiP has tone)	violates DEPTONE x 1 per MaP
c	satisfies MiP→T (every MiP has tone)	violates DEPTONE x 1 per MiP
d	satisfies PWd→T (every PWd has tone)	violates DEPTONE x 1 per PWd
e	satisfies Ft→T (every Ft has tone)	violates DEPTONE x 1 per Ft
f	satisfies σ→T (every σ has tone)	violates DEPTONE x 1 per σ
g	satisfies μ→T (every μ has tone)	violates DEPTONE x 1 per μ

Having established the ranking of the key markedness and faithfulness constraints, that is between P→T & DEPTONE, the next section explores the interaction of these with interface constraints on the mapping between morphosyntactic structure and prosodic structure at the word level.

6.2.2 Pitch accent distribution and the mapping of lexical words to PWds in EA

This section treats the mapping of lexical words to PWds in EA. It would be possible to insert fewer accents without violating PWD→T, the better to satisfy DEP_{TONE}, if fewer PWds were formed. On the assumption that purely phonological constraints such as the

[149] Candidate (b.), reproduced from previous tableaux, is included to demonstrate that failing to accent a PWd incurs fewer violations of DEP_{TONE}.

[150] "The mapping /A/ →[B] harmonically bounds the mapping /A/ →[C] iff the /A/ →[B] mapping incurs a proper subset of the constraint violations incurred by the /A/ →[C] mapping." (McCarthy 2002:23)

T↔P constraints can only ‘see’ phonological categories, PWD→T will not itself penalise unaccented words which are not mapped to a PWd constituent. Such forms would however violate LEXWD:PWD, the interface constraint on the mapping between morphosyntactic structure and prosodic structure at the word level:

(6.31) LEXWD:PWD A lexical word maps to a PWd[151].

The preference for a candidate which satisfies LEXWD:PWD over a candidate which better satisfies DEP_{TONE} suggests that LEXWD:PWD outranks DEP_{TONE}:

(6.32) LEXWD:PWD >> DEP_{TONE}

	/maama$_{lex}$ bitit9allim$_{lex}$ yunaani$_{lex}$ bi-l-layl$_{lex}$ /	LEXWD:PWD	DEP_{TONE}
☞	a. |([máama] [bitit9állim])([yunáani] bi-l-[láyl])|		****
	d. |(maama [bitit9állim])([yunáani] bi-l-[láyl])|	*!	***

Thus the ranking established so far for EA is:

(6.33) PWD→T, LEXWD:PWD >> DEP_{TONE} (>> Ft→T)

	/maama$_{lex}$ bitit9allim$_{lex}$ yunaani$_{lex}$ bi-l-layl$_{lex}$ /	PWD→T	LEXWD: PWD	DEP_{TONE}
☞	a. |([máama] [bitit9állim])([yunáani] bi-l-[láyl])|			****
	b. |([maama] [bitit9állim])([yunáani] bi-l-[láyl])|	*!		***
	d. |(maama [bitit9állim])([yunáani] bi-l-[láyl])|		*!	***

This is the key section of the phonological grammar which I propose accounts for rich pitch accent distribution in EA. The grammar is also demonstrated in a sentence of greater complexity in (6.34) and (6.35) below.

[151] This constraint is equivalent to MCAT=PCAT in McCarthy & Prince (1993), and conflates a left/right edge sensitive pair of constraints, ALIGN(LEX,L; PWD,L) and ALIGN(LEX,R; PWD,R) (cf. 5.4.2 above).

(6.34) Rich pitch accent distribution in a complex sentence.

<table>
<tr><td>HaSalit</td><td>9ala</td><td>minHa</td><td>min</td><td>is</td><td>sifaara</td><td>9ala</td><td>šaan</td><td>e</td><td>tiruuH</td><td>tidris</td><td>fi</td><td>?amriika</td><td></td></tr>
<tr><td>she-received</td><td>of-</td><td>grant</td><td>from-</td><td>-the-</td><td>embassy</td><td>in-</td><td>order</td><td></td><td>she-goes</td><td>she-studies</td><td>in-</td><td>America</td><td></td></tr>
<tr><td>| LH*</td><td><</td><td>LH*</td><td>---</td><td>></td><td>LH*</td><td><</td><td>LH*</td><td></td><td>LH*</td><td>LH*</td><td><</td><td>LH*</td><td>L-L% |</td></tr>
<tr><td>[[V</td><td>[PP</td><td>[NP]$_{NP}$]$_{PP}$</td><td>[PP</td><td></td><td>[NP]$_{NP}$]$_{PP}$</td><td>[PP</td><td>[NP]$_{NP}$]$_{PP}$</td><td>[C</td><td>[v</td><td>[V</td><td>[PP</td><td>[NP]$_{NP}$</td><td>]$_{PP}$]$_{VP}$]$_{vP}$]$_{CP}$]$_{VP}$]$_{S}$</td></tr>
</table>

'She got a grant from the embassy to go and study in America.'

(6.35) PWD→T, LEXWD:PWD >> DEP$_{TONE}$ in a complex sentence.

<table>
<tr><td></td><td>/HaSalit$_{lex}$ 9ala-minHa$_{lex}$ min-is-sifaara$_{lex}$ 9ala-šaan$_{lex}$ tiruuH$_{lex}$ tidris$_{lex}$ fi-?amriika$_{lex}$/</td><td>PWD→T</td><td>LEXWD: PWD</td><td>DEP$_{TONE}$</td></tr>
<tr><td>☞</td><td>a. |([HáSalit] 9ala-[mínHa] min-is-[sifáara] 9ala-[šáan] [tirúuH] [tídris] fi-[?amríika])|</td><td></td><td></td><td>*******</td></tr>
<tr><td></td><td>b. |([HáSalit] 9ala-[minHa] min-is-[sifáara] 9ala-[šáan] [tirúuH] [tídris] fi-[?amríika])|</td><td>*!</td><td></td><td>******</td></tr>
<tr><td></td><td>c. |([HáSalit] 9ala-minHa min-is-[sifáara] 9ala-[šáan] [tirúuH] [tídris] fi-[?amríika])|</td><td></td><td>*!</td><td>******</td></tr>
</table>

Thus far it has not been possible to establish the ranking of PWD→T and LEXWD:PWD with respect to each other, since in the data examined to date they do not directly conflict in their requirements. There is however some slight variation in accentuation of lexical words. The corpus survey in chapter 3 revealed that 97% of lexical (content) words in EA bear a pitch accent. Exceptional productions, in which a lexical word is unaccented (and not mapped to a PWd) represent only 3% of renditions of lexical words.[152]. In the sentence below (from the read narratives section of the corpus) speakers vary in their treatment of the pre-head adjectival modifier [Tuul] 'all':

(6.36) Speakers' read speech phrasings of a 5PWd monoclausal sentence.

speaker	guHa	kaan	Tuul	9umr	-uh	9aayiš	fi	-l-	?ariyaaf
fna	LH*	---	LH*	LH*	<	LH*	<	<	!LH*
fsf	LH*	<	<	LH*	<	LH*	<	<	!LH* L-L%
meh	LH*	<	(LH*)	LH*	<	LH*	<	<	LH* H-
miz	LH*	<	LH*	LH*	<	LH*	<	<	LH* H-
mns	LH*	<	<	LH*	<	LH*	<	<	LH* L-

guHa	kaan	Tuul	9umr	-uh	9aayiš	fi	-l-	?ariyaaf
Guha	was	all	life-	his	living	in	the	villages
[[NP]	AUX	[AdvP		]	[V	[PP		]]]$_S$

"Guha had lived all his life in the countryside."

The most common rendition (3/5) is the one in which the modifier is accented and is thus as predicted by the grammar established so far (as in (6.37).

(6.37) Three speakers - normal grammar:

	/guHa$_{lex}$ kaan Tuul$_{lex}$ 9umr-uh$_{lex}$ 9aayiš$_{lex}$ fil-?ariyaaf$_{lex}$/	PWD→T	LEXWD: PWD	DEPTONE
☞	a. |([gúHa] ken [Túul]) ([9úmr-uh] [9áayiš]) (fil- [?ariyáaf])|			*****
	b. |([gúHa] ken Tul [9úmr-uh]) ([9áayiš] (fil- [?ariyáaf])|		!*	****
	c. |([gúHa] ken [Tuul]) ([9úmr-uh] [9áayiš]) (fil- [?ariyáaf])|	!*		****

[152] For evidence that unaccented words are also unstressed see section 5.4.4. A number of reasons were suggested for these exceptions, including fast speech rate, high frequency of the word in question (thus loss of semantic and lexical content) or pre-head modifier position (see discussion in section 3.3.2).

For the other speakers, some higher ranked constraint could outrank LEXWD:PWD[153]:

(6.38) Two speakers - exceptional grammar (something outranks LEXWD:PWD)

	/guHa$_{lex}$ kaan Tuul$_{lex}$ 9umr-uh$_{lex}$ 9aayiš$_{lex}$ fil-?ariyaaf$_{lex}$/	PWD→T	CONSTRAINTX	LEXWD: PWD	DEPTONE
	a. |([gúHa] ken [Túul])([9úmr-uh] [9áayiš])(fil- [?ariyáaf])|		*!		*****
☞	b. |([gúHa] ken Tul [9úmr-uh]) ([9áayiš] (fil- [?ariyáaf])|			*	****
	c. |([gúHa] ken [Tuul])([9úmr-uh] [9áayiš])(fil- [?ariyáaf])|	*!			****

If this is the case, these exceptional patterns of accentuation suggest that when the usual winning candidate (a.), which fully satisfies both PWD→T and LEXWD:PWD, is dispreferred for other reasons, it is the candidate that best satisfies PWD→T that wins. Under this scenario, these speakers' renditions would support ranking of PWD→T above LEXWD:PWD as follows:

(6.39) possible ranking: PWD→T >> LEXWD:PWD

A more likely explanation of the speakers' renditions however is simply that some speakers fail to analyse the modifier [Tuul] as a lexical category. In the absence of further evidence at present, the constraints PWD→T and LEXWD:PWD will still be shown as mutually unranked in the remainder of the analysis. A further alternative explanation would be to reverse the ranking between PWD→T and LEXWD:PWD, predicting a language in which lexical words may achieve PWd status (and be stressed) but not be accented. This is not true of EA (see 5.4.4 above), but appears to be true of Spanish, in which approximately 30% of content words are unaccented in spontaneous speech (Face 2003:121-2). This distinction may explain the discrepancy between pitch accent distribution in spontaneous speech in EA and Spanish (EA always accents whereas Spanish shows sensitivity to speech register).

The next section explores variation in the accentuation of function words, which is argued to be due to rhythmic well-formedness constraints.

[153] The higher-ranked constraint in question could plausibly be a NOCLASH constraint, violated by candidates in which pitch accents fall on adjacent or near-adjacent syllables(Nespor & Vogel 1989). Although many of the unaccented content words listed in chapter 3 (as in example 3.11) are indeed followed by an initial-stressed word, further investigation is needed to determine the exact restrictions if any on inter-accent intervals in EA (see also discussion in 6.2.3 and 7.3.2 below).

6.2.3 Variation in accentuation of function words

Function words are quite often ‘promoted’ to PWd status in EA, provided that they are of sufficient prosodic size, as discussed in section 5.4.1 above. The minimal word in EA is bimoraic (Watson 2002), and this is here analysed as being due to a constraint on foot size, FTBIN, which is widely argued for (McCarthy & Prince 1990, Yip 2002).

(6.40) FTBIN Feet must be binary under syllabic or moraic analysis.

Assuming strict layering of the prosodic hierarchy (i.e. that HEADEDNESS is unviolated) every PWd must be composed of at least one well-formed (bimoraic) Ft. Watson (2002:88-89) notes that the minimal word constraint is strictly enforced in EA[154]. Whereas other spoken dialects tolerate subminimal words such as /?ab/ ‘father’ and /?ax/ ‘brother’[155], in EA these words are expanded when pronounced in isolation by gemination: [?abb], [?axx]. As already discussed in section 5.4.1.1, similar repair processes apply to commonly used subminimal function words. Thus /kam/ ‘how many?’ emerges in EA as [kaam], /man/ ‘who?’ as [mi:n], and /ma9/ ‘with’ as [ma9a].

FTBIN must thus outrank not only LEXWD:PWD (hence no subminimal content words in EA) but also whatever mechanism regulates promotion of function words to PWd status in EA. As a result FTBIN is not included in the analysis below, and only function words of sufficient size are considered in the following discussion.

What might account for promotion of function words to PWd status, and which are therefore accented? In the following example (from the retold narrative section of the corpus), the speaker accents the auxiliary verb [?akuun] ‘might’, which we expect, as a function word, to emerge unaccented (as is, for example, the pronoun [?ana] ‘I’):

(6.41)

?ana	mumkin	?akuun	baddii-k	kiilu	bi	balaaš
>	LH*	LH*	LH*	H*	<	!LH* L-L%
I	maybe	could	give-you	a-kilo	for	free
[[NP]	ADV	AUX	[V	[NP]	[PP	]]$_{VP}$]$_{S}$

‘I could maybe give you a kilo for free.’

[154] Also termed an ‘absolute ban’ on ‘degenerate’ non-binary feet (Hayes 1995:85).
[155] Recall that in EA final consonants are extrametrical.

In a similar example (from the read narratives section of the corpus) four of the five speakers accent the pronoun [huwwa] 'he' in the sentence fragment illustrated in (6.42) below.

(6.42)

	guHa	naTT	min	il-	kursi	illi	huwwa	kaan	?aa9id	9alayh
fna	LH*	LH*	<	xxx	LH*	---	xxx	LH*	LH*	LH*
fsf	LH*	LH*	<	<	LH*	<	LH*	<	LH*	LH*
meh	(LH*)?	LH*	<	<	LH* H-	<	LH*	<	LH*	LH*
miz	LH*	LH*	<	<	LH*	<	LH*	<	LH*	LH*
mns	LH*	LH*	<	<	LH* H-	<	LH*	<	LH*	LH* H-

guHa	naTT	min	il-	kursi	illi	(huwwa)	kaan	?aa9id	9alayh
Guha	jumped	from-	the-	chair	that	he	was	sitting	on-it
[[NP]	[V	[PP	[NP		[C	[NP]	AUX	[V	[PP] $]_{CP}]_{NP}]_{PP}]_{VP}]_{S}$

'Guha jumped up from the chair he was sitting on [and]..'

It could be argued that this accentuation is related to emphasis on the pronoun, in its role as head of the relative clause. However one of the speakers (*fna*) omits the head pronoun altogether, suggesting that its semantic weight may in fact be relatively low (the relative clause is still grammatical when the pronoun is omitted). In her production of the example she accents the auxiliary verb [kaan] 'was' instead. This suggests that whatever causes *fna* to accent the function word [kaan] is the same constraint that leads the other four speakers to accent the overt pronoun [huwwa].

Accentuation of a function word implies that it has been assigned PWd status, and assignment of PWd status to a function word incurs a violation of the constraint PWd:LexWd (see section 5.4.2 for discussion of this constraint; cf. also (Selkirk 1996)). Examples such as (6.42) above, which include a sequence of two or more function words, reveal a potential reason for the promotion of function words. A sequence of unaccented function words results in an unusually long sequence of unaccented syllables (recall that there are relatively few long words in EA due to the operation of vowel syncope, discussed briefly in chapter 2 section 2.3.1, so sequences of unaccented syllables are rare).

Cross-linguistically, long sequences of unaccented syllables are often rhythmically ill-formed(Nespor & Vogel 1989), and this tendency can be captured by means of a constraint, NoLapse, which requires regular rhythmic prominences at some level of the prosodic hierarchy. A standard formulation of NoLapse is as a requirement that every weak 'beat' be adjacent to a strong one (Elenbaas & Kager 1999):

(6.43) NoLapse: Every weak beat must be adjacent to a strong beat or the word edge.

In the EA examples above a sequence of four or more syllables seems to be repaired by insertion of an additional prominence (which is best achieved by promotion of a function word to PWd status). A working definition of NoLapse for EA is therefore:

(6.44) NoLapse (EA): Sequences of four or more unstressed syllables are not allowed.

The application of this constraint is illustrated in the tableaux in (6.45) and (6.46) below. Note that an additional constraint would also be required to rule out routine realisation of more than one pitch accent on a single PWd.[156]

[156] There are instances in the LDC spontaneous speech corpus of words bearing more than one pitch accent, such as in a very emphatic expression of relief /?axiiran/ 'at last' which was produced with a pitch accent on [xii] and on [ran] (*4682A 439.04 440.44*).

(6.45) NOLAPSE(EA) >> PWD:LEXWD based on example (6.42) (except speaker *fna*)[157]

	guHa$_{lex}$ naTT$_{lex}$ min il-kursi$_{lex}$ illi huwwa kaan ?aa9id$_{lex}$ 9alayh$_{lex}$	NOLAPSE(EA)	PWD: LEXWD	DEP$_{TONE}$
☞	a. |([gúHa][náTT])(min-il-[kúrsi] illi [húwwa])(kaan [?áa9id] [9aláyh])|		*	******
	b. |([gúHa][náTT])(min-il-[kúrsi] illi huwwa kaan [?áa9id] [9aláyh])|	*!		*****

(6.46) NOLAPSE(EA) >> PWD:LEXWD based on example (6.41)

	?ana mumkin$_{lex}$?akuun baddii-k$_{lex}$ kiilu$_{lex}$ bi-balaaš$_{lex}$	NOLAPSE(EA)	PWD: LEXWD	DEP$_{TONE}$
☞	a. |(?ana [múmkin] [?akúun])([baddíi-k] [kíilu])(bi-[baláaš])|		*	*****
	b. |(?ana [múmkin] ?akuun [baddíi-k])([kíilu] bi-[baláaš])|	*!		****

[157] Lexical categories in the input form are marked with a subscript: "lex". Vowels bearing a pitch accent in output forms are indicated with an acute accent mark on the accented vowel: eg "á" or "ú". The edges of prosodic constituents are marked at **[PWd]**, **(MiP)** and **|MaP|** level.

A potential alternative analysis would be to say that, due to the high-ranking of FTBIN in EA, a sequence of two PWds in EA will very often contain four or more syllables. Recall also that a minimum of two PWds is the working definition assigned in EA to the MiP, a prosodic constituent which is rhythmically rather than morphosyntactically defined (section 5.3.3). It might be therefore be MIPBIN ('MiPs must be binary, that is, composed of two PWds') that outranks PWD:LEXWD, and thus that in order to avoid a non-binary MiP (if there are not enough PWds in the sentence) then an available function word of suitable prosodic size is promoted.

The test of this alternative is whether or not sequences of unstressed syllables *within* a PWd attract an extra pitch accent. If so this suggests that NOLAPSE is the correct analysis. There are very few words containing four or more unstressed syllables in EA, since as noted above polysyllabic cognates in Classical Arabic are frequently shortened due to application of vowel syncope in EA. One exception to this however is the word /miGanawaati/ 'singer' which when prefixed with the definite article contains four unstressed syllables before its own stressed syllable: [il.mi.Gan.na.waa.ti]. This word is found in some of the sentences elicited for the phrasing pilot study whose results were reported in chapter 5.

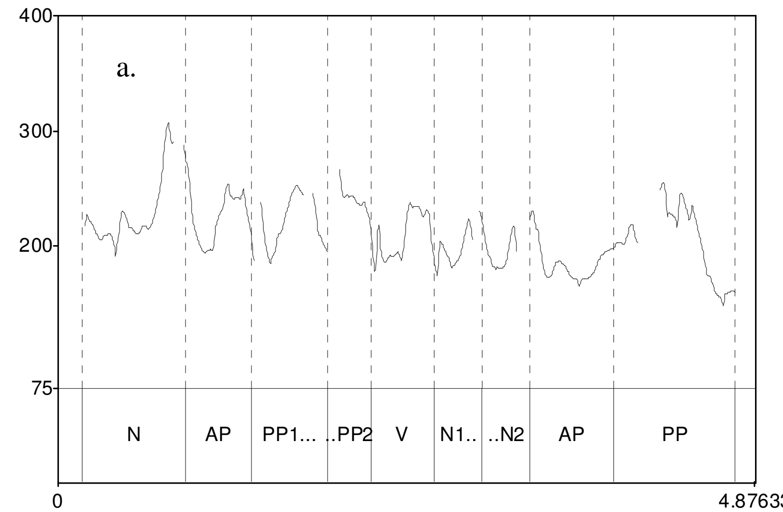
Two tokens of the sentence containing [il.mi.Gan.na.waa.ti] were illustrated in Figure 5.9 (in chapter 5; reproduced below as Figure 6.1) and in both cases there is a single pitch accent on the word. These were however both cases in which the polysyllabic word was sentence-initial, which might be a position less conducive to accent insertion due to tolerance of anacrusis. Figure 6.2 below shows a speech example elicited to test whether sentence-medial polysyllabic words received trigger additional pitch accent insertion due to the effects of NOLAPSE. The word /mutadayyina/ 'devout (f.s.)' syllabifies with the prefixed definite article: [il.mu.ta.da.yi.na][158] so that there are four unstressed syllables between accents. Nonetheless the word bears only a single pitch accent.

[158] The speaker has produced this word with full vowels, as if in Classical or Modern Standard Arabic; in EA the vowel deletion and reduction should result in [mitdayina]. It is possible that the religious meaning of the word has triggered a rendition in a higher register.

Figure 6.1 Sample pair of utterances: speaker MF
Long double-branching subject + long double-branching object

il-miGannawaati	l-muhimm	fi-nihaayit-il-film	biyGumm…
the-singer	the-important	in-end-the-film	upsets…
… banaat 9ammit-i	l-9ayaniin	min-l-?iskandariyya	
…cousins(f)-my	the-ill	from-Alexandria	

	N	AP	PP	V	N	AP	PP
a.	(						)
b.	(		)	(			)



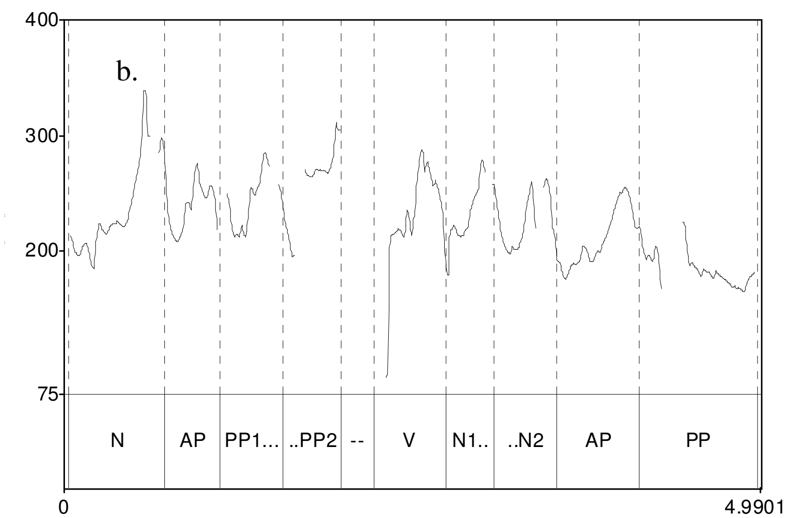
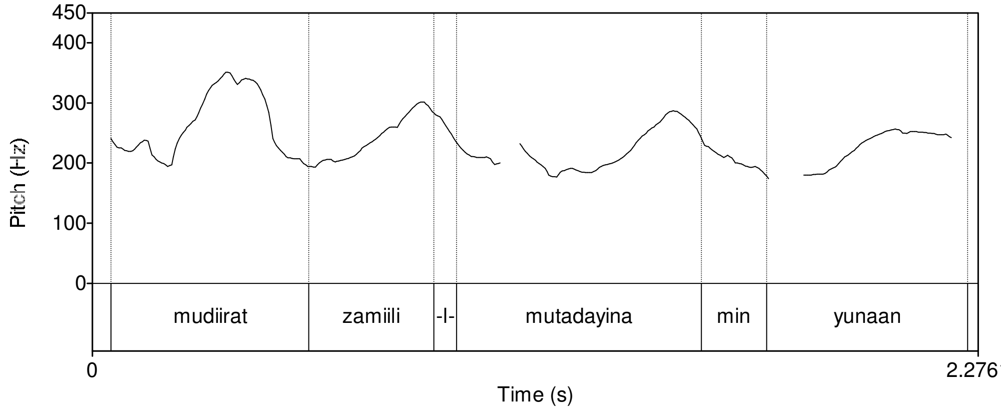


Figure 6.2 Sentence containing polysyllabic word [il.mu.ta.da.yi.na] which bears a single pitch accent.



mudiirat	zamiil-i	-l-	mutadayinna	min	yunaan
the boss (f.)	colleague-my	the	devout	from	Greek

"My colleague's devout boss from Greece..."

However the matter is in need of further investigation since some sequences of four unstressed syllables are tolerated, as in the sentence illustrated in (6.47) and (6.48) below, reproduced from (5.48)[159].

(6.47)

HaSalit	9ala-minHa	min-is	sifaara	9ala-šaan	tiruuH	tidris	fi	?amriika
obtained	at-grant	from-the	embassy	in-order	she-goes	she-studies	in	America

'She got a grant from the embassy to go and study in America.'

(6.48) /([HáSalit] [mínHa]) (min-is-[sifáara] 9ala-[šáan])…/.

This section has explored evidence to suggest that NoLapse is the most plausible explanation for the promotion of function words to PWd status. The exact definition of NoLapse, and indeed clarification of rhythmic well-formedness in EA in general is however a topic raised here for further investigation.

The next section concludes the formal analysis in EA by exploring constraints on the direction of cliticisation of unaccented function words.

[159] Compare also Heliel's (1977:395-6) finding that "the maximum number of syllables between two consecutive stresses in [Egyptian] Arabic is six syllables and the minimum one syllable".

6.2.4 Direction of cliticisation in EA

Finally, the analysis needs to be able to account for the choice in EA to procliticise function words to a following lexical word with PWd status, within a higher phrase-level constituent, rather than to encliticise them to a preceding lexical word. By analogy with Selkirk's (1996) analysis of procliticising weak function words in English, I adopt the view that proclisis arises due to constraints on the alignment of edges of prosodic constituents at different levels. For example, in English, Selkirk argues that proclisis is preferred in English in order that the right edge of every MaP (or Phonological Phrase, in her terminology) is aligned with the right edge of a constituent at the next level below in the hierarchy (the PWd in the notion of the hierarchy used in that analysis). The constraint is formulated as follows:

(6.49)	ALIGN MAP,R:PWD,R:	For any MaP in output representation, align its right edge with the right edge of some PWd.

Given the full range of prosodic constituents assumed here (and in later work by Selkirk), the relevant left/right pair of constraints on MiP edges would be:

(6.50)	ALIGN MIP,R:PWD,R:	For any MiP in output representation, align its right edge with the right edge of some PWd.
	ALIGN MIP,L:PWD,L:	For any MiP in output representation, align its left edge with the left edge of some PWd.

Since MiPs are preferably binary (formalised here as a constraint BINMIP), in order to establish whether the left or right edges of MiPs align with PWd edges, an example containing an odd number of PWds is needed, such as the sentence illustrated in (6.47). In the tableau in (6.51) below, in a sentence with an odd number of PWds proclisis in (candidate a.) satisfies ALIGNMIP,R,PWD,R (unaccented function words cliticise to the following lexical word, indicated by ">") but violates ALIGNMIP,L,PWD,L and BINMIP. In contrast, enclisis (candidate b.) satisfies ALIGNMIP,L:PWD,L (unaccented function words cliticise to the previous lexical word, "<") but violates ALIGNMIP,R:PWD,R and BINMIP. Since both candidates violate BINMIP we can establish that ALIGNMIP,R:PWD,R outranks ALIGNMIP,L:PWD,L. Note that this analysis does not *prove* proclisis to be the correct analysis in EA, since we could assume enclisis and re-rank the constraints. However since it is hard to determine the direction of cliticisation auditorily (see section 5.4.4), proclisis is assumed here in order to demonstrate how the theory might account for the consensus in the literature that proclisis is the norm in Arabic (see inter alia Watson 2002, Shlonsky 1997).

(6.51) ALIGNR MIP,R:PWD,R >> ALIGNR MIP,L:PWD,L based on example (6.41) above

	HaSalit$_{lex}$ 9ala-minHa$_{lex}$ min-is-sifaara$_{lex}$ 9ala-šaan$_{lex}$ tiruuH$_{lex}$ tidris$_{lex}$ fi-?amriika$_{lex}$	BINMIP	ALIGN MIP,R:PWD,R	ALIGN MIP,L:PWD,L	ALIGN MAP,R:PWD,R	ALIGN MAP,L:PWD,L
☞	a. |([HáSalit] 9ala>[mínHa]) (min>is>[sifáara] 9ala>[šáan]) ([tirúuH] [tídris]) (fi>[?amríika])|	*		**		
	b. |([HáSalit]<9ala [mínHa]<min<is-) ([sifáara]<9ala [šáan])([tirúuH] [tídris]<fi) ([?amríika])|	*	**!			

candidate a. has proclisis: satisfies ALIGNR MIP,R:PWD,R violates ALIGNR MIP,L:PWD,L & BINMIP
candidate b. has enclisis: satisfies ALIGNR MIP,L:PWD,L violates ALIGNR MIP,R:PWD,R & BINMIP

6.2.5 Summary of the analysis

This concludes the analysis of the basics of EA pitch accent distribution using tone↔prominence constraints. The data motivate the following ranking for EA:

(6.52) PWD→T, LEXWD:PWD >> DEP_{TONE} >> Ft→T

This is the sub-section of the grammar which is argued to be responsible for rich pitch accent distribution in EA. A candidate which satisfies PWD→T harmonically bounds a candidate which satisfies FT→T, in terms of violations of DEP_{TONE}. This stringency relation is thus confirmed to be inherent to the fixed hierarchy of P→T constraints, assuming strict layering of prosodic constituents such that intermediate levels of prosodic representation are never skipped or out of order (that is to say, that LAYEREDNESS and HEADEDNESS are undominated as assumed in Selkirk 1996[160]).

It was proposed that variable accentuation of function words is due to a preference for avoiding sequences of unstressed syllables (NOLAPSE). Finally an analysis of proclisis (rather than enclisis) of unaccented function words was outlined.

Crucially the analysis shows that it is possible to account for the distribution of pitch accents in EA by means of a purely phonological constraint, requiring a prominence at a certain level of the prosodic hierarchy to be associated with tone. In EA the relevant constituent is the PWd. The interaction of the phonological constraint with other constraints on prosodic structure (interface constraints with morphosyntactic structure at the word-level and rhythmic well-formedness constraints) is shown to result in the surface facts of EA pitch accent distribution.

The next section (6.3) explores potential advantages and implications of analysing EA rich pitch accent distribution in terms of tone↔prominence relations.

---

[160] Recall from chapter 2 (section 2.2.2) that these prosodic domination constraints are defined as follows: LAYEREDNESS:No C^i dominates a C^j, where j>i. [*eg: no σ dominates a Ft]*; HEADEDNESS: Any C^i must dominate a C^{i-1} (except if C^i - σ) [*eg: a PWd must dominate a Ft]*.

6.3 Discussion: a cross-linguistic theory of tone~prominence relations

This section concludes the chapter by discussing some potential advantages, implications and applications of tone↔prominence theory, for EA and other languages.

6.3.1 Tone↔prominence theory and word-prosodic typology

The basic distinction among word-prosodic systems were set out in chapter 4. A key observation is the one made by Hyman (2001) that lexical use of tone is *paradigmatic* whereas postlexical (accentual) use of tone is *syntagmatic*. Recall that for Hyman the definitional feature of a tone language is the fact that the function of pitch in the language is (lexically) distinctive (paradigmatic), whereas the definitional feature of an accentual language is the fact that the function of pitch in the language is contrastive, marking out a single obligatory syllable as most prominent among the other syllables of the word (syntagmatic).

A potential advantage of the formulation of tone↔prominence relations by means of a *pair* of hierarchies as set out here, regulating the relationship of tone to prominence and of prominence to tone respectively, neatly encodes Hyman's observation. The two fixed rank hierarchies are reproduced here from (6.21) above.

(6.53)
a. Tone-to-Prominence constraints [T→P]
T→μ >> T→μ(σ) >> T→μ(Ft) >> T→μ(PWd) >> T→μ(MiP) >> T→μ(MaP) >> T→μ(IP)

b. Prominence-to-Tone Constraints [P→ T]
μ(IP)→T >> μ(MaP)→T >> μ(MiP)→T >> μ(PWd)→T >> μ(Ft)→T >> μ(σ)→T >> μ→T

As noted already in 6.1.3 above, these hierarchies mean that in the unmarked case tones will be associated to moras, and (minimally) the head mora of every IP will bear tone. In the case of lexically contrastive tones, the unmarked case will look like a tone language with every contrastive tone realised on some mora, and the effect will be paradigmatic. In the absence of lexical tones, the unmarked case will be an intonational stress accent language with tone inserted to mark the head of a phrase level domain, and the effect will be syntagmatic.

Jun (2005b:432) similarly notes that postlexical and lexical properties of language are independent from each other: "postlexical prosodic pitch properties cannot be predicted from lexical pitch properties". A language with lexical tone may also have or not have intonational pitch accents and boundary tones[161]. There is no typological implicational relationship that can be inferred from patterns of postlexical and lexical pitch properties observed across different languages, and this supports the analytical need for both the T→P and P→T constraint hierarchies.

6.3.2 Tone↔prominence theory and intonational typology

Section 6.2 shows that tone↔prominence relations can be used successfully to analyse EA, but this is a language which has an accentual system, and thus only provides evidence in support of association of tones to prominent head positions in prosodic structure. What sort of mechanism would have to be added to tone↔prominence theory in order to capture the fact that there are both lexical and postlexical tone languages in which tone is attracted not to the head of a prosodic constituent but to its edge? I suggest that in these languages we see the effects of an ANCHOR constraint which 'pulls' tone away from prominent positions in constituents to their edges, to serve a demarcative rather than culminative function at that level of the hierarchy.

(6.54) ANCHORT:α (R/L) Tones are anchored at the right/left edge of constituent α.

So for example in Korean, in which both left and right edges of MiP level constituents are tonally marked, both ANCHORT:MIP(R) and ANCHORT:MIP(R) would outrank T→MIP (recall that T→MiP is shorthand for 'T→μ(MiP)') and DEP_{TONE}.

An additional advantage of tone↔prominence theory is that it predicts cross-linguistic variation to result not only from interaction of the two sets of T↔P constraints with both faithfulness constraints MAX_{TONE} and DEP_{TONE}; in languages with lexical tone, the role of *T ('avoid tones altogether'). The *T constraint could be ranked differently with respect to the T→P hierarchy than DEP_{TONE}, so that not all underlying tones are allowed to surface: this appears to be what occurs in Japanese and Basque (Gussenhoven 2004). Full exploration of this prediction is however beyond the scope of the present thesis.

[161] She also notes that the types of prosodic units above the word which are tonally marked cannot be predicted from the timing unit used within words: a syllable-timed language may have or not have a tonally marked MiP.

6.3.3 'Epenthetic tone' and insertion vs. spreading of intonational pitch accents

The P→T constraints proposed here require insertion of tone to mark the head of a prosodic constituent if there is no underlying lexical tone available to do the job. Two analytical questions arise from this conception of tone insertion.

Firstly, the 'epenthetic tone' observed to be inserted routinely in EA is a rising pitch movement, which we have to date analysed as a bitonal pitch accent (L+H*). I have thus assumed in the above analysis that the constraints are blind to the type of phonological tone inserted. This conflicts with claims made by DeLacy (2002, 2004) that there is a privileged relationship between more prominent tones (such as H) and prosodically prominent positions. It could however be argued that the most perceptually salient tonal combination is however a rise; if the role of tone in EA is purely functional, for example to facilitate word segmentation as discussed in section 5.5, then insertion of a complex rising tone is perhaps to be expected. The exact phonological representation of the EA rising pitch accent is discussed in detail in chapter 7 below.

Secondly, recall that Gussenhoven (2004) analysed tone spreading in Roermond Dutch as arising due to a constraint requiring every TBU to be associated to tone, and he thus argued for a careful distinction between association and alignment of tones (see section 6.1.1 above)[162]. This raises the question of how to explain why we see insertion of tone in EA, in the form of 'default' pitch accents, rather than tone spreading? I suggest that, by analogy with analyses of tone insertion vs. tone spreading in tone and lexical pitch accent languages, the explanation lies in the relative ranking of a constraint NOSPREAD, which requires tones to associate to at most one constituent (at some level of the hierarchy) (Gussenhoven 2000, Yip 2002, cf. Selkirk 2005b).

The issue of tone insertion vs. tone spreading in relation to EA may throw light on the problem of how to analyse phrase-final 'nuclear' pitch accents in EA. A key explanatory advantage of T↔P theory for EA is that it formalises the notion that pitch accents in EA are purely 'epenthetic' (inserted to meet a phonological constraint), and thus in turn be said to predict the fact that the language will have a small pitch accent

[162] Alternative constraints such as ALIGN(T,TBU) & ALIGN(T,TBU) would only be able to re-position tones already available in the representation; cf. Gussenhoven (2004:149, 2000:section 5.4)

inventory: all pitch accents in EA are 'default' pitch accents[163]. Nonetheless, as discussed in chapter 3 section 3.4.3, phrase-final 'nuclear' pitch accents frequently display a very different shape to that observed in the ubiquitous rising pitch movement on all pre-nuclear PWds. Instead, in final position the pitch accent seems often to be falling (Rifaat 1991), and Rifaat (2004) thus proposes a distinct final pitch accent type, phonotactically constrained to occur only in phrase-final position: HL#[164]. In chapter 3 I proposed that the final pitch accent was instead an instance of the default pitch accent LH* but with an early peak due to proximity to the utterance boundary (cf. Prieto et al 1995, Chahal 2001) as well as tonal crowding from upcoming boundary tones (L-L% or H-H%).

In Spanish and Greek, which are observed to share EA's property of having rich pitch accent distribution (Jun 2005b), some authors have proposed distinct pre-nuclear and nuclear pitch accents. For example in Spanish the pre-pitch accent observed on every content word is L*+H whereas in nuclear position the pitch accent has an earlier peak and is analysed as L+H* by Face (2002:19-20)[165]. In Greek, the predominant pitch accent is analysed as L*+H, but a greater variety of pitch accents are seen in nuclear position, including H* and L* (Arvaniti & Baltazani 2005).

The P→T constraints permit two possible analyses of EA nuclear pitch accents therefore. Firstly, one could follow Rifaat (and other authors for Spanish and Greek) and assume that the nuclear pitch accent is indeed of a different shape. In a P→T analysis however there is no need to propose a distinct 'nuclear' pitch accent, phonotactically restricted to utterance-final position. Instead one simply states that PWD→T outranks NOSPREAD in EA. In utterance-final position there is no need to insert a pitch accent to fulfil the need to mark the prominence of the final PWd, instead the final phrase tone of the utterance will 'spread', that is associate to the final stressed syllable as well as align to the phrase boundary. This analysis would avoid the problem faced by a 'distinct pitch accent' analysis of how to explain the fact that in phrases ending with high boundary tones there is no L target or fall between the final stressed syllable and the phrase-final H-H% combination. In a 'distinct pitch accent' analysis

[163] The question arises why we do not see a small pitch accent inventory in English, in which epenthetic tone is arguably inserted to mark every MiP. Selkirk has explored the role of morphemic tone in English (Selkirk 2005b), and this might go some way to explaining additional pitch accent variety in English.
[164] Where # indicates adjacency to an utterance boundary.
[165] Other authors have proposed a single pitch accent in Spanish, with the early peak in nuclear position ascribed to boundary effects (e.g. Nibert 2000).

one would have to stipulate that the nuclear accent only occurs with low boundary tones, whereas in a 'phrase-tone spreading' analysis co-variance between the properties of the final pitch accent and the following boundary tones is expected.

Alternatively, one could maintain the view that EA has a single pitch accent in its inventory (LH*) and explain the choice to insert that pitch accent in final position rather than allow spreading of phrase tones by the following ranking: NOSPREAD >> PWD→T. This analysis captures the fact that a L turning point is observed between the peak of the penultimate stressed syllable of the utterance and the H peak of the final stressed syllable[166]; this pitch valley goes unexplained in a 'phrase-tone spreading' analysis.

I suggest that the single pitch accent inventory analysis, based on NOSPREAD >> PWD→T, is the stronger of the two because of the facts of phrase-final L turning points. However further investigation of the properties of EA nuclear pitch accents, and of the distribution of various pitch shapes observed in that position, may yet reveal evidence to support the alternative analysis. For our present purposes it is however useful to note that a T↔P analysis affords insights into potential new explanations for puzzling facts of EA intonation.

6.3.4 Pitch accent distribution as an independent parameter of prosodic variation

If tone↔prominence theory is correct we should expect to see a great deal of variation cross-linguistically in the nature of the relationship between phonological tone and prosodic structure. Taking a specific area of variation, what evidence is there for variation across the prosodic hierarchy in P→T constraints? That is, if we see the effects of PWd→T in EA, in which languages do we see the effects of other constraints in the P→T hierarchy? A full answer to this question lies beyond the scope of this thesis, which had as its primary goal to establish the empirical facts of EA pitch accent distribution. However as already seen in chapter 5, there is evidence to suggest that there may be variation in the domain of pitch accent distribution in intonational languages, with both MiP and MaP claimed to be relevant domains; in tone languages it is uncontroversial to talk of variation in the TBU between mora and syllable. The specific predictions of the theory however, such as the existence of, say, a language in which we see the effects of Ft→T, must remain the subject of future research.

[166] See for example Figures 5.11-13 in chapter 5.

Nonetheless, the T↔P theory does make a strong claim about the nature of the mechanism behind pitch accent distribution, placing it firmly within the phonological component of the grammar. In this conception, variation in the domain of pitch accent distribution is a completely independent parameter of prosodic variation, and this makes the prediction that that there is not necessarily any correlation between pitch accent distribution and other prosodic factors.

The remaining two chapters of the thesis test whether this prediction holds of EA by examining other potentially correlated factors. Chapter 7 explores whether rich pitch accent distribution affects surface alignment properties in EA pitch accents. This is a test of the independence of pitch accent distribution as a prosodic parameter if we adopt the view that T→P constraints regulate the distribution of all tones, whether lexical or postlexical. Specifically T↔P theory predicts that the domain of pitch accent distribution may vary independently of the choice of 'TBU'. Chapter 7 seeks to determine what the TBU is in EA, on the basis of surface pitch accent alignment properties. If the TBU in EA is similar to that observed in languages which display very different pitch accent distribution, then this provides support for the prediction that pitch accent distribution is a truly independent parameter of prosodic variation.

Chapter 8 explores the prosodic reflexes of information structure in EA. As discussed in chapter 5 (section 5.1.1), many authors have argued that pitch accent distribution is not a phonological parameter at all but a reflex of information structure (as observed in Germanic languages for example). If there is no link between pitch accent distribution and any aspect of information structure then this supports the prediction that pitch accent distribution is a truly independent parameter of prosodic variation.

In effect then, chapters 7 & 8 seek to support the claim that the requirement to have a pitch accent on every PWd is purely phonological (due to PWD→T) as argued in this chapter, by excluding the possibility that the requirement is instead a by-product of some other component of the grammar (such as phonetics or syntax/semantics).

6.4 Summary

This chapter reviewed mechanisms which have been proposed in the literature to express the fact that the relationship between phonological tone and prosodic prominence is a two-way relation, and that the relevant prosodic constituent whose head attracts tone may vary. The attraction of tones to prominent positions at higher levels of the hierarchy was explored in order to find a way of explaining the empirical fact that EA marks every PWd with a pitch accent, rather than marking a phrase-level constituent as observed in many other intonational languages.

The properties of a particular conception of tone-prominence relations were described - tone↔prominence theory - in which surface relations between tone and prosodic prominence result from the interaction of a pair of inherently ranked fixed hierarchies of markedness constraints which regulate association of tone to prosodic prominence, and of prosodic prominence to tone, respectively.

A formal analysis of data from the EA thesis corpus established the section of the grammar which results in EA rich pitch accent distribution: it is better to insert tones than to leave PWds unaccented in EA (PWD→T >>DEP$_{\text{TONE}}$) and it is better for a lexical word to lose its PWd status and go unstressed than to be a PWd and be realised without an accent (PWD→T >>LEXWD:PWD). The stringency relation between individual P→T constraints at different levels of the hierarchy was confirmed (a candidate satisfying PWD→T harmonically bounds a candidate satisfying FT→T). Variation in accentuation of function words was argued to be due to the effects of a rhythmic well-formedness constraint on the proximity of accents (NOLAPSE). Proclisis of unaccented function words was ascribed to a constraint requiring a PWd at the right edge of every MiP level constituent: ALIGNMIP,R, PWD,R.

Key advantages of tone↔prominence theory were argued to be its potential for encoding aspects of word-prosodic typology (paradigmatic vs. syntagmatic tone) and the possibility of an explanation for the distinct realisation of pitch accent in final and non-final positions in the phrase (tone insertion vs. tone-spreading).

The main contribution of the analysis is however that it demonstrates that the EA pitch accent distribution can be analysed by means of a purely phonological constraint on the relations between prosodic prominence and phonological tone. Thus pitch accent distribution, as an parameter of prosodic variation, is predicted to be independent of other aspects of the grammar.

The next two chapters test this prediction empirically by investigating whether there is any necessary correlation between rich pitch accent distribution and other factors: in chapter 7 pitch accent alignment properties are explored in order to determine whether the association properties of tones to prominence (the TBU) are indeed independent of the attraction of tone to prominence; chapter 8 explores the prosodic reflexes of focus in EA to determine whether or not pitch accent distribution is truly independent of information structure.

7 Pitch accent alignment in Egyptian Arabic

7.0 Outline and aims

The interaction of tone and prominence is argued in chapter 6 to be governed by markedness constraints. These constraints are by definition indifferent as to the origin of the tones whose distribution they constrain, that is, whether those tones are present in the input (lexical) or not. This property of the tone↔prominence hierarchy captures formally the 'unity of pitch phonology' expressed representationally in autosegmental-metrical (AM) theory.

Even in intonational languages then, in which all tones are 'postlexical' and absent from the input, the theory predicts that the tone bearing unit (TBU) may vary, since tone could in principle be associated to prominence at any level of the prosodic hierarchy. A standard assumption in AM theory is that pitch accents associate with the main stress of the word which is located in the main stress foot and that this association is inherited by the stressed syllable. This chapter seeks to establish whether the syllable or the foot is the TBU in EA, on the basis of experimental evidence from newly collected data.

In addition, as has been seen the overwhelming majority of pre-nuclear pitch accents in EA were observed to be of a single type, namely 'rising' pitch accents. The present experiment should facilitate proposal of a plausible phonological representation of this ubiquitous pitch accent, since the surface F0 *alignment* properties of pitch accents across syllable types have been argued to allow generalisations to be made about the inherent phonological *association* of EA pitch accents. The results of the study thus also contribute to a growing body of research into cross-linguistic variation in the surface phonetic alignment properties of rising pre-nuclear pitch accents.

Section 7.1 outlines the findings of prior research on the notion of an underlying phonological association of pitch accents, inferable from surface phonetic alignment. A summary is provided of what is known already about EA rising pitch accents. The null and alternative hypotheses of the experimental investigation are set out in detail.

Section 7.2 describes the design of the experiment, in which the prosodic weight of the stressed syllable of target words was systematically varied (light vs. open vs. heavy syllables), as well as the methodology used to collect and analyse the data.

Section 7.3 presents the results of the investigation of alignment across syllable types which reveals the precise patterns of alignment of L and H targets in EA rising pitch accents. These are broadly similar to results observed for Dutch, in that alignment of the H peak appears to vary with stressed syllable type (Ladd et al 2000). The EA results are compared to those reported for alignment of pre-nuclear pitch accents in other Arabic dialects and also in other languages.

In section 7.4 a phonological representation for EA pitch accents is proposed based on their general alignment patterns and perceptual properties: it is argued that EA rising pitch accents consist of a single phonological object (L+H*) phonologically associated with the foot as TBU. This incorporates the notion of a TBU in EA into the tone↔prominence analysis proposed for EA pitch accent distribution in chapter 6. It is argued that in EA, as well as in Dutch, association of the pitch accent to the foot arises because the constraint requiring any tones to be associated to foot as head of the PWd: [T→PWD] is ranked higher than DEP_{TONE}, which mitigates against the insertion of postlexical tones. Hence it is possible for Dutch and EA to share the property of having pitch accents which associate to the foot, but not to share the property of rich pitch accent distribution.

7.1 Background

The experiment reported in this chapter has two goals. The primary goal is to identify the 'tone-bearing unit' (TBU) in EA. In terms more widely used in intonational phonology, this means to identify the underlying phonological association of EA rising pre-nuclear pitch accents, from the surface phonetic alignment of pitch targets. In so doing the study also contributes novel data, from a new language, and from a range of syllable types, to the theoretical debate regarding the phonological mechanisms which might underlie cross-linguistic variation in tonal alignment. The secondary goal is to determine how best to interpret the surface tonal alignment patterns of pre-nuclear rising accents in EA and thus establish a plausible phonological representation for them.

7.1.1 Pitch accent alignment in autosegmental-metrical (AM) theory

Pre-nuclear (i.e. non-final) rising pitch accents are observed in many languages, such as English (Pierrehumbert 1980), Dutch (Ladd et al 2000) and German (Atterer & Ladd 2004). In Greek (Arvaniti et al 1998, Arvaniti & Baltazani 2005) and Spanish (Face 2002, Nibert 2000, Prieto et al 1995), as in EA, the rising pitch accent occurs on (almost) every content word in non-phrase final positions (cf. chapter 3 and Jun 2005b). Rising pre-nuclear pitch accents are also reported in other dialects of Arabic such as Lebanese Arabic (Chahal 2001) and Moroccan Arabic (Yeou 2004).

Within the general framework of autosegmental-metrical (AM) theory (Ladd 1996), autosegmental tonal events (pitch accents and boundary tones) are associated with the heads and edges (respectively) of metrical or prosodic constituents. Pitch accents may be monotonal or bitonal, composed of one or two phonological tones, known as pitch targets (H and L). The underlying phonological *association* of the pitch accent is widely thought to be reflected in the surface *alignment* of F0 turning points to landmarks in the segmental string. There have been various theoretical positions as to how the alignment properties of rising pitch accents should be phonologically interpreted and thus as to how cross-linguistic alignment variation is to be captured.

In one school of thought (emanating from the analysis in Pierrehumbert 1980) there are two possible phonological representations for rising pitch accents, L*+H and L+H*, in which the rise occurs either early or late in the stressed syllable respectively. The starred tone (*) notation indicates which of the two pitch targets, L or H, is aligned with the stressed syllable, and the other tone is said to 'lead' or 'trail' at a fixed distance from the starred tone. Under this view cross-linguistic differences in alignment arise from differences in the underlying phonological representation of the rising accent in a particular language.

The simplicity of this analysis was challenged by Grice (1995a) who argued that the details of cross-linguistic variation in alignment reveal the need for a distinction between: i) a Pierrehumbert-style 'contour' pitch accent, in which only one of the two tones is associated with some phonological landmark, and ii) a 'cluster' pitch accent in which both tones are associated to segmental landmarks at either edge of the stressed syllable. There is now a body of evidence to suggest that individual L and H pitch targets in a rising pitch accent may each display independent 'segmental anchoring' to landmarks in the segmental string (Arvaniti et al 1998).

Further research has thus since focussed on how best to understand the mechanism underlying segmental anchoring. One view states that whilst the whole pitch accent is phonologically associated with the stressed syllable, individual tones (pitch targets) may show ‘secondary association’ to fixed points in the segmental structure (Grice et al 2000, Gussenhoven & Rietveld 2000). In this view cross-linguistic differences in alignment are due to differences of secondary association, and variation is predicted to be categorical. An alternative view states that a rising accent could in fact be the same phonological object in different languages, but show differences in phonetic implementation giving rise to differences in the detail of tonal alignment (Atterer & Ladd 2004, Ladd 2004). This latter view predicts a continuum of cross-linguistic variation in tonal alignment and has been likened to parallel cases in segmental phonology such as a single underlying phonological feature [+voice], but which is interpreted phonetically quite differently in different languages by means of variation in voice onset time (VOT).

Other recent research has suggested that the segmental landmarks to which individual pitch targets are aligned may be structurally defined rather than purely segmentally defined (a nuance which I will distinguish by the term ‘structural anchoring’). In Dutch for example, the H peak has been shown to align at “the end of the syllable” in both CV and CVV syllables (Ladd et al 2000:2693). The authors of that study note that for alignment purposes in Dutch the ‘end of the syllable’ falls “in the following consonant if the vowel is short” (ibid.). This detail raises the possibility that the relevant structural domain in Dutch is perhaps not in fact the syllable but the metrical stress foot (a sequence of two light syllables form a moraic trochee in Dutch, Hayes 1995). Similarly, Ishihara (2003) found that alignment of the H peak in Japanese pitch accented words (with accent on the initial syllable) was phonologically rather than segmentally conditioned, occurring consistently at the start of the second mora, across a range of syllable types (CV.CV, CVN and CVV). Compare also the possibility that English pitch accents may display phonological association to a domain larger than the syllable, as discussed in 6.1.2 above (Ladd 2005).

Recent research has thus suggested at least three possible mechanisms that might explain surface phonetic alignment, each of which make different predictions about the nature of the alignment relation between bitonal pitch events and the segmental string.

An analysis using contour tones predicts one pitch target event to be very stable with the other following or preceding at a fixed distance, and this might be termed the *fixed* *duration* hypothesis[167]. The *segmental anchoring* hypothesis predicts the L and H pitch targets to display independence of alignment, to segmentally defined landmarks. The *structural anchoring* hypothesis is similar but predicts that the pitch targets align to the edges of some structurally defined domain. A fourth possible mechanism, the *fixed slope* hypothesis, arises from alternative theories of intonation which analyse F0 events as a sequence of contours rather than levels (such as Ashby 1978), predicting that the degree of F0 excursion in a rising accent will co-vary with its duration and thus that a rising pitch accent will have a fixed slope.

The analysis of EA proposed in this thesis, by means of tone↔prominence theory (as discussed in chapter 6), brings a further hypothesis to the debate. The prediction of T↔P theory is that cross-linguistic variation in surface alignment may reflect variation in underlying phonological association of phonological tones to different constituent levels of the prosodic hierarchy as TBU. This claim appears however to be compatible with the structural anchoring hypothesis, on the assumption that the structurally defined domain to which pitch accents display alignment is a constituent of the prosodic hierarchy (such as the foot or the syllable). What is not known is whether this underlying association is to the head or the edge of the domain, and how, in either case, association would be reflected in surface alignment.

Cross-linguistic variation in the alignment properties of rising pitch accents could in principle be due to the operation of different mechanisms (from among those listed above) in different languages. The strongest hypothesis to test is the claim that just one of these mechanisms can explain the alignment properties of all intonational languages, with variation a matter of differences in phonetic implementation (cf. Atterer & Ladd 2004). These questions are particularly apposite when approaching a language or dialect whose alignment properties have yet to be established since in principle any of these mechanisms could be relevant.

This chapter takes structural anchoring (to a constituent of the prosodic hierarchy as TBU) as its working hypothesis and in particular explores the possibility that underlying association in EA may be to the foot instead of to the syllable. In addition, in providing detailed data from a language new to the typology of alignment variation, the chapter

[167] Similar predictions are arrived at for entirely different reasons in the Fujisaki model of Japanese prosody (Fujisaki & Sudo 1971).

also explores to what extent the strong hypothesis of a single mechanism behind alignment variation can be maintained.

Pursuit of the structural anchoring hypothesis is inspired in part by Ladd et al's (2000) finding that patterns of phonetic alignment are *phonologically* conditioned in Dutch . Ladd et al. were able to make this claim conclusively due to a particular property of certain Dutch vowels. In Dutch the 'long' high vowels /i:/, /y:/ and /u:/ are in fact phonetically quite short. Ladd et al tested the alignment of pitch targets to stressed vowels in phonological near minimal pairs, containing a phonologically short or phonologically long vowel flanked by segmentally parallel consonants, such as 'striemende' [ˈstriːməndə] (*pouring*) ~ 'trimmende' [ˈtrɪməndə] (*jogging*).

Comparison of the mean duration of the two types of vowel showed no significant difference between the two, yet there was a significant difference in alignment, with the F0 peak aligned earlier in phonologically long syllables containing [iː], than in phonologically short syllable containing [ɪ][168]. Ladd et al therefore analyse the Dutch rising pitch accent as having two anchor points: the initial L turning point is aligned stably at the beginning of (the onset consonant of) the stressed syllable, and the H peak is aligned at the end of the stressed syllable. It is unfortunately not possible to reproduce this exact experimental paradigm in EA because all qualities of phonologically long vowel are phonetically long in the language[169].

Ladd et al make it clear that (2000:2693):

> "in our materials, 'the end of the syllable' is at the end of the vowel if the vowel is phonologically long, and in the following consonant if the vowel is short".

The choice to generalise over the two contexts of 'the end of the vowel' in a long vowel and 'in the following consonant' after a short vowel, and to describe them as being structurally parallel - 'the end of the syllable' - involves accepting an ambisyllabic analysis of the intervocalic consonant in a CVCV sequence in Dutch. For example, in

[168] Ladd et al (2000:2693) also found evidence of time pressure on H alignment in phonologically long stressed syllables, which they suggest is due to time pressure to realise a correctly scaled F0 peak in the restricted time available.

[169] This is demonstrated, with respect to the experimental data reported here, in Figure 7.5 below.

the target above containing a short vowel, 'trimmende' [ˈtrɪməndə], this means stating that the [m] is ambisyllabic and is structurally affiliated to *both* foot internal CV syllables in the CVCV head foot of the word (a moraic trochee). This is a plausible assumption for Dutch short vowels in this position, since 'short' or lax vowels in Dutch are required to occur in a position which is "in close contact with" a following consonant (Hayes 1995:306).

I would like to suggest that Ladd et al's generalisation is not the only one that could be applied to the set of alignment patterns observed in Dutch. One could equally well reduce the two contexts ('the end of the vowel in a long vowel' and 'in the following consonant after a short vowel') to 'during the second mora in a bimoraic trochee'. This opens up the possibility that the alignment in Dutch is in fact evidence of association not to the syllable but to the foot[170]. This is consistent with arguments in the phonological literature that apparent effects of ambisyllabicity can be successfully reanalysed as foot-internal phenomena[171].

In order to test whether the sensitivity to phonological structure observed Dutch is reproduced in EA the experimental investigation described in this chapter elicited targets containing three types of stressed syllables of forms CV, CVC and CVV. Specifically, testing three types of syllables makes it possible to determine whether such alignment in EA (if found) indicates phonological association to the syllable or to the foot. Yip (2002:133) notes that in a tone language in which all syllables are monomoraic it is impossible to determine whether the mora or the syllable is the TBU, since all syllables contain an equal number of moras and syllables. To test in EA whether the syllable or the foot is the TBU we need to look at alignment in targets of association in which there is a mismatch between the number of syllables and the number of feet: such an example would be a bimoraic foot composed of two open syllables.

The foot in EA is the bimoraic trochee and stress may in principle be attracted to both monosyllabic and bisyllabic feet of any of the following kinds. Monosyllabic feet are CVC or CVV heavy syllables. Bisyllabic feet are always a sequence of two light

[170] Note that the initial L target in fact aligned just before or just after the beginning of the consonantal onset of the stressed syllable (Ladd et al 2000:2689 Table 1).
[171] Re-analysis of ambisyllabicity as foot-internal phenomena is discussed further in section 6.4.1 below (Jensen 2000, Harris 2004).

syllables: CVCV. This gives us an ideal testing ground to determine the TBU, assuming that surface alignment of pitch targets proves to display sensitivity to phonological structure in EA. If the foot is the TBU in EA, then the patterns of alignment observed in CV stressed syllables (which are initial in a CVCV bisyllabic foot) will be different from those observed in CVC and CVV syllables (in a monosyllabic foot). If the syllable is the TBU then patterns of alignment are not expected to vary significantly across the three stressed syllable types.

The primary hypothesis of the experimental investigation described below is therefore that the TBU in EA is the foot, with the alternative hypothesis that the TBU is the syllable. This hypothesis is consistent with what I have described above as the 'structural anchoring' hypothesis. In addition the predictions of the remaining three main hypotheses entertained in the literature will also be tested against the EA data: the fixed duration hypothesis, the fixed slope hypothesis and the segmental anchoring hypothesis. The means by which all four hypotheses are tested are outlined below in section 7.2.3.

7.1.2 Surface alignment and the phonological representation of pitch accents

The literature outlined in section 7.1.1 above shares the assumption that the details of surface phonetic alignment of F0 movements can give us information about the underlying phonological affiliation of phonological pitch targets.

Much of the debate above has therefore been part of a wider discussion over the correct way to translate the alignment behaviour of pitch targets into a phonological representation for the pitch accent concerned. There are two schools of thought which express objection to this assumption however.

Firstly, there are those who dispute the validity of surface phonetic alignment as a source of any reliable information at all about phonological affiliation (Xu 2002, Xu & Liu 2005). Xu argues that surface F0 movements may well reflect the results of an attempt to produce phonologically specified targets but that it is unwise to assume that the actual F0 movements observed do not also reflect some degree of articulatory compromise[172]. The rate at which the human voice can vary pitch is limited (Xu 2002)

[172] Xu et al suggest that the phonological target in question is prosodically defined (the syllable) but argue that it is always the syllable that is the target of tonal association, in *all* languages(Xu & Liu 2005).

and as a result Xu suggests that not every aspect of an aimed for target will be produced (the F0 peak may be scaled too low, or the position of the turning point aligned too early or too late). There is however awareness of articulatory constraints on the production of tonal targets in the general alignment literature (such as that as outlined in section 7.1.1), with specific studies being implemented to determine the extent of the effects of tonal and prosodic context on tonal scaling and peak alignment (for example Prieto et al 1995, Prieto 2003).

A second source of dissent regarding the extent to which alignment data provide reliable information about phonological association, relates more specifically to the task of designating a phonological representation for a particular pitch accent, as observed in a particular language. As already noted, many of the alignment studies mentioned above were undertaken with the express aim of working out whether the phonological categories proposed within AM theory for pitch accents (notions such as the associated starred '*' tone, as well as leading and trailing tones) are in fact adequate. For example Grice (1995b) sought to determine whether leading/trailing tones always precede/follow the starred tone at a fixed distance in all languages, as had been tacitly assumed. Similarly, Arvaniti et al (2000, 1998) have sought to determine the degree to which surface phonetic alignment can be captured by means of the notion of the starred tone. They found that Greek rising pitch accents were composed of two tones (L and H) but that neither of these tones exhibited surface alignment patterns that could be said to be 'aligned with the stressed syllable'. The L was aligned consistently just before the beginning of the onset of the syllable, whilst alignment of the H target was more variable but was usually in 'the posttonic syllable'[173]. The authors in fact conclude that "association cannot be based on phonetic alignment in any straightforward way and .. a more abstract and rigorously defined notion of starredness is required" (Arvaniti et al 2000:121).

Prieto et al (2005) have thus recently advocated a return to a more perceptual basis for designating the phonological representation of pitch accents. They suggest that the basic or primary association of a pitch accent is between "the *whole* accent and its tone bearing unit" (italics mine, Prieto et al 2005:section 3). In a bitonal accent the star '*' notation should be assigned "on perceptual grounds", that is, by how native listeners designate the accent (as a peak or a valley).

[173] The study used target words containing open CV stressed syllables such as [paˈɾanoma] 'illegal'.

These differences in interpretation lead to different approaches to the task of assigning a phonological representation (such as L+H* or L*+H) to a rising pitch accent. For example, Face (2002) argues that the pre-nuclear rising pitch accent in Madrid Spanish is L*+H, assigning the star to the L target and thus designating it as the target associated to the stressed syllable. Having observed that the F0 valley was aligned stably at the onset of the stressed syllable but that alignment of the H target was more variable, Face argues that: "since the associated tone has priority in establishing the alignment of the pitch accent, this analysis explains the consistent alignment near the onset of the stressed syllable" (Face 2002:19, cf. Pierrehumbert & Beckman 1988:125). An alternative analysis claims that the same Spanish pitch accent is (L+H)*, whereby the whole tonal sequence is associated to the stressed syllable (Hualde 2002). Face rejects this because it fails to explain the stability of L against the instability of H, and because it allows for a possible three-way contrast among rising accents: L*+H/L+H*/(L+H)*.

Prieto et al propose a different designation for the rising pitch accent in Catalan, even though it displays very similar alignment properties to those observed by Face in Madrid Spanish. They propose L+H* for the most frequent pre-nuclear Catalan pitch accent, assigning the star to the H tone because it is perceived as high by native listeners. They argue that: "in bitonal accents.. the stronger tone will be starred according to perception of the prominent syllable: that is, depending on whether the prominent syllable is heard with a 'high tone' or with a 'low tone' by native speakers of the language" (Prieto et al 2005 §3.1).

Prieto et al suggest that surface alignment would be determined from underlying phonological association of the whole accent as follows. Following Pierrehumbert and Beckman (1988), the main association is 'central', that is to the head of the TBU. Indeed they suggest that in English association is to the foot ("English permits at most one pitch accent per metrical foot" (ibid.)) and this association is passed on to the head of the foot, that is, to the stressed syllable. They suggest that surface implementation of this underlying association results from the following mapping procedure: "both the left and right edges of accented syllables are the basic anchor points for target tones in pitch accents. The starred tone will first be aligned to the right periphery of the metrically prominent syllable... leading tones will be linked to the left edge of the syllable, .. trailing tones will be linked somewhere in the right periphery of the syllable" (ibid.).

Prieto et al's view of the designation of pitch accents appears to be highly compatible with the 'structural anchoring to TBU' hypothesis adopted in this chapter, since they emphasise the notion that it is the whole accent which is phonologically associated with (the *head* of) some unit of prosodic structure.

Once the facts of EA alignment are known they will be compared against the predictions of these various suggestions about how best to determine the phonological specification of rising pitch accents, and a designation proposed.

7.1.3 Pitch accent alignment in Egyptian Arabic (EA)

In prior studies of EA intonation, both descriptive and instrumental, there are indications that EA pre-nuclear accents are rising. Mitchell's (1993) pronunciation guide describes the EA pitch contour as a sequence of 'see-saw jumps' and even notes that "pitch dips markedly.. [on] pre-accentual syllables.. from which a 'jump' takes place to the height of the.. accented syllable" (Mitchell 1993:222-3).

Norlin (1989:47) describes the pitch accent as "a phonological High", which is realised as a peak or as a "rise towards the peak". He notes that the peak tends to be at or near the end of the stressed syllable. Similarly, in another instrumental study, Rifaat (1991 ch.1) notes that non-final EA words have "a late peak situated on the last point of the [stressed] syllable.. [and] are all rising".

In a later phonological study of the EA pronunciation of Standard Arabic, Rifaat (2004) notes the same descriptive generalisation. His phonological analysis of the pre-nuclear accent is as monotonal H*.

As discussed in chapter 3 (section 3.4.3.2), El Zarka (1997:239ff.) describes very similar surface alignment facts, analysed by means of alignment to a 'Tone Domain' which is composed of a stressed syllable and all following unstressed syllables. She proposes that the EA pitch accent is H*L: the H target aligns to the start of the Tone Domain (that is, to the stressed syllable) and the L target aligns to the end of the stressed syllable (that is, just before the next stressed syllable). As discussed in chapter 3, this analysis although conceptually different still predicts an essentially rising pitch movement localised around the stressed syllable.

As demonstrated in chapter 4 (section 4.2.3), investigation of pre-nuclear accents on CVC stressed syllables in word-medial position, confirms the presence of a clear L target at or near the start of the stressed syllable. The L valley (labelled 'L2') following the H peak always occurred inside the following PWd, and displayed association to the onset of the stressed syllable of that following word. It is thus appropriate to investigate EA pre-nuclear accents as *rising* pre-nuclear accents, and make cross-linguistic comparison therefore with the alignment properties of pre-nuclear accents shown to be rising in other languages.

7.2 Methodology

7.2.1 Materials and data collection

The experiment measures the alignment of pitch peaks and valleys against segmental landmarks in stressed syllables occurring initially in words of 2 or 3 syllables. The methodology is modelled on that of Atterer & Ladd (2004) and Ladd et al (2000).

In order to test the 'structural anchoring to TBU', and specifically the hypothesis that the TBU is the foot (or syllable), target syllables elicited are of three types: short open (CV), short heavy (CVC) and long heavy (CVV). To facilitate location of F0 events in the pitch track, target words were sought in which the flanking consonants to the stressed vowel are sonorants [l], [m] or [n]; the limited number of such lexical items in EA means that there is variation in the quality of the stressed vowel across targets ([a], [i] or [u]). There were six targets of each syllable type.

Each target word was placed early in a sentence frame to elicit a pre-nuclear pitch accent. The 18 target sentences were interspersed with distractors and presented to speakers typed in Arabic script, ten sentences to a page; the first and last sentence on each page was always a distractor sentence. EA spelling conventions, and lexical items exclusive to the dialect, were used to correctly elicit spoken EA dialect and to minimise register interference from the use of written prompts (which in the Arabic diglossic situation are ordinarily associated with higher register Standard Arabic); all of the speakers found it easy to produce the targets as elicited in EA dialect. A list of the target sentences is provided in the table in (7.1) below.

(7.1) Alignment target words in their carrier sentences
(Target words in *italics* with stressed syllable underlined.)

set	*target*	*transliteration/gloss*
1	01	Zahar *namaš* 9ala gism l-walad wa-9arifna ?innul HaSba *Spots* appeared on the boy's body and we knew it was measles.
	02	šufna *malik* il-?urdun lamma ruHna l-?urdun We saw the *king* of Jordan when we went to Jordan.
	03	al-?asmaa wa *nimar* it-tilifoon bititnisi bis-sur9a Names and telephone *numbers* are quickly forgotten.
	04	9ammitna *lili* bitit9allim yunaani Our aunt *Lily* is learning Greek.
	05	xaalitna *muna* hatsaafir faransa Our aunt *Muna* will travel to France.
	06	bu?? *muna* munimnim xaaliS wa ša9riha Tawiil Muna's mouth is very tiny and her hair is long.
2	07	?akalna *manga* laziiza giddan min-is-suu? We ate a very delicious *mango* from the market.
	08	warayna *manhag* il-kurs il-gidiid lil-mudarrissiin We showed the new *timetable* to the teachers.
	09	HaSalit 9ala *minHa* min is-sifaara 9alašaan turuuH tidrus fi ?amriika She obtained a *grant* from the embassy in order go study in America.
	10	iddiini *nimra* tilifoonik wa hatiSSil biik Give me your telephone *number* and I will phone you.
	11	šuft film *mumti9* 9ala t-taariix il-yunaani I saw a *nice* film about Greek history.
	12	miš *mumkin* id-duxuul taani ba9d il-xuruug It's not *possible* to come in again after you have left.
3	13	ir-ruzz da *maaliH* ?awwi wiTa9muh waaHiš That rice is very *salty* and it tastes bad.
	14	fii *maani9* kibiir bayni wa bayni-d-diraasa l-9ulya wa huwwa l-filuus There is a big *obstacle* between me & higher education & it's money.
	15	šufna *miina* buur sa9iid l-gidiid lamma ruHna buur sa9iid We saw the new Port Said *port* when we went to Port Said.
	16	in-nahr *niili* ba9d Guruub iš-šams The river looks *turquoise* after sunset.
	17	šufna *nuunu* SuGayyar ?awwi fil-mustašfa We saw a very small *baby* in the hospital.
	18	šufna *muulid* il-wali lamma ruHna l-Hussayn We saw the *festival* of the saint when we went to AlHussayn district.

The sentences were read 3 times each by 15 speakers of EA (9 male, 6 female). All were mother tongue speakers of CA, born and raised in Cairo, aged 21-34 years, and at pre-intermediate level or lower in English, with no reported auditory or speech production difficulties; subjects were paid a small fee for their participation. The recordings thus potentially yield 810 tokens (18 targets x 15 subjects x 3 repetitions).

Digital recordings were made in Cairo at 44.1KHz 16bit using ProTools 6.0 on MBox in a draped classroom. A head microphone was used to reduce the effects of inevitable background noise. Each sound file (re-sampled at 22.05KHz 16bit) was segmented and labelled with reference to F0 contour and spectrogram extracted using Praat 4.2 at default settings (Boersma & Weenink 2004). Tokens containing a disfluency (N=59) or phrase boundary (N=42) on or near the target word were discarded, leaving 709 tokens for analysis.

7.2.2 Data processing and analysis

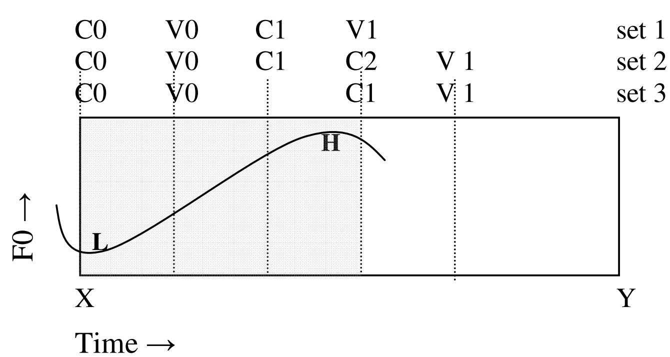
The following segmental landmarks and pitch events were labelled in each target syllable (see Figure 7.1 below):

(7.2)	C0	start of initial consonant of target syllable
	V0	start of vowel of target syllable
	C1	start of next consonant
	V1	start of following vowel
	Y	right edge of the word
	L	the start of the F0 rise (pre-peak F0 minimum)
	H	the end of the F0 rise (local F0 maximum)

The following dependent variables were extracted (in milliseconds):

(7.3)	L-C0	distance from L to onset of stressed syllable
	L-V0	distance from L to onset of stressed vowel
	H-C1	distance from H to end of stressed vowel
	H-V1	distance from H to onset of following vowel
	H-C0	peak delay (distance from H to onset of stressed syllable)

Figure 7.1 Schematised labelling diagram of segmental landmarks and pitch events



[malik]	m	a	l	i	k		set 1
[manga]	m	a	n	g	a		set 2
[ma:liH]	m	a:		l	i	H	set 3

Measures of stressed syllable and foot duration were calculated as follows:

(7.4) Definitions of syllable & foot duration used during analysis

		set 1	*set 2*	*set 3*
Syllable Duration #1 (canonical syllable)	*sylldur1*	C1-C0	C2-C0	C1-C0
Syllable Duration #2 ('ambisyllabic' syllable)	*sylldur2*	V1-C0	C2-C0	C1-C0
Foot Duration[174]	*footdur*	Y-C0	C2-C0	C1-C0
Word Duration[175]	*worddur*	Y-X	Y-X	Y-X

Note that two definitions of the 'syllable' are included in these calculations, both the canonical syllable (that is, CV, CVC and CVV across the three sets respectively), and an ambisyllabic conception of the syllable, which means inclusion of the foot internal intervocalic consonant in the sequences of light syllables in set 1 (thus, CVC, CVC and CVV).

In addition to the basic variables listed in (7.3) above, and in order to facilitate comparison of the EA facts with published data from other studies, another common measure of peak alignment was calculated. 'Relative peak delay' (RPD) is a measure of 'peak delay' as a proportion of the duration of the stressed syllable, and for the present purposes four versions of RPD were calculated: RPD#1 (peak delay/sylldur#1), RPD#2 (peak delay/sylldur#2), RPD#3 (peak delay/footdur), and RPD#4 (peak delay/worddur).

7.2.3 Experimental hypotheses

The dataset provide descriptive data regarding the surface alignment of EA pitch targets, which is used below in the first instance to determine whether alignment patterns differ in CV vs. CVC/CVV syllables (suggesting that the TBU is *not* the canonical syllable, but instead the foot) or are the same in all three syllable types (suggesting that the TBU is the canonical syllable). In addition the descriptive data are used to make comparison with the alignment facts of other languages.

Next the EA dataset is used to formally test each of the potentially explanatory association mechanisms against the facts of EA alignment.

---

[174] This definition of foot duration assumes incorporation of an extrametrical final consonant into the stress foot in targets 01-03.

[175] In six of the target sentences (03, 04, 09, 12, 13, & 14) the target word is preceded by an unaccented function word. In the analysis assumed here such function words procliticise to the following lexical word within a phrase-level constituent (MiP; see 6.2.4). These target words are assumed to be mapped alone to, and therefore co-extensive with, a PWd; thus word duration = PWd duration.

In order to test the main 'structural anchoring to TBU' hypothesis the degree of correlation between peak delay (H-C0) and the various potentially relevant structural domains (sylldur#1, sylldur#2, footdur, worddur). Measures of relative peak delay (peak delay as a proportion of each of the three domains), will be compared across syllable types, in order to determine whether the appropriate reference points for target alignment generalise consistently to a structurally defined domain.

In addition the predictions of alternative hypotheses are also tested. If the fixed duration hypothesis holds of the EA data, there should be no correlation between rise duration (H-L) and syllable duration (both sylldur#1 and sylldur#2 will be tested). If the fixed slope hypothesis holds there should be a correlation between the degree of F0 change (F0 at H - F0 at L, measured in semitones) and the time taken to complete the rise (rise duration) (cf. Arvaniti et al 1998). In order to test the segmental anchoring hypothesis the alignment of L and H to segmental landmarks will be compared across different syllable types, to determine whether the appropriate reference point for target alignment generalises to a point which is segmentally defined (such as 'end of the stressed vowel').

7.3 Results

7.3.1 Choice of variables for comparison

There is some variation in the existing literature as to the most appropriate measurements and derived variables for investigation of alignment.

Some authors use durational variables, which measure the distance between an F0 event and a segmental landmark, such as L-C0 and H-C1 etc (e.g. Atterer & Ladd 2004). However as Atterer & Ladd point out in an appendix, the choice of which landmark to measure against is not trivial. If an F0 event is relatively 'far' from a particular segmental landmark, changes in speech rate (both intra- and inter-speaker) will affect the degree of variance of that derived variable much more than if the event were relatively close to the landmark.

A common alternative measure of peak alignment used by other authors is 'peak delay', being the distance between the F0 peak ('H') and the syllable onset (with syllable onset defined most often as the start of the syllable-initial consonant, that is, C0) (for example, in Arvaniti et al 1998). Extending use of this measure further, still other authors calculate 'relative peak delay', which is a measure of 'peak delay' as a proportion of the duration of the stressed syllable. This measure is used by Prieto et al (1995) for Spanish, Chahal (2001) for Lebanese Arabic and also Yeou (2004) for Moroccan Arabic. Note that in their Appendix Atterer & Ladd (2004) suggest that a potential solution to the difficulty of identifying which segmental landmarks to measure against would be to measure instead the position of the F0 event as a percentage of some relevant domain, such as the stressed syllable, and this equates to the use of the various relative peak delay variables here.

In contrast, there is almost total consensus as to which variable to use to investigate the start of the F0 rise in rising pre-nuclear accents. Most authors use as a measure the distance between the start of the F0 rise (L) and the start of the syllable-initial consonant (C0). Some simply make the assumption that the start of the rise and the start of the syllable are the same point (as in for example, Arvaniti et al 1998, who base this assumption on the findings of Arvaniti & Ladd 1995). In order to facilitate comparison with all of these studies, all of these possible measurements and variables were extracted from the EA alignment data.

As a preliminary to full statistical analysis, an initial comparison was made of the frequency distributions of the various variables, since most statistical tests are based on the assumption that data are normally distributed. The mean, median and skewness of all of the potential variables, across all fluent tokens in the dataset, are reported in Appendix C.1. All but two of the variables are normally distributed, the exceptions being peak delay (H-C0) and RPD#3 (relative to foot duration). However, the distributions of RPD#1, #2 and #4 are not skewed, which suggests that in general relative peak delay is a more reliable measure than absolute peak delay. The proportional measure may alleviate the problem of speech rate variation between speakers, which would be a particular acute in a 'long-distance' measure such as peak delay (H-C0), by normalising the data against a relevant measure within the same token.

In the remainder of the chapter relative peak delay will be used as the primary measure of peak (H) alignment, and L-C0 as the primary measure of L alignment, with other variables (such as H-C1) reported and discussed when appropriate in order to facilitate comparison with the existing literature.

7.3.2 Pre-analysis: potential effects of factors external to the experimental design

A pre-analysis was carried out to assess the potential effects of three factors which were external to the design of the experiment but which could affect the results: gender, clash context and vowel quality.

Firstly, since both male and female subjects were recruited it is possible that gender has an effect on either F0 events (due to differences in pitch range), or on speech rate, or both.

Secondly, the experiment requires EA lexical items which have a specific number of syllables and stress position, and in which the stressed vowel is flanked by sonorant consonants [m], [n] or [l]. The number of such words was limited, and it was not possible to include words with only one vowel quality. An equal number of words with each of vowels [a], [i] and [u] were used in each set in order to balance out as much as possible any potential effects of vowel quality, since it is known that vowel quality affects intrinsic vowel duration (e.g. Chahal 2001 for Lebanese Arabic).

Finally, in an effort to elicit the colloquial register of the spoken dialect, rather than a more formal standard variety of Arabic, the sentences were kept as natural as possible, and it was not possible to fully control the position of stress in words preceding and following the target word. It is known that a following 'clash' can affect peak alignment (Prieto et al 1995), and there is some evidence to suggest that a preceding clash may also have an effect (Prieto 2005a). Counting the distance between stressed syllables, in the alignment dataset there was either 0 or 1 intervening syllables between the stressed syllable of the target word and the stressed syllable of the preceding word, and a range of 1-4 intervening syllables between the stressed syllable of the target word and the stressed syllable of the following word. Cross-linguistic clash resolution strategies vary, but are known to affect F0 alignment and/or duration, and it is thus important to check whether such effects are found in the present dataset. A series of one-way ANOVAs was carried out on the relevant measures to decide whether or not these potentially confounding factors should be taken into account in the main statistical investigation.

Starting with the effect of gender, a one-way ANOVA by gender shows that there is no significant effect on durational variables such as syllable duration (sylldur#2: p=0.657), but that gender does have a significant effect on F0 alignment variables (eg rpd#3:

p<0.01). This suggests that there are not major gender-related variations in speech rate, but that, as might be expected, differences between genders in pitch range, and possibly also pitch register, do have an effect on F0 alignment. These differences are retrievable from the planned by speaker analysis in the main statistical investigation so it was decide to proceed as planned but to look out for potential grouping of speakers by gender with regard to aspects of F0 alignment and scaling.

Turning next to potential effects of vowel quality, a one-way ANOVA of stressed vowel duration across all tokens, factored by vowel quality ([a] vs. [i] vs. [u]) shows that whilst stressed [i] vowels are on average slightly shorter than [a] and [u], the difference is not significant (p=0.053, α= 0.05). I assume that any effect of vowel quality on the alignment of F0 events would be *indirect*, mediated via a direct effect on stressed vowel duration. In the absence of a direct effect on stressed vowel duration in the dataset, it was deemed appropriate to set aside vowel quality as a potentially confounding factor in the overall alignment results.

The effects of clash in the dataset, as revealed by a series of one way ANOVAs of relevant variables, factored by preceding/following clash, are complex and must be set in the context of what is known about clash resolution in EA (see discussion in chapter 2 section 2.3.1). The type of clash effects that might be expected in the present dataset are as follows: in the case of a possible preceding clash, accent shift in the target could result in later alignment of either the L or H targets (or both), and/or reduction in the duration of the stressed syllable; a following clash might result in an increase in the duration of the stressed syllable of the target or earlier alignment of the H target due to accent shift.

A oneway ANOVA across all tokens by preceding clash interval (0 vs 1 intervening syllables) suggests that a preceding clash has no significant effect on the local alignment of the L target (L-C0: p= 0.375). There appears to be an effect of preceding clash on local alignment of the H target (H-syllend2: p< 0.01), but it is in the opposite direction than that which might be expected, since the H in fact aligns earlier after a clash (of 0 intervening syllables) instead of later as expected. There is an effect on stressed syllable durations, which are significantly shorter in strictly adjacent clash contexts (0 intervening syllables) than in non-adjacent clash contexts (1 intervening syllable) (sylldur#2: p<0.01), suggesting that there may be a rhythmic stress-shift effect in EA, sensitive to strictly adjacent clash only, but with no accompanying effect on F0 alignment. These mixed effects do not create an overall picture of systematic effects of preceding clash on F0 alignment to the segmental string.

Turning to the effects of a following clash, a oneway ANOVA by following clash (1 vs 2+ intervening syllables) shows a local effect on the alignment of the H target (H-C1), which falls on average slightly before C1 in clash contexts (1 intervening syllable) but after C1 in non-clash contexts (2+ intervening syllables) (H-C1: $p<0.001$). However, in other measures of H alignment (such as peak delay (H-C0) and H-syllend2) this effect disappears. There are no significant differences in stressed syllable duration between clash and non-clash contexts (1 vs 2+ intervening syllables). On the basis of this pre-analysis, together with that on preceding clash, it was decided to pursue the main statistical investigation without including clash as a factor.

To confirm the findings of these individual pre-analyses, a series of linear regressions were run including the above factors as potential predictors of Peak Delay (*PD*), alongside Speaker and Syllable duration[176]. The full model therefore comprised Syllable Duration (*sylldur*), Speaker (*spkr*), Vowel quality (*vowqual*), Preceding Interval (*preint*), Following Interval (*follint*) and Gender[177]. The regression analysis was repeated leaving one factor out of the model at a time, in order to determine which variables were predicting a significant percentage of the variance in the model.

The full model accounts for just over 36% of the variation in the model ($R^2 = 0.364$). Syllable duration is the major predictor, accounting for 22.59% of the variation. The next most important predictor is Speaker, which however accounts for only 3.09% of the variation in the model. A following clash was shown to have only a very small effect (Following Interval accounts for 1.26% of the variation) and the effect of gender, vowel quality and preceding interval are confirmed to be negligible. The results of the analysis for each version of the model is shown in the table in (7.5).

[176] Syllable duration was included as a likely meaningful predictor. The sylldur#2 definition was used in these regressions analyses, since it was found to be the best predictor of peak delay (as discussed below).
[177] $PD = \alpha + \beta_1$ sylldur $+ \beta_2$ spkr $+ \beta_3$ vowqual $+ \beta_4$ preint$+ \beta_5$ follint $+ \beta_6$ gender (where α is a constant).

(7.5) Multivariate linear regressions to determine the predictive value of potentially confounding factors.

model	*R squared*	*Constant*	*Sylldur*	*Speaker*	*VowQual*	*PreInt*	*FollInt*	*Gender*
all	0.364	109.80	0.58	-3.76	-0.34	-0.38	-7.70	2.16

without:								
Speaker	0.333	94.68	0.59		0.09	-0.02	-8.16	-26.02
VowQual	0.364	108.73	0.58	-3.76		-0.03	-7.77	2.14
PreInt	0.364	109.08	0.58	-3.76	-0.25		-7.69	2.15
FollInt	0.351	96.30	0.57	-3.85	-2.00	1.08		2.76
Gender	0.364	109.31	0.58	-3.56	-0.32	-0.37	-7.73	
Sylldur	0.138	245.19		-4.26	-8.53	-23.02	-4.41	6.49

As a result of these pre-analyses it was deemed appropriate to set aside all of these potentially confounding factors.

7.3.3 The descriptive facts of EA alignment

Mean values of the basic measures of L and H alignment (as in (7.3) above) provide a first picture of the alignment facts of EA. The results are listed in (7.6) below.

The details of individual speaker behaviour in L alignment (based on speaker means by set) reveals that there are two speakers who align L on average before C0 (*mrf* and *mun*, mostly in set 1), but that no one speaker consistently aligns L before C0 across sets; the overall tendency is to align L just after C0. There is similar variation in individual speaker behaviour in H alignment with speakers divided roughly equally in both CV syllables (just before or just after V1) and CVV syllables (just before or after C1), whilst in CVC syllables all speakers align H between C1 & C2.

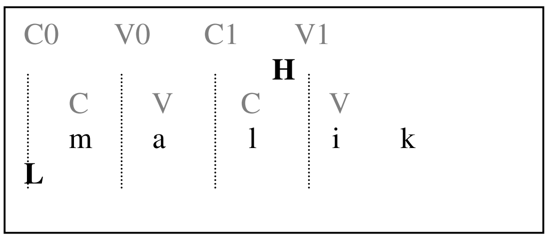
(7.6) Mean & standard deviation alignment variables in milliseconds by syllable type

Set #		L1-C0	L1-V0	H-C1	H-C2	H-V1	H-C0
Set 1	Mean	7.21	-58.24	43.45		-5.31	177.93
CV	N	215	215	215		215	215
	SD	26.33	27.47	37.36		36.04	48.37
Set 2	Mean	16.24	-44.85	25.89	-45.04	-114.49	155.48
CVC	N	249	249	249	249	249	249
	SD	29.33	31.68	29.08	40.02	46.66	36.63
Set 3	Mean	10.41	-57.91	3.57		-48.23	177.36
CVV	N	245	245	245		245	245
	SD	33.61	34.29	46.54		46.74	57.35
Total	Mean	11.49	-53.42	23.51	-45.04	-58.49	169.85
	N	709	709	709	249	709	709
	SD	30.23	32.01	41.53	40.02	62.49	49.22

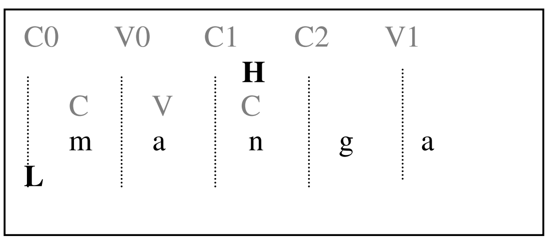
The descriptive generalisations for EA are thus that L alignment is stable at a point just after the onset of the stressed syllable whereas H alignment appears to vary between syllable types (light vs. heavy): in a CV syllable H aligns inside the foot internal intervocalic consonant (after C1 and just before V1); in a CVC syllable H aligns on average a third of the way through the coda consonant (between C1 and C2); and in a CVV syllable H aligns at the end of the stressed long vowel (just before or just after C1). These generalisations are illustrated in (7.7a-c) below.

(7.7) Schematic summary of L and H alignment across syllable types in EA

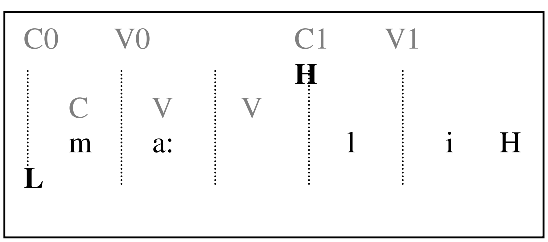
a.



b.



c.



The surface alignment facts in EA seem to pattern closely with those observed by Ladd et al (2000) in Dutch. The L target aligns stably at the left edge of the syllable, whilst the position of the H target appears to vary according to syllable type. Specifically in open stressed syllables the H peak falls well outside the CV syllable, inside the ‘ambisyllabic’ foot internal intervocalic consonant, whilst in both CVC and CVV syllables the H peak falls just inside the stressed syllable (within the coda in CVC and just before the end of the stressed vowel in CVV).

Recall the hypotheses regarding the potential TBU in EA, set up in section 7.1.1 above and repeated here:

i) if the syllable is the TBU in EA, then the patterns of alignment observed in CV stressed syllables (which are initial in a CVCV bisyllabic foot) will be different from those observed in CVC and CVV syllables (in a monosyllabic foot);
ii) if the foot is the TBU then patterns of alignment are not expected to vary significantly across the three stressed syllable types.

The descriptive generalisations appear to support the notion that the foot is the TBU in EA. Recall however that it is not possible to compare across phonological categories in EA which are also phonetically identical in duration, as was possible for Ladd et al (2000) in Dutch (cf. section 7.1.1 above). Hence it is important to formally exclude other hypotheses such as the fixed duration hypothesis, as is done in the next section.

7.3.4 Hypothesis testing

7.3.4.1 The secondary hypotheses: fixed duration/slope and segmental anchoring

To test the fixed duration hypothesis the degree of correlation between rise duration and syllable duration (both sylldur#1 and sylldur#2) was tested. If there is co-variation in rise length and syllable length this undermines the notion that the F0 rise is purely durationally defined. A scatter plot (Figure 7.2 below) shows that rise duration and syllable duration are highly correlated, and the correlation is highly significant whichever definition of the syllable is used (Pearsons' correlation coefficient: risedur x sylldur#1= 0.261; risedur x sylld#2 =0.406; $p< 0.01$ in both cases). It appears that rise duration is highly sensitive to syllable duration, and thus that the fixed duration hypothesis does not hold of EA.

Similarly, to test the fixed slope hypothesis the degree of correlation between the rise duration and the degree of F0 change (F0 at H - F0 at L) was tested. A scatter plot in Figure 7.3 below shows that F0 change (in semitones) and rise duration are highly correlated. A non-parametric measure of correlation was used which indicates that there is significant correlation between Rise Duration and F0 Change. The fixed slope hypothesis therefore cannot be excluded as an explanation of the positioning of H and L targets in EA rising pitch accents. A summary of statistical results is provided in (7.8).

(7.8) Testing the fixed slope hypothesis

Tests of normality (Kolmogorov-Smirnov)	risedur: $p< 0.05$ in sets 1 & 3; F0 change: $p< 0.05$
Spearman's Rank	$R= 0.514$, $p<0.01$ 2-tailed
Kendall's Tau	$\tau = 0.357$, $p < 0.01$ 2-tailed

Moving on to the segmental anchoring hypothesis, since the alignment of L is stable, here the alignment of H to segmental landmarks was compared across different syllable types. If patterns of H alignment can be generalised to a particular segmental landmark this would constitute evidence in support of the segmental anchoring hypothesis. Already from the descriptive results (as in (7.7) above) it appears highly unlikely that any single segmental landmark can be used to define H alignment in EA; instead, different segmental landmarks appear to be relevant to different syllable types (CV: C1/V1; CVC: C1/C2; CVV: C1). The most likely candidate as a consistent segmental landmark seems to be C1 (the end of the stressed vowel). However, as shown in Figure 7.4, values of H-C1 vary significantly between syllable types, and a one-way ANOVA by set confirms that the differences are significant ($F= 62.571$; $p<0.01$).

As an aside, recall that in Dutch it was possible to prove conclusively that the difference in alignment between long and short vowels was phonologically conditioned by comparing alignment in high vowels which were phonologically distinct (long/short) but phonetically almost identical in duration. There is no indication that EA high vowels are phonetically short, but a measure was taken of the duration of the stressed vowel across syllables types (by calculating the distance between C1 and V0 in all syllables types). The results were classified by vowel quality (recall that the materials included both long and short exemplars of [a], [i] and [u].

Figure 7.5 below shows mean values of Stressed Vowel Duration classified by syllable type and by vowel quality. It is clear that vowels of all types are significantly longer in set 3, CVV syllables, than in sets 1 & 2, CV and CVC syllables (significant at $p< 0.05$ level for the difference betweens set 3 and Sets 1&2 using Tamhane's test (variances are not equal)). It is unlikely to be possible to tease apart the structural and durational hypotheses in EA, since phonologically long and short vowels of all quality differ in absolute duration.

Figure 7.2 Scatter plot of rise duration x syllable duration.

The plot shows correlation between the two variables, which is *incompatible* with the 'fixed duration hypothesis'.

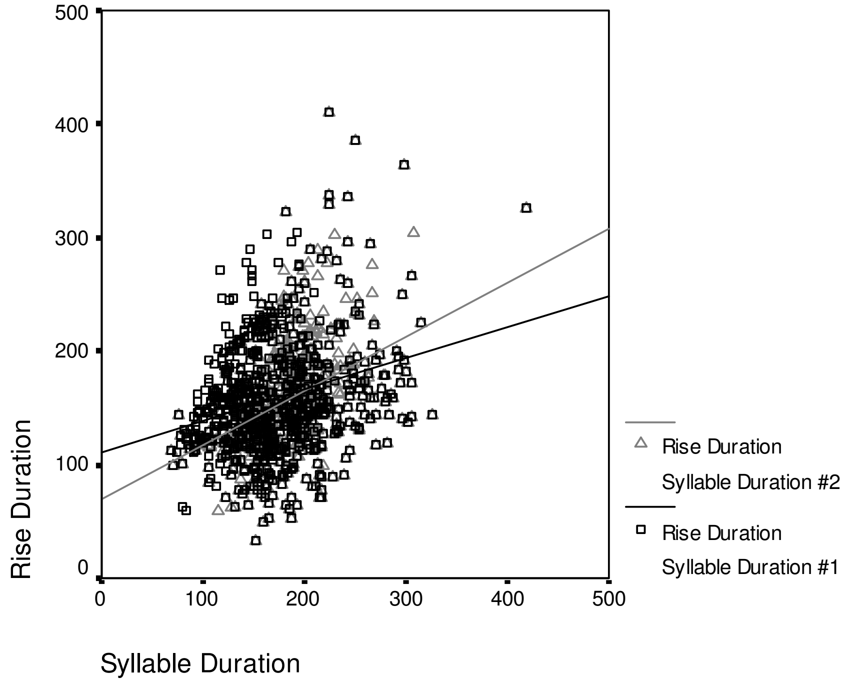


Figure 7.3 Scatter plot: F0 Change (semitones) x Rise Duration (by syllable type).

The plot shows correlation between the two variables in all three syllable types, which is *compatible* with the 'fixed slope hypothesis'.
Set 1 = CV; set 2 = CVC; set 3 = CVV.

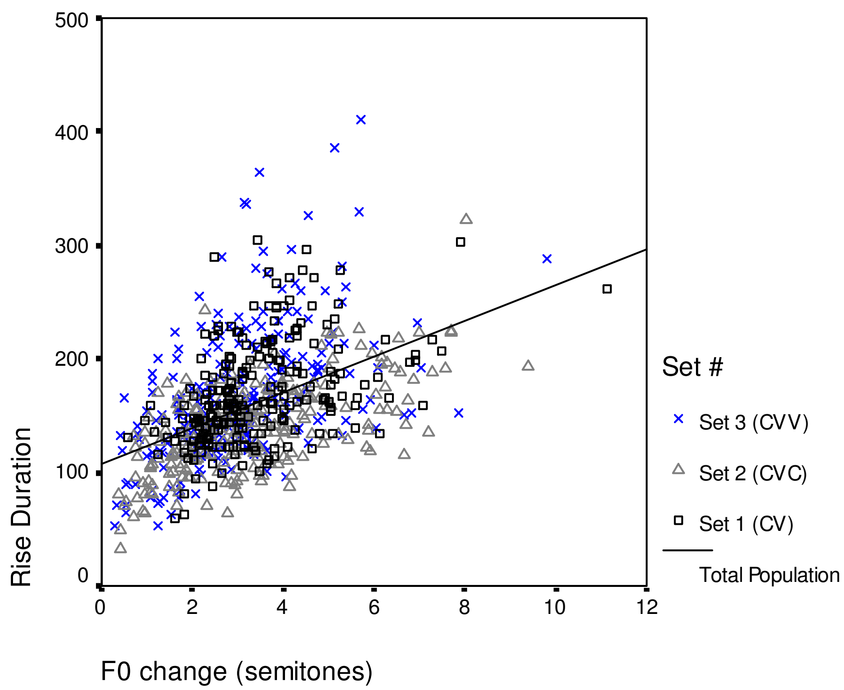


Figure 7.4 Confidence Intervals for H-C1 by syllable type.

The plot shows mean values of H-C1 (the distance between the H peak and the end of the stressed vowel). C1 is the most likely candidate to act as a uniform target for segmental anchoring of the H peak in all syllable types. The plot indicates however that the values of H-C1 vary significantly between syllable types, which is *incompatible* with the 'segmental anchoring hypothesis'.

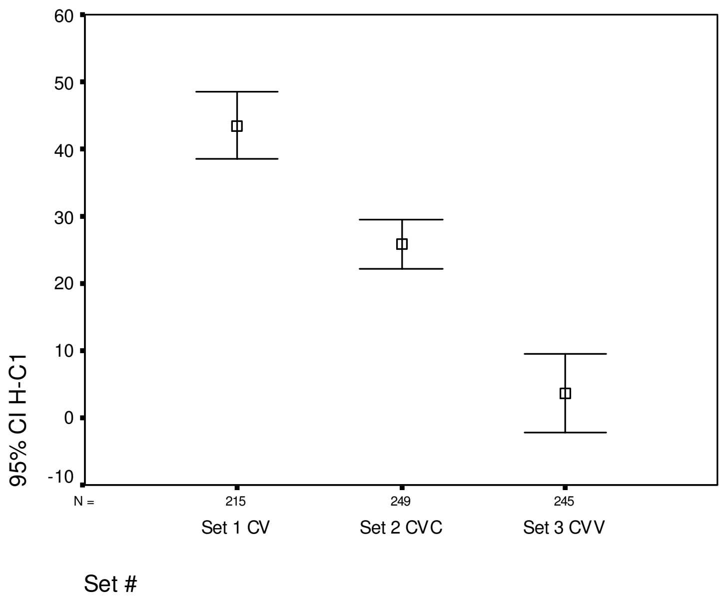
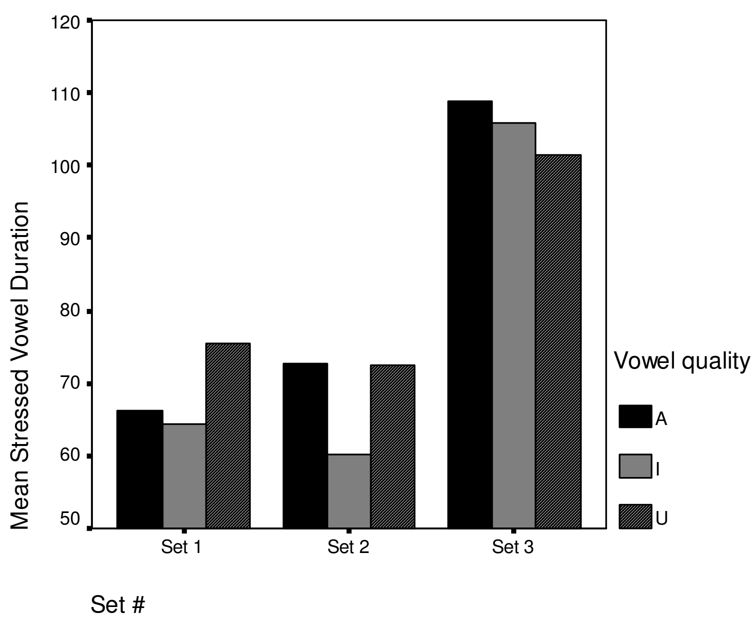


Figure 7.5 Mean Stressed Vowel Duration by syllable type and by vowel quality.

The plot shows mean values of stressed vowel duration, grouped by syllable type (Set 1 = CV; set 2 = CVC; set 3 = CVV). Black bars indicate vowel [a]; light grey bars indicate vowel [i]; dark grey bars indicate vowel [u].



7.3.4.2 The primary hypothesis: structural anchoring to the Foot as TBU

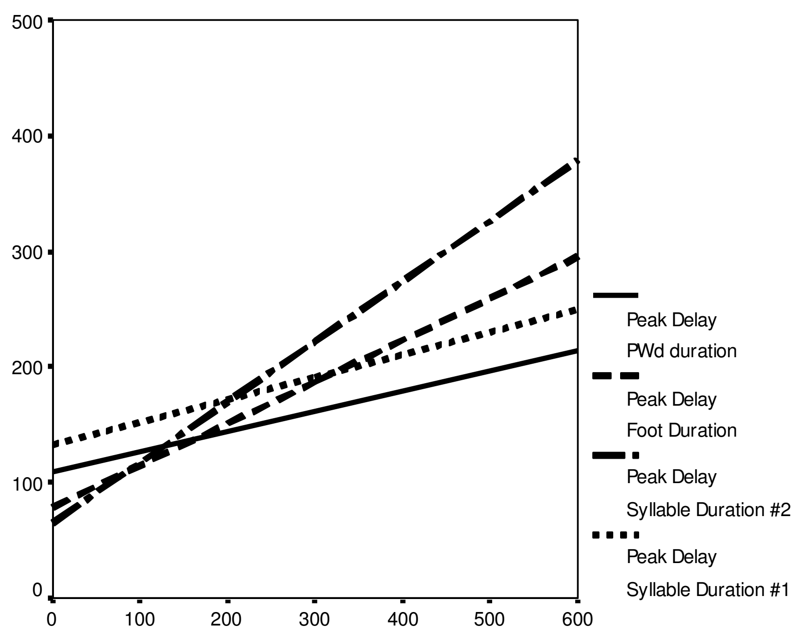
The 'structural anchoring to TBU' hypothesis predicts a correlation between peak delay (H-C0) and some prosodic domain. Four potential domains are tested here: the canonical syllable, the 'ambisyllabic' syllable (whose status as a prosodic constituent is discussed below), the foot and the PWd. Best-fit regression lines shown in Figure 7.6 below show that the durations of all four potentially explanatory domains are highly correlated to peak delay, and all of these correlations prove to be significant. Note that the highest correlation is between peak delay and sylldur#2 (the 'ambisyllabic' syllable), suggesting that the 'ambisyllable' is the domain most likely to be relevant to structural anchoring. Individual scatter plots of each of the four correlations are provided in Appendix C.2. A summary of the statistical results is provided in (7.9).

(7.9) Testing the structural anchoring hypothesis

	Spearman's rho[178]	
peak delay x sylldur#1	0.346	$p < 0.01$
peak delay x sylldur#2	0.516	$p < 0.01$
peak delay x footdur	0.430	$p < 0.01$
peak delay x wordur	0.253	$p < 0.01$

Figure 7.6 Regression lines: peak delay x four different structural domain durations.

The lines shows correlation between the two variables, in all four measures of structural domain, which is *compatible* with the 'structural anchoring hypothesis'. Sylldur#1 = 'canonical syllable; Sylldur#2 = ambisyllabic syllable.



[178] The distribution of peak delay values is slightly positively skewed (skewness = 1.163) hence a non-parametric test was used.

Finally, a series of linear regression analyses were carried out with Speaker and Following Interval included as predictive factors of peak delay (PD) in the model[179], alongside each of the four potentially explanatory domains: the canonical syllable (*sylldur#1*), the ambisyllabic syllable (*sylldur#2*), the foot (*footdur*) and the PWd (*pwddur*)[180]. The results of each of the regressions are shown in the table in (7.10) below.

The results indicate that the domain which best predicts the variation in the model is the ambisyllabic syllable (36.4%), followed closely by the foot (30.7%)[181].

(7.10) Multivariate linear regressions to detect predictive value of confounding factors

domain tested:	*R squared*	*Constant*	*Speaker*	*FollInt*	*'Domain'*
Syllable Duration #1	0.2399	142.9826	-3.63091	-3.45976	0.372536
Syllable Duration #2	0.36363	108.2521	-3.55357	-7.78668	0.582462
Foot Duration	0.307187	150.5585	-3.57786	-9.89774	0.327436
PWd duration	0.244621	146.6789	-4.03273	-3.59093	0.18851

7.3.4.2 Summary: hypothesis testing

In summary then, the alignment data suggest that it is appropriate to reject both the fixed duration hypothesis and a (purely-)segmental anchoring hypothesis for EA.

There is evidence however to support both the fixed slope hypothesis and the 'structural anchoring to TBU' hypothesis. In the latter case, in apparent contradiction to the hypothesis that the TBU in EA is the foot, the evidence suggests that the 'ambisyllabic syllable' is the structural domain relevant to pitch accent alignment in EA.

7.3.5 Explaining the apparent fixed slope effects

From the data analysis two mechanisms appeared to be at work in EA: there is evidence in support of both the fixed slope hypothesis and the structural anchoring hypothesis. Although fixed slope effects have also been noted in the EA pronunciation of MSA (El Zarka p.c.), I will argue here that attempting to realise the F0 rise with a fixed slope is

[179] These were found to be factors responsible for a small amount of variation in the model during pre-analysis (see 7.3.2 above).

[180] *Peak Delay* = $\alpha + \beta_1$ speaker + β_2 follint + β_3 'domain' (where α is a constant).

[181] In the discussion section below I will argue for reanalysis of the ambisyllable as the foot; hence the distinction here is between alignment to a structural domain ending at the second mora of the foot (the ambisyllable) or at the extreme edge of the foot itself (the end of the second mora).

not the primary mechanism driving pitch accent in EA. Instead I suggest that the observed correlation between F0 change and rise duration is best understood as a by-product of target undershoot under time pressure.

Languages appear to vary in the realisation of multi-tonal sequences under time pressure, giving priority to accurate realisation of either the alignment or the scaling of pitch targets, but not both. For example, Ladd et al (1999) show that in English rising pre-nuclear accents display constant alignment to the segmental string across speech rates, whereas Prieto & Torreira (2004) found that alignment in Spanish was not constant across speech rates.

There is obviously a high degree of adaptation of the tonal contour to speech rate in EA, as evidenced by the high correlation between syllable duration and rise duration. I suggest that this is achieved in EA at the expense of full realisation of the F0 scaling of one or more of the tonal targets. In syllables of relatively shorter absolute duration, the positional alignment of the pitch targets will be maintained at some cost to their F0 height specification. Recall that the corpus survey in chapter 3 noted instances where the L valley between adjacent pitch accents (with a single intervening unstressed syllable) was realised at a much higher frequency than the pitch register of the utterance would predict, and that these were analysed as cases of target undershoot due to 'tonal repulsion' from the close proximity of pitch accents[182]. Pressure of this sort would result in a measure of correlation between the degree of F0 change and the duration of the rise: less time between tones results in decreased F0 excursion (realised by means of undershoot of the L tone).

If we adopt this explanation for the apparent fixed slope effects in EA, then the remainder of the evidence points clearly towards the structural anchoring to TBU hypothesis, and I propose that this is the primary mechanism underlying EA surface alignment patterns.

7.3.6 Comparison with pitch accent alignment in other languages

Descriptive comparison with the results of studies on other dialects of Arabic, using parallel measures of alignment extracted from the EA data, and compared in parallel syllable types to those used in the published studies, reveals variation in the basic

[182] See chapter 3 section 3.4.2.2, with an example of target undershoot provided in Figure 3.10.

alignment patterns of pre-nuclear rising accents in different Arabic dialects and in different languages. The data are compared as much as possible across parallel syllables types and position in the word[183].

In CVC syllables in Lebanese Arabic (LA) L aligns before/after C0 (depending on word position), and H aligns outside the stressed syllable (Chahal 2001). This compares with alignment in CVC syllables in EA CVC syllables of L just after C0 but of H inside the stressed syllable. In CV and CVC syllables in Moroccan Arabic (MA) L aligns “close to the onset of the syllable” whilst H aligns outside the (canonical) syllable[184] in CV syllables and inside the syllable in CVC syllables (Yeou 2004).

In CVC syllables then, EA and MA pattern together in that H is aligned inside the syllable, but differ from LA in which H aligns outside the syllable. All three of the dialects seem to align L just after the onset of the stressed syllable. These results are summarised in the table in (7.11) below.

(7.11) Comparison of alignment variables[185]: LA vs EA and MA vs EA

a.	*LA (range)*		*EA (mean)*	
	L (L1-C0)	*H (RPD)*	*L (L1-C0)*	*H (RPD#1)*
initial CVC syllables	21...49 ms	1.16...1.27	16.24 ms	0.79
medial CVC syllables	-50...-39ms	1.24...1.36		

b.	*MA (mean RPD)*	*EA (mean RPD#1)*
CV syllables	1.165	1.337
CVC syllables	0.897	0.789

Turning to comparison of the EA data with those of other languages, here comparison is made in CV syllables only (as used in all of the studies listed in the table in 7.12 below). EA seems to pattern with English and Dutch in aligning L nearer the onset of the stressed syllable than German. As regards H alignment, EA again patterns with English, aligning the peak after the end of the stressed vowel (within the ambisyllabic

[183] Recall that in variables expressing the distance of a pitch event from a segmental landmark (such as L-C0) a negative value indicates alignment of the pitch event *before* the segmental landmark. In measures of relative peak delay (RPD) a value < 1 indicates alignment of the peak within the stressed syllable (however defined) and a value > 1 indicates alignment of the peak outside the stressed syllable.
[184] Yeou does not specify which segmental landmarks form the edges of the syllable in his calculations; his measure of RPD is thus assumed to be peak delay as a proportion of the *canonical* syllable.
[185] Lebanese Arabic data from Chahal (2001), based on data from 4 speakers; Moroccan Arabic from Yeou (2004), based on data from 5 speakers.

consonant); both EA and English differ from Dutch which aligns H later, and from German which aligns H later still.

(7.12) Comparison of alignment variables[186] (CV syllables)

language	*# speakers*	*L-C0 (mean)*	*H-C1 (mean)*	*H-V1 (mean)*
EA	15	7.21	43.45	-5.3
Southern German	7	60.2		26.1
Northern German	7	49.9		10.7
English (normal rate)	6	-5.5		-3.6
Dutch (short vowel)	5	-1.2	24.8	

As mentioned already, there is good evidence for phonological conditioning of pitch accent association in Dutch (Ladd et al 2000), in which the descriptive alignment facts appear to be broadly similar to those in EA: stable alignment of L at the left edge of the syllable and phonologically conditioned alignment of H in the second mora of the (ambisyllabic) syllable (reanalysed here as the second mora of the foot). Nonetheless the fine-grained detail of those facts seem to be different in the two languages, with the whole pitch movement aligned earlier in Dutch than in EA.

If it is true that the same mechanism explains alignment in both Dutch and English (given the broadly similar descriptive facts), then, whilst cross-linguistic comparisons of these kinds are necessarily limited in scope (since the various studies use different materials, sample sizes and alignment measures), the EA data provide additional evidence to support the following claim: that whilst there may be a single model of underlying phonological association of pitch targets, such as that proposed in section 7.4 below, there is also likely to be a continuum of cross-linguistic variation in the surface phonetic detail of pitch target alignment (Atterer & Ladd 2004, Ladd 2004).

7.4 Discussion

7.4.1 Pitch accent alignment patterns across different syllable types in EA

The structural domain which is most closely correlated to peak delay in the EA data appears to be the ‘ambisyllabic’ syllable. This matches the results obtained for Dutch by Ladd et al (2000). Similarly, whilst Prieto & Torreira (2004) found no evidence to support anchoring of H to the syllable edge in Spanish, their calculations were based on a canonical definition of the syllable; their finding that H aligned systematically later in

[186] The studies are: S & N German (Atterer & Ladd 2004); English (Ladd et al 1999); Dutch (Ladd et al 2000).

CVC syllables than in CV syllables, if couched in an ‘ambisyllabic’ notion of CV syllables, suggests that ‘structural anchoring’ may hold in Spanish, and that the relevant domain is also an ‘ambisyllabic’ syllable.

There is however no such prosodic constituent as the ‘ambisyllable’. Many authors have made appeal to the notion of ambisyllabicity in the analysis of a wide variety of phonological phenomena, such as stop allophony and r-intrusion in English (McCarthy 1993), fricative allophony in German (Hall 1989), and consonantal assimilation and stød association in Danish (Borowsky et al 1984). Recently however, these analyses have been challenged in favour of a view whereby the context of these alternations generalises to one simple, prosodically defined, environment: they occur foot-internally. This view has been proposed for phonological phenomena by Jensen (2000), and for both phonetic and phonological phenomena by Harris (2004).

I propose therefore that alignment of H peaks inside the foot internal intervocalic consonant in CVCV feet is telling us that the syllable is not the TBU in EA at all: rather, alignment of the H target is foot internal. The H target does not align to the edge of the foot, but targets the second mora, hence the greater correlation between peak delay and ‘ambisyllabic syllable duration’ than between peak delay and foot duration. I suggest that this is in line with the null hypothesis put forward at the beginning of this chapter that the TBU in EA is the *foot* rather than the syllable. Note however that the head of the foot is leftmost in EA (in a bimoraic trochee), whereas the second mora is the weak mora of the foot, even though it attracts the perceptually salient H tone. The implications of this anomaly for designation of a phonological representation for the EA pitch accent are discussed in detail in section 7.4.2 below. In section 7.4.3 I explore implementation of the foot as TBU by means of T→P constraints.

7.4.2 Discussion: towards a phonological representation of the EA pitch accent

The results of the alignment experiment described in section 7.3 above were argued in section 7.4.1 to support proposal of the foot as the TBU in EA. This section explores how this finding might be implemented in the form of a designation for the phonological representation of the EA rising pitch accent. Specifically, three potential representations are examined, and the advantages and disadvantages of each analysis explored, as follows:

(7.13) Potential phonological representations for the EA pre-nuclear pitch accent.

i) (L+H)* ii) L*+H iii) L+H*

A compound (L+H)* representation was proposed by Hualde (2002) for Spanish, adopting the notion of a 'cluster' pitch accent in which both tones are associated to segmental landmarks at either edge of the stressed syllable (Grice 1995a). Face (2002) rejected this analysis for Spanish on the grounds that it doesn't answer to the asymmetric alignment properties of the L and H targets in Spanish: the stability of L (at the left edge of the syllable) compared to the relative instability of H. This same asymmetry is exactly what we find in EA pitch accents and thus would be good cause to reject a compound (L+H)* analysis of the EA pitch accent.

Face (2002) argued that the asymmetry of L and H alignment in Spanish is best captured by the representation L*+H. He chooses to assign the '*', and thus the primary role in 'establishing association' (Pierrehumbert & Beckman 1988), to the L tone, due to its stability of alignment. An advantage of using this analysis for EA would be the fact that the 'strong' tone status is assigned to the L tone, which is aligned most closely to the (leftmost) head mora of the foot in EA. Conversely however, this involves assigning 'strong' tone status to the least perceptually salient of the two tones (L and H), and which appears to be prone to undershoot in tonally crowded environments (see 3.4.2 and 7.3.5 above).

Prieto et al (2005) offer an analysis of the rising pre-nuclear pitch accent in Catalan, which shares the same asymmetry in the alignment of L and H as observed in both EA and Spanish. They propose that the correct representation is L+H*. As mentioned already in section 7.1.1 above, they assign 'strong' status within a bitonal accent not in the light of alignment properties but instead of which of the tones is perceived by native listeners as the more salient: is the rise perceived as 'a high tone' or 'a low tone'. The less salient tone is assigned the subordinate leading/trailing tone status (a leading tone in this case). They suggest that the default alignment of the pitch target will be to the edges of the TBU, but with primary association established by the starred tone.

The most salient tone in the EA rising pitch accent is the H tone. This is evidenced by the fact that both Mitchell (1993) and Norlin (1989) describe the pre-nuclear accents as 'peaks', and in Rifaat's (2004) AM analysis of the EA pronunciation of MSA he analyses pitch accents in non-phrase-final position as H*. In addition, H scaling in EA pitch accents appears to be resilient, whereas the L target is observed to undergo undershoot, and it is plausible to assume that the tone whose F0 scaling is preserved is the one whose phonetic properties are phonologically meaningful or contrastive, and that this is the 'starred tone'.

At first sight, a possible disadvantage of adopting L+H* for EA is the fact that if the H tone is the 'strong' tone it is hard to understand why it aligns within the rightmost, weak mora of the foot. However, if in fact the H tone is phonologically *associated* with the leftmost head mora of the foot, then the type of production restrictions discussed by Xu & Liu (2005) mean that we would expect it to consistently align just after the strong mora - that is, early in the second mora of the foot, as observed.

On the balance of the available evidence then, I therefore propose a L+H* representation for the ubiquitous EA rising pre-nuclear pitch accent. Under this analysis the positioning of the H tone is best defined relevant to the whole foot as a TBU domain. This implication is explored in the next section in the context of the theory of tone↔prominence relations.

7.4.3 Intonational TBUs and tone~prominence relations

In this section I argue that in EA association of the pitch accent to the foot as TBU arises due to a constraint requiring any tones in output representation to be associated to the foot in its role as the *head of the PWd*. This is based on the notion of the TBU as the "prosodic head of some level of prosodic constituency" (Yip 2002:141). In the T↔P theory outlined in chapter 6, association of the pitch accent to the foot as TBU is due to ranking of the constraint, T→PWD ('every tone must associate to the head of a PWd')[187] higher than DEP_{TONE} ('don't insert tone').

I suggest also that the notion of the foot as TBU in its role as the head of the PWd is consistent with the alignment facts, whereby the position of the H peak in EA is best defined relative to the foot domain as a whole, rather than in terms of any particular

[187] Recall that 'T→PWD' is shorthand for 'T→μ(PWd)', so that the tone will be attracted to the head or DTE of the PWd.

position within the foot. Thus the H aligns at some proportion of the foot which coincides most often with the second mora[188].

What sort of arguments are there in favour of formalising association of pitch accents in intonational languages by means of T→P constraints in this way? A more economical analysis would perhaps be to simply state that the T→P hierarchy only refers to lexical tones, and that in intonational languages pitch accents associate to the stressed syllable (perhaps in its role as head of the foot) as they have long been assumed to do. There are two central reasons why the influence of T→P constraints should not be restricted only to lexical tones. Firstly, as markedness constraints, which 'see' only output representation, some additional analytic mechanism would have to be introduced in order to effect the restriction. Whilst this is possible to do, it is undesirable for a second reason, which is that in AM theory it is in fact an advantage if the theory does encode the notion of the 'unity of pitch phonology' (Ladd 1996:147ff.). It is this aspect of the theory which means it can be used to analyse any language, regardless of the *function* of pitch in the language.

Decisive evidence for a role for the full T→P hierarchy in intonational languages would come from variation in the TBU across levels of the prosodic hierarchy in them. Yet from the limited survey of languages touched on here, and based on alignment evidence only, there appears to be no indication of such variation. In many of the languages mentioned, although usually analysed as being associated to the stressed syllables, pitch accents display properties suggesting that the correct conception of the TBU is the foot. For example, in Dutch we have evidence from alignment that the H tone falls outside of the stressed syllable, but inside the foot in a CVCV bimoraic trochee (Ladd et al 2000).

So what we have are indications that in EA and Dutch the intonational TBU is the foot, and thus that the active T→P constraint in Dutch would also be T→PWd. Indeed the other studies which compared pitch accent alignment across different syllable types, in Japanese and Spanish, also describe surface alignment patterns that are consistent with the analysis proposed here for EA: the foot is TBU, in its role as head of the PWd (due to T→PWd). Thus to date there is no evidence of variation in intonational TBU across levels of the prosodic hierarchy.

[188] Cf. definition of F0 peak alignment in English as a function of the distance between accented syllables (Silverman & Pierrehumbert 1990:99ff.).

Nonetheless the fact that these languages do not vary in TBU still provides an argument in favour of the T↔P view of prosodic variation. Specifically, the question arises as to how it is possible for Dutch and EA to share the property of having the foot as TBU but not to share the property of rich pitch accent distribution[189]. Regardless of whether or not the proposal that T→P constraints hold of intonational languages proves to be correct, the alignment facts of for EA, and the parallels with those observed in Dutch, provides compelling evidence in favour of the P→T analysis of EA pitch accent distribution.

The property of having rich or sparse pitch accent distribution is thus *fully independent* of other intonational properties, such as the association and alignment of individual pitch accents. This supports the notion that density of pitch accent distribution is an independent parameter of prosodic variation, as proposed in this thesis, and as captured in the formalisms of Tone-Prominence Theory. With a view to finding additional evidence of this kind, the next chapter explores another environment (namely, the reflexes of focus) in which pitch accent distribution might be expected, and usually is assumed, to reflect some other aspect of the grammar, rather than displaying independence.

7.5 Summary

This chapter investigated the surface phonetic alignment of pitch targets in EA rising pre-nuclear pitch accents, on the basis of newly collected experimental data. The investigation establishes the descriptive facts of EA pitch accent alignment, which are that the L tone aligns stably to the left edge of the stressed syllable, whilst the H tone falls somewhere in the second mora of the foot.

In addition several hypothesis were tested against the EA data as candidates for the phonological mechanism underlying surface alignment facts: fixed duration, fixed slope, segmental anchoring and structural anchoring. The structural anchoring arises directly from the tone↔prominence relations analysis adopted in this thesis. The hypothesis was that EA pitch accents associate to the foot. Although there was some indication of 'fixed slope' effects, these were ascribed to tonal undershoot of pitch targets in tonally

[189] Dutch has relatively sparse pitch accent distribution, similar to that observed in English (Gussenhoven 1983).

crowded environments, and the majority of the evidence supports the hypothesis. Association of the pitch accent to the foot as TBU, is analysed as due to ranking of the constraint T→PWD above DEP_{TONE}.

Possible phonological representations of the EA pre-nuclear pitch accent were discussed and arguments set out in favour of its designation as L+H*. The chapter closed by discussing how association of tone to the foot in EA might be accounted for by means of a T→P constraint. Whether or not this particular conception of postlexical tonal association proves to be correct, the alignment facts in EA are argued to provide definite support for the P→T analysis of EA pitch accent distribution proposed in chapter 6: density of pitch accent distribution is *fully independent* of other intonational properties and is an independent parameter of prosodic variation, as tone↔prominence theory predicts. The next chapter explores the reflexes of focus which is another area in which pitch accent distribution might be expected to reflect some other aspect of the grammar, rather than displaying independence.

8 Prosodic reflexes of focus in Egyptian Arabic

8.0 Outline and aims

The distribution of pitch accents has been analysed in many languages as essentially focus-marking, with the presence of pitch accents marking a distinction between new and old information: in English, old, ‘given’ information is usually ‘de-accented’ (Gussenhoven 1983, Selkirk 1984, Ladd 1996).

This chapter has two goals: firstly to confirm whether or not pitch accents occur in EA in the contexts where in Germanic languages pitch accents are deleted (or in some accounts drastically reduced in pitch range cf. Xu & Xu 2005). Norlin (1989) observed that after a focus pitch accents in EA were produced in compressed pitch range, but does not define the exact type of focus elicited in his study. There is syntactic evidence to support two semantically distinct notions of focus in Arabic (Moutouakil 1989): *information* focus and *contrastive* focus are obligatorily expressed using different syntactic strategies in Modern Standard Arabic (MSA). The second goal of this chapter is therefore to reproduce Norlin’s result in a freshly collected experimental dataset and if successful to clarify what type(s) of focus induce pitch range compression on ubiquitous EA pitch accents.

As a whole then, this chapter permits us to explore in what ways focus is expressed in a language like EA in which every content word bears a pitch accent. The claim of this thesis is that rich pitch accent distribution in EA can be ascribed entirely to the workings of the phonological part of the grammar, and specifically, in Tone↔Prominence Theory, to the effects of a constraint requiring every PWd-level prominence to bear tone: PWD→T. A valid test of this claim is therefore to demonstrate whether or not there is any relationship between pitch accent distribution in EA and the syntactic and semantic parts of the grammar.

Section 8.1 sets out in detail the reasons for choosing to distinguish these two types of focus in the present study are set out in section 8.1, together with a review of what is known from prior work about the syntax and the prosody of focus in EA. Section 8.2 describes the design and implementation of the focus experiment.

Qualitative analysis of the focus data was reported briefly in chapter 3 and showed that even when post-focal and 'given', EA words bear a pitch accent. These results are reviewed again here at the start of section 8.3, which then goes on to present the results of a quantitative analysis of the focus dataset. This reveals that there are gradient effects of focus in the form of pitch range manipulation and that the effects appear to reflect contrastive focus status, but not information focus status.

In a number of languages, such as Spanish and varieties of Italian, which share with EA the property of having a pitch accent on every content word, the distinction between 'broad' and 'narrow' focus has been shown to be expressed by means of distinct pitch accent types. These pitch accents differ largely in the alignment of the H peak, relative to the segmental string. Section 8.3 also reports the results of an investigation into the alignment of the pitch accents in the focus dataset, in order to determine whether or not distinct pitch accent types are used in EA to mark different degrees of focus.

In section 8.4, the use of pitch range manipulation to mark contrastive focus in EA is discussed in the light of debate in the literature regarding emphatic pitch range variation in other languages, and in particular as to whether or not focus marking of this kind is categorical or gradient. The chapter concludes in section 8.5 by discussing the apparent lack of marking of information focus in EA in the context of the syntactic properties of the language, and the implications of this finding for the purely phonological analysis of EA pitch accent distribution proposed in this thesis.

8.1 Focus in EA and other languages

8.1.1 Types of focus

The notion of focus can be used to describe what is in fact a range of different degrees of emphasis or highlighting. A number of different ways of dividing up this continuum have been proposed, which can usefully be grouped into two main ways of thinking about focus: the scope of the focus (how much of the sentence is highlighted) and the nature of the focus (what properties set it apart from other parts of the sentence).

The scope of focus is frequently split into just two categories: sentence focus and constituent focus, which highlight respectively all of the sentence and some constituent part of it. Moutouakil (1989) for example demonstrates that this distinction holds in Standard Arabic. In (8.1a) below, the whole sentence is in focus and thus could be a felicitous answer to the question "What happened today?"; in contrast, in (8.1b) and (8.1c) placement of what Moutouakil calls the 'tonic accent' (that is, the nucleus, or main prominence of the utterance, indicated in italics) signals constituent focus on that item, so that these two sentences would be infelicitous in response to the question "What happened today?" (Moutouakil 1989:25-26).

(8.1)	a.	raja9a	zayd-un	min	s-safar-i
		returned	Zayd-NOM	from	the-journey-GEN
		'Zayd has returned from his journey'			
	b.	saafara	zayd-un	*l-baariHat-a*	
		travelled	Zay-NOM	yesterday-ACC	
		'Zayd set out *yesterday*'			
	c.	saafara	*zayd-un*	l-baariHat-a	
		travelled	Zay-NOM	yesterday-ACC	
		'*Zayd* set out yesterday'			

As indicated here a common indicator of the location of constituent focus is by means of nucleus placement, and Moutouakil's examples suggest that prosodic marking of focus scope by means of nucleus placement is possible in Standard Arabic.

Ladd has described the same basic notion as a distinction between broad and narrow focus (Ladd 1980, Ladd 1996). Whereas a broad focus utterance carries 'all new' information such that the whole sentence is in focus, in a narrow focus utterance just some part is new or informative, and this is set against a background of 'old' or given information, often repeated from earlier in the discourse.

A key observation is that in certain cases broad and narrow focus utterances may be prosodically indistinct, since nuclear prominence on a sentence-final word could in principle be the reflex of either broad focus on the whole sentence (triggering 'default' utterance final prominence) or narrow focus on the sentence-final word itself, unless disambiguated by additional paralinguistic emphasis (cf. discussion in Ladd 1996:199ff.). The manner in which the semantic impact of a sentence-final nucleus 'percolates' from the final position in a phrase to the rest of the phrase has been analysed by means of focus projection from a focussed item to all other items within the syntactic constituent of which it is a part (Selkirk 1984, Selkirk 1995).

In defining the scope of narrow focus above it becomes necessary already to describe the nature of the relationship between the focussed item and the remainder of the sentence against which it is highlighted. This is the second common way in which types of focus are categorised.

Moutouakil makes a distinction in Standard Arabic between what he terms ‘new focus’ and ‘contrastive focus’. A new focus highlights information that is simply new whilst a contrastive focus highlights information that is in conflict with other information in the discourse (Moutouakil 1989:21; nuclear prominence marked in italics):

(8.2)	a.	*šaay-an*	šariba	xaalid-un		contrastive focus
		tea-ACC	drank	Khalid-NOM		
		“It was tea that Khalid drank”				
	b.	šariba	xaalid-un	*šaay-an*		new focus
		drank	Khalid-NOM	tea-ACC		
		“Khalid drank tea.”				

Note that according to Moutouakil, in Standard Arabic the focussed item must undergo syntactic movement in order to express contrastive focus, placement of the nucleus alone is not enough. It is on the basis of this kind of syntactic evidence from a number of languages that Kiss (1998) similarly distinguishes ‘informational focus’, which expresses the notion of given vs. new information, and ‘identificational focus’, which expresses contrastive and/or exhaustive identification[190].

These two types of focus will frequently be abbreviated for practical purposes here to *focus* [±f] for information focus, and *FOCUS* [±F] for contrastive focus (following Selkirk 2002, Selkirk 2004a).

Not all authors agree that the *focus*/*FOCUS* distinction is valid[191]; however, it is maintained in the present study for two reasons. Firstly, as has been seen the distinction is argued to be valid in Standard Arabic, on the basis of syntactic evidence, and may

[190] Kiss (1998) details syntactic evidence to show that languages may differ as to whether an identificational focus may optionally or obligatorily express exhaustive and/or contrastive identification. She notes that Standard Arabic requires an identificational focus to be both exhaustive *and* contrastive (following Ouhalla 1994). Since the main aim of the present study is to determine if there is a prosodic distinction between the main two focus categories (*focus*/*FOCUS*), it is beyond the scope of the present study to attempt to further distinguish prosodic reflexes in contrastive vs. exhaustive contexts.

[191] For example Göksel & Özsoy argue that the prosody of focus in Turkish can be analysed without making a distinction between new and contrastive focus (Göksel & Özsoy 2000, Göksel & Özsoy 2003).

therefore plausibly be relevant in spoken dialects of Arabic. Ouhalla (1999) points out that in Moroccan Arabic (MA) a contrastive focus is most naturally expressed prosodically by leaving the focussed item in-situ and marking it with a special pitch accent (indicated in *italics*), as in (8.3a) below, but can also be expressed syntactically using a structure which he describes as being “close, though not identical, to the structure of English cleft-sentences”, as in (8.3b) (Ouhalla 1999:338):

(8.3)	a.	naadia	shr-at	*ktab*	(maši	magalla)	in-situ
		Nadia	bought-she	book	(not	magazine)	
		‘Nadia bought a book (not a magazine)					
	b.	šaf	l-*bnt*	(maši	l-wld)		cleft structure
		saw-he	the-girl	(not	the-boy)		
		‘It was the girl he saw (not the boy).’					

As well as cross-linguistic syntactic evidence for the *focus*/*FOCUS* distinction, there is also prosodic evidence from other languages which suggests that the two types of focus may not only exist but also have distinct prosodic reflexes.

Another common way in which languages use exclusively prosodic means to mark the distinction between information focus and contrastive focus is through a choice of two distinct pitch accent types. The surface difference between the two pitch accents has in most cases been reported to be in alignment of the H peak of the pitch accent relative to the segmental string. This focus-marking strategy has been reported for Neapolitan Italian, European Portuguese and Madrid Spanish (see D'Imperio 1997, Frota 2000, Face 2002 respectively , cf. also Ladd 1996:127). Pitch accent (alignment) choice has also been explored as a marker of the *focus*/*FOCUS* distinction in English (Selkirk 2002).

Whilst a number of studies exist on the prosodic effects of focus in Arabic dialects, none have systematically explored whether information and contrastive focus are prosodically distinct, so the present study makes a potentially novel contribution to our knowledge about the prosody of Arabic. Since alignment of pitch targets is thought to be relevant to the *focus*/*FOCUS* distinction cross-linguistically it is one of the potential prosodic reflexes of focus examined in the experimental data presented here.

The second reason for maintaining the distinction between information focus and contrastive focus relates to the claim regarding rich pitch accent distribution in EA. In

English, a prime context in which a word must be 'de-accented' is when it occurs after a contrastively focussed (+*FOCUS*) item. However, as discussed already in chapter 5 (5.2.2.2), it has been pointed out that in English de-accenting in this position is not purely a function of post-*FOCUS* position; if a word is new (+*focus*) in this position it is argued to be more often than not accented (pitch accents are here denoted by an acute accent) (Selkirk 2000:p247 ex. 27iii & p251 ex. 36):

(8.4) I heard a rumour that she is selling all her stuff, but it says here that...
a. she lóaned$_{[+F]}$ [her róllerblades$_{[+f]}$] to Róbin. *object is discourse new*

I thought she sold her rollerblades to Robin, but it says here that...
b. she lóaned$_{[+F]}$ [her rollerblades$_{[-f]}$] to Robin. *object is discourse old*

Since one of the primary purposes for eliciting the focus data described here was to confirm whether or not words are ever de-accented in EA, the *focus* status of target words was varied, as well as their position (after a *FOCUS* or not), so as to be sure of creating parallel conditions to those which would be most conducive to de-accenting in a Germanic language such as English.

The working criteria used to define and/or identify the two types of focus during design of the datasets were the notion of 'givenness' for *focus* and 'contrast among alternatives' for *FOCUS*. These notions relate roughly to the general concept of focus as defined in Schwarzschild (1999) and Rooth (1996) respectively, even though these authors do not themselves make a distinction between two different types of focus[192].

The most robust context for an item to be considered 'given' was deemed to be one in which it has been previously uttered in the same discourse chunk (eg in the same or a preceding sentence). This is parallel to the Hallidayan notion of being 'textually given', as opposed to 'situationally given' (which refers to items that are salient in the discourse but have not actually been uttered, such as near synonyms or items that are 'understood' from context: Halliday 1967:23)[193].

The most robust context for an item to be considered contrastively focussed was deemed to be one in which the focussed item was picked out from a set of two overt

[192] Use of these two independent semantic notions to define a single *focus*/*FOCUS* distinction follows Selkirk & Kratzer (2004).
[193] Situationally given items may also be de-accented in English (Brown 1983, Cruttenden 2006).

alternatives, in the spirit of Face's (2002:4) definition: "explicit contrast between an extremely limited set of two discourse elements". In most cases this was achieved by overt prior mention of the alternative ('I saw Y, but he saw X.') or insertion of an explicit continuation ('I saw X, not Y'), or both.

8.1.2 Focus in Egyptian Arabic

As seen above a commonly described prosodic reflex of focus location is placement of the nucleus, or main prominence, in a sentence. The 'default' position for sentence prominence in other Arabic dialects has been shown to be on the final content word (Benkirane 1998, Chahal 2001), but the example above from Moroccan Arabic (MA, see 8.3 above) shows that in at least some Arabic dialects the nucleus can be moved to a non-final content word in order to highlight that item. There are some reports in descriptive grammars of EA to suggest that this kind of 'nuclear mobility' is also observed in EA.

Mitchell (1993:230) 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".

Although he is generalising across Arabic dialects in this statement, the first example he gives (as mentioned already in chapter 3) is from EA (reproduced in (8.5) below). He points out that, whilst a final nucleus on [maSri] 'Egyptian' is the unmarked form, locating the nucleus on either [?itneen] 'two' or [gineeh] 'pounds' is "perfectly possible and natural" (Mitchell 1993:230):

(8.5)	?itnéen	ginéeeh	máSri
	two	pounds	Egyptian
	'Two Egyptian pounds"		

Crucially for the overall claim of this thesis, as well as that of this chapter, Mitchell does not specify whether material following an early nucleus is de-accented or not (see discussion in chapter 3 section 3.1.1).

Mitchell does however point out that it is only possible to move the nucleus in EA when the focussed semantic notion is expressed using an independent lexical item. Many elements of the sentence in Arabic, such as object pronouns, are expressed using clitic particles which form a single PWd with their host lexical head. He gives the example of the English four word sentence ‘Did you see him?’ which in Arabic is expressed by means of a single word [šuftuh] ‘saw-masc.sing.+him’ (Mitchell 1993:231-2):

(8.6)	a.	Did you see *him*? ie not *her*	šuftuh saw-him	huwwa he
	b.	Did *you* see him? ie not someone *other than you*	?inta you	šuftuh saw-him

In a parallel case in English the location of the nucleus could be moved around to shift the focus within the sentence. This option is not available in EA, since it is not possible to shift lexical stress within the word[194], and instead additional particles must be used to express focus (as in 8.6a above). This property of EA is something that we will return to in the discussion in section 8.5.

Gary & GamalEldin (1981:49) state that what they call ‘emphasis’ can be expressed: “in a variety of ways, one way being to stress that part of the sentence being emphasised, at the same time as raising its intonation contour to a higher pitch ... however emphasis can also be expressed by emphatic particles, movement of constituents and repetition of certain constituents”.

Gary & GamalEldin distinguish distinct prosodic reflexes of two types of emphasis which they describe as ‘non-contradictory’ and ‘contradictory’. These terms seem to equate to *focus* vs *FOCUS* respectively (the examples of non-contradictory emphasis do not seem to involve exhaustive selection of one option from among alternatives, but rather of stating new information about an established, given, topic). They note that non-contradictory emphasis involves ‘raising’ the whole sentence (by which they appear to mean the pitch range of the entire sentence), with rising pitch and lengthening of the final syllable, followed by either slightly rising or falling pitch, as illustrated in their examples reproduced in (8.7) below.

[194] Gussenhoven (2004) discusses the notion of cross-linguistic variation in the ‘minimal domain of focus’, with English allowing accent shift within a word in order to convey contrast, as in the famous: “This whisky should be DEported not EXported” (after Bolinger, cited in Ladd 1996:177-8). EA does not permit accent shift to express contrast in this way.

(8.7) Examples of 'non-contradictory emphasis' (Gary & Gamal-Eldin 1981:49)[195].

a. SaHb-ak da raGGaay bi-šaakl
friend-your that talkative in-appearance
'Your friend is extremely talkative!'

b) 9amm-ak da šaxsiyya zariifa
uncle-your that character pleasant
'Your uncle is such a pleasant character!'

In contrast, contradictory emphasis is expressed by stating the negated presupposition first, followed by a pause, and then the correcting statement is made, with falling pitch, characterised by "greater stress and higher intonation" than the denied portion:

(8.8) Examples of 'contradictory emphasis' (Gary & Gamal-Eldin 1981:50).

?il-kitaab da miš ?azra? / da ?aHmar
the-book that NEG blue that red
'That book isn't blue; it's red.'

(8.9) ?ana ma-?ult-iš ?in-ni raayiH ?iskandariyya / ?ana raayiH TanTa
I NEG-said-NEG that-I going Alexandria I going Tanta
'I didn't say that I was going to Alexandria; I am going to Tanta.'

They report that almost any constituent can be emphasised " by a combination of stronger stress and higher intonation" (ibid. p50). Exceptions include definite articles, prepositions (eg [*fi*] 'in', [*li*] 'to') and the relative pronoun [*?illi*]. Increased stress and 'higher' intonation on a focussed constituent, following a negated presupposition, creates contrastive stress (Gary & Gamal-Eldin 1981:51):

(8.10) miš beeD, zeet
not eggs oil
"Not eggs... oil!"

As regards syntactic expression of focus in EA, Gary & GamalEldin note that noun phrases, constituents of main clauses, adverbial phrases and whole verb phrases can all be clefted or pseudoclefted (ibid. p52). It is not clear here whether they are referring to the expression of information focus or contrastive focus. If the latter, then the situation in EA would be parallel to that observed in MSA by Moutouakil (1989).

[195] Interpretation of these sentences is slightly complicated by the use of [da] 'that' which could be used to single out the subject from other possible reference; if so this is an exhaustive~contrastive distinction.

Another author who mentions the possibility of nuclear mobility is Heliel (1977), in the introduction to a study of the rhythmic properties of EA. Heliel gives examples in which his notation appears to suggest that focus conditions an IP boundary[196], and hence two nuclear accents are produced (in 8.11) or nuclear mobility (in 8.12 & 8.13). The examples are from Heliel (1977:125, 132), with the nucleus marked in bold type, following his notation.

(8.11)	a.	?ana	gib**ta**ha		'I brought it.'
		I	brought-it		
	b.	**?a**na	gib**ta**ha		'I brought it, not anyone else.'

(8.12)	a.	?a xu:k	**saafir**	imbaariH	*interested in travelling*
		brother-your	travelling	yesterday	
	b.	?a**xu:k**	saafir	imbaariH	*interested in the traveller*
	c.	?axu:k	saafir	im**baariH**	*interested in time of travel*
		'Your brother travelled yesterday.'.			

(8.13)	a.	si**mi9na**	9ali	'We heard Ali.'
		heard-we	Ali	
	b.	simi9na	**9ali**	'Ali heard us.'
		heard-us[197]	Ali	

Heliel uses the term 'tonic' to denote the nucleus and, in the terms employed here, suggests that *FOCUS* is expressible in-situ by prosodic means in EA: "the place of the tonic in Arabic is not fixed but varies meaningfully thus creating an independent set of choices" (ibid. p132).

The consensus from the descriptive literature then, is that it is possible to express *focus*, and possibly also *FOCUS*, in-situ by means of nucleus placement. However the evidence from instrumental studies appears to conflict with this conclusion, as outlined in the next section.

---

[196] Heliel states that (in his notation) bold type indicates the tonic of the sentence, yet in the example reproduced in (8.11b) he has two words with the stressed syllable in bold: suggesting that there are two tonic accents in this two-word sentence (the first of which is a pronoun). If we interpret this as indeed indicating a sentence with two tonic accents, in the model of EA intonation used here (and as motivated in chapter 3 section 3.4.3) this implies insertion of an IP boundary and boundary tone (either L% or H%), even though it is hard to justify [?ana] 'I' and [gibtaha] 'brought-it' as separate root clauses. Re-interpretation of his notation to indicate merely the presence of a pitch accent is implausible since each of the content words in (8.12a-c) would routinely be accented in EA. An alternative re-interpretation would be that his notation does not indicate nuclear accents but *relatively more prominent* accents. This latter will prove to be more sustainable in the light of the results of the focus experiment described in this chapter (see section 8.3 ff. above).

[197] The 1st person plural suffix in the Perfect, [-na], is identical in form to the pronominal object suffix 'us'.

8.1.3 Categorical vs. gradient marking of focus in EA: focus in-situ

Hellmuth (2005) reports the results of a small pilot study carried out to investigate the interaction between focus structure and pitch accent distribution in EA, following the methodology of Swerts et al. (2002). Semi-spontaneous scripted utterances were recorded during the course of a card game, with cards manipulated so as to elicit target structures in a particular Focus context. The target structures in question were either a single NP or a sequence of phrases (NP-NP-PP). The Focus status of each item in a phrase was potentially new (+*focus*), contrasted (+*FOCUS*), or given (-*focus*). The contexts of particular interest for our present purposes are sequences of PWds where the first is [+*FOCUS*] and the second is [-*focus*]. In the pilot study there was only one token (out of 48) in which the second PWd in such a sequence was produced without a pitch accent.

The lack of de-accenting after a *FOCUS* is in line with the findings of the corpus survey described in chapter 3. Indeed, in the light of the results of Norlin (1989) discussed below, it may not conflict with the idea that in EA the nucleus, or at least the 'most salient' pitch accent in the utterance, can be moved around in order to express focus.

Norlin (1989) reports gradient reflexes of focus in a small pilot study carried out with one speaker of EA. He elicited parallel renditions of an SVO sentence, embedded in different frame paragraphs in order to elicit either a statement or a question[198], and with either broad focus over the whole sentence or narrow focus on just one part (the subject, the verb or the object). The target sentence is provided in (8.14) below. The passages were recorded with a single speaker of EA, hence Norlin is able to report absolute F0 figures in Hertz in his results.

(8.14)	muniir	il-marin	rama	l-lamuun	il-murr
	Munir	the nimble	threw	the lemons	the bitter
	'Nimble Munir threw the bitter lemons.'				

Norlin carefully describes the F0 properties of the neutral declaratives in the dataset ("a neutral statement is realized globally by a continuous declination") then documents in what ways the F0 contour of non-neutral utterances vary.

[198] Recall that in EA a 'declarative question' has the syntactic form of a statement but with question status indicated by intonation (see discussion in chapter 3 section 3.4.1). Since the present study involved only statements Norlin's observations in declarative questions are not included in the survey of his results here.

Statements with the subject noun phrase in focus do not display continuous downdrift (as compared to a neutral statement), instead the pitch peak(s) of the focussed phrase are higher than in the neutral statement, and then the focus F0 contour dips below that of the neutral statement until the end of the utterance. Norlin found that statements with the sentence-final object in focus shared the same contour as neutral statements in the pre-focus part of the utterance, which suggests that the effects of focus were directional, affecting the focussed item and any linearly *subsequent* items, rather than also affecting pre-focus items. These results suggest then that focus can be expressed in EA by both expanding F0 excursion on the focussed item and compressing F0 on following items. As Norlin points out, the focus/post-focus distinction is achieved by manipulating the pitch range of the whole of the remainder of the sentence, not just the focussed part.

Unfortunately however Norlin does not report the exact design of the frame paragraphs that were used for elicitation purposes in his study, so whilst we know that some kind of focus was elicited on the target items (S, V or O) there is no way of knowing exactly what type of focus, nor the focus status of other elements in the surrounding sentence. A key goal of the experimental study described in this chapter is therefore to reproduce Norlin's results and to clarify which type(s) of focus are marked by means of gradient pitch range manipulation.

Norlin's results are paralleled in a much larger study by Chahal (2001:124ff.) on acoustic cues to focus in Lebanese Arabic (LA). Chahal elicited instances of a series of double-object sentences (such as [lama Hamet muna min lima] 'Lama protected Muna from Lima'), by means of a question-answer paradigm. By varying the question Chahal elicited the sentence with focus on any one of the three proper name noun phrase arguments, which acted as the target words for the study: the subject, direct object or indirect object. The names were varied so as to elicit focus in each position on a name containing each of the three vowels: [a], [i] and [u]. Chahal then measured a range of potential different acoustic cues to focus in the target words: F0, intensity, duration and F1/F2 values. These were investigated alongside categorical cues to focus in the dataset, which Chahal also reports, and which include insertion of a prosodic boundary and post-focal de-accenting.

As regards the gradient correlates of focus status, Chahal found that each of the four potential correlates were *enhanced* in focussed words compared to their counterparts in a neutral sentence: F0 was higher, duration was longer, intensity was increased, and vowels were more dispersed. In addition, she found that in such sentences the potential acoustic cues to focus were *reduced* in non-focussed words, as compared to the same correlates in their counterparts in a neutral sentence: F0 was lower, duration was shorter, intensity was decreased, and vowels were more centralised.

In contrast to Norlin's findings, Chahal found that the gradient acoustic cues were reduced in *both* of the non-focus target words in a sentence, regardless of their linear position with respect to the focussed word, before or after. She analyses these findings as 'hyperarticulation' of focussed words accompanied by under articulation of non-focus words (based on proposals made by Lindblom 1990).

Whatever the precise details of the directionality of gradient effects of focus in EA, a more pressing question is to know exactly what type of focus context conditioned the observed effects. In the case of Chahal's study, the eliciting questions are reported and thus we know that she elicited a narrow (exhaustive) focus (that is, '+*FOCUS*') on her primary target word, and that other target words were always given in context (that is, '*-focus*'), since they were appeared in the questions and were thus 'textually given' in the speaker's answering response.

8.1.4 Categorical vs. gradient marking of focus in EA: syntactic focus

A further matter for investigation arises because neither Norlin (1989) nor Hellmuth (2005) elicited data which differentiated an in-situ contrastive focus from a syntactic contrastive focus. As seen in section 8.1.1 above, Moutouakil (1989) and others have noted that MSA is a language like Hungarian in which contrastive focus (*FOCUS*) may only be expressed via a syntactic strategy (such as a cleft).

One description of focus strategies in EA suggests that it may pattern with MSA in this respect: Gary & Gamal-Eldin (1981:126) report that use of contrastive stress is "limited to sequences where the contrasted elements are explicit". The example they provide, reproduced in (8.15) below, is one in which explicit contrast is expressed by means of a syntactic clefting strategy (introduced by the relative clause marker *lli*) (Gary & Gamal-Eldin 1981:126):

(8.15) humma-lli-?insaHabu-miš-iHna
they-who-withdrew-NEG-us
'They were the ones who withdrew not us'

It is not clear from this example whether it is use of the syntactic strategy (clefting) that licenses a *FOCUS* interpretation, or merely the presence of an explicit contrast.

Elicitation with a single EA speaker of parallel target utterances to those in Hellmuth (2005), but with an explicit contrast added[199], suggested that explicit contrast could only be expressed using a cleft: the speaker found it difficult to formulate a response when a fixed word order was required. Auditory analysis of the resulting clefted utterances showed clear pitch accents on contrasted items and on the negative marker, and subsequent words appeared to be produced in a compressed pitch range.

Frota (2000) has suggested that in some languages prosodic effects of focus may be contingent on implementation of a syntactic focus strategy (which via the syntax-phonology interface results in a change in prosodic phrasing). It is therefore possible that de-accenting could be conditioned in EA only by use of syntactic-*FOCUS*.

In order to address this issue the experiment described in the remainder of this chapter differentiates between in-situ *FOCUS* and syntactic *FOCUS*, placing -*focus* (given) targets after both *FOCUS* types. These two contexts are difficult, if not impossible to elicit in a semi-spontaneous manner, and this was a primary factor in the decision to use a read speech experimental design here, placing target sentences within paragraph frames to manipulate context. This type of methodology has been used successfully by a number of authors (Norlin 1989, Sneed 2004). The potential for register interference from use of written prompts was mitigated by the use of lexical items exclusive to EA, placed in the target sentences wherever possible, and used liberally in filler sentences interspersed with targets in the final design.

8.1.5 Rationale of the experimental focus investigation

The remainder of this chapter describes the methodology and results of an experiment carried out to identify the prosodic reflexes of both *FOCUS* and *focus* in EA.

[199] This was achieved by adding 'struck through' game cards to those used in the original study, and eliciting sequences either with the cards in fixed 'slots' or placed in any order by the subject.

Some aspects of the focus corpus have already been discussed in earlier chapters, and in particular the fact there appear to be no categorical effects of any type of focus on pitch accent distribution. Given the results of Norlin's study we do expect however to find gradient reflexes of focus but it could be either *focus* or *FOCUS* (or indeed both) that conditions gradient effects. The experiment should provide evidence regarding the prosodic reflexes (if any) of both types of focus.

The experiment also investigates whether the prosodic effects of *FOCUS* are enhanced (or possibly contingent upon) use of a syntactic *FOCUS* strategy (such as fronting by means of a cleft). If gradient prosodic reflexes are observed only in syntactic-*FOCUS* conditions this would indicate that EA is a language in which the prosodic effects of focus are contingent on the syntactic strategy. If such prosodic reflexes as are observed occur equally in *FOCUS* in-situ and syntactic-*FOCUS* contexts then this would indicate that the prosodic focus strategies can be used by speakers independently of their choice of syntactic structure.

The next section describes in detail the methodology used to carry out the experiment.

8.2 Methodology - data collection and analysis

8.2.1 Materials

In order to clarify the empirical facts of EA focus effects, two lexically distinct SVO target sentences were each placed in one of four frame paragraphs designed to manipulate the relative focus relations within the sentence.

Both the *FOCUS* status of the subject of the sentence (referred to as the 'trigger') and the *focus* status of the direct object (the 'target') are varied, resulting in four possible *FOCUS~focus* combinations between the trigger and target (note that the subject trigger word was designed to be 'new' in all contexts):

(8.16) Four-way combination of FOCUS~focus conditions in target sentences.

		trigger	*target*	
Set 1	[+F+f]	+FOCUS	+focus	$\text{Mum}_{[+F]}$ learns $\text{Greek}_{[+f]}$.
Set 2	[+F-f]	+FOCUS	-focus	$\text{Mum}_{[+F]}$ learns $\text{Greek}_{[-f]}$.
Set 3	[-F+f]	-FOCUS	+focus	$\text{Mum}_{[-F]}$ learns $\text{Greek}_{[+f]}$.
Set 4	[-F-f].	-FOCUS	-focus	$\text{Mum}_{[-F]}$ learns $\text{Greek}_{[-f]}$.

The two target sentences and a translation of the eight context paragraphs used in the experiment are set out in the tables in (8.17) and (8.18) below. The context paragraphs in Arabic with interlinear gloss for lexical set A are provided in Appendix D.1.

The basic technique used to generate differences in focus status was whether or not a word is repeated from earlier in the paragraph (ie whether or not the word is ‘textually given’ (Halliday 1967)). Note that the subject trigger word was designed to be ‘new’ [+f] in all contexts. This was achieved by alternating between the more formal [?ummi] ‘my mother’ and more usual colloquial [maama] ‘mum’. Whilst these two words appear to be highly synonymous and of course co-referent, the difference in register (formal vs. colloquial) was deemed sufficient to prevent interpretation of the word [maama] ‘mum’ as textually given in the context paragraphs. Crucially, in these ‘trigger’ words it is [±F] (contrastive focus) status that is at issue rather than the [±f] (information focus) distinction, for the purposes of the experiment.

The final adverbial phrase [bil-layl] ‘in the evenings/at night’ was included in order to elicit a pre-nuclear rather than nuclear pitch accent on the preceding target word, since the properties of phrase-final pitch accents are known to be affected by their proximity to the phrase boundary[200]. The context paragraph was designed so that the final adverbial phrase would always be given and not inadvertently create a sense of identificational focus on the adverbial by generating alternatives (such as study in the evenings vs. (inferred) study during the day).

(8.17) SVO sentences used in the focus experiment.

	⇩ trigger		⇩ target	
A	*mama*	*bitit9allim*	*yunaani*	*bil-layl*
	mum	learns	Greek	in-the-evening/night
	‘Mum	is learning	Greek	in the evenings’
B	*mama*	*bitnayyim*	*in-nounou*	*bil-layl*
	mum	puts-to-bed	the-baby	in-the-evening/night
	‘Mum	puts	the baby to bed	at night’

200 Nuclear pitch accents in EA are analysed in chapter 3 as being of the same phonological specification as their pre-nuclear counterparts but subject to final lowering and/or peak retraction (for alternative analyses see the discussions in chapter 3 section 3.4.2.4 and also chapter 6 section 6.3.3).

(8.18) Context paragraphs used in the focus experiment.
Trigger = [maama] 'Mum' — contrastive focus status varied ±F
Target = [yunaani] 'Greek' or [nounou] 'baby' — information focus status varied ±f

A1	[+F+f]	'My colleague said they heard my dad went to university in the evenings but I told him no. *Mum is learning Greek in the evenings.* Dad sits at home and watches TV.'
A2	[-F+f]	'My mother loves learning new things. *Mum is learning Greek in the evenings* and she also studies history.'
A3	[+F-f]	'My colleague said they heard my dad was learning Greek in the evenings but I told him no. *Mum is learning Greek in the evenings.* Dad sits at home and watches TV.'
A4	[-F-f]	'My mother loves Greek. *Mum is learning Greek in the evenings* and she likes to watch films on Greek history.'
B1	[+F+f]	'My aunt said she heard my dad puts the kids to bed at night for my sister but I told him no. *Mum puts the baby to bed at night*. Dad reads a story to the girls.'
B2	[-F+f]	'My sister is ill at the moment so my mum helps her get the kids to bed at night. *Mum puts the baby to bed at night* and reads him a story.'
B3	[+F-f]	'My aunt said she heard my dad puts the baby to bed at night for my sister but I told him no. *Mum puts the baby to bed at night.* Dad reads him a story.'
B4	[-F-f]	'My sister is ill at the moment so my mum is helping her with the baby at night. *Mum puts the baby to bed at night* and reads him a story.'

In addition to the plain SVO target sentences a further three variations of each of the two sentences were created in which the subject was highlighted using either a phrase-initial pseudo-cleft construction or a phrase-final negative continuation or both. These are set out in (8.19) below (sentences 1A and 1B are the plain SVO versions, repeated from 8.17 above).

The 'SVO+' extended sentences were placed in each of the four context paragraphs and three EA speakers (who did not participate in the later recordings) were asked to provide grammaticality judgements about the paragraphs. It was expected that the extended sentences (in which a syntactic strategy is used to contrastively focus the subject) would be judged infelicitous in context paragraphs designed to generate a [-F] subject. This was indeed the case, and was unanimous across the three speakers.

As a result the SVO+ sentences were placed only in the two paragraphs with +F subjects ([+F+f] and [+F-f]) for subsequent recording and analysis. The grammaticality judgements also serve however to confirm that the context paragraphs used in the study do successfully create plausible and retrievable focus structure.

(8.19) SVO and ‘SVO+’ sentences used in the focus experiment.

	⇩ trigger			⇩ target		
1A	*mama*		*bitit9allim*	*yunaani*	*bil-layl*	
	mum		learns	Greek	in-the-evening	
	‘Mum		is learning	Greek	in the evenings’	
2A	*mama*		*bitit9allim*	*yunaani*	*bil-layl*	*miš baaba*
	mum		learns	Greek	in-the-evening	not Dad
	‘Mum		is learning	Greek	in the evenings,	not Dad’
3A	*mama*	*hiyya illi*	*bitit9allim*	*yunaani*	*bil-layl*	
	mum	she who	learns	Greek	in-the-evening	
	‘It’s Mum	who	is learning	Greek	in the evenings’	
4A	*mama*	*hiyya illi*	*bitit9allim*	*yunaani*	*bil-layl*	*miš baaba*
	mum	she who	learns	Greek	in-the-evening	not Dad
	‘It’s Mum	who	is learning	Greek	in the evenings,	not Dad’
1B	*mama*		*bitnayyim*	*in-nounou*	*bil-layl*	
	mum		puts-to-bed	the-baby	at-night	
	‘Mum		puts	the baby to bed	at night’	
2B	*mama*		*bitnayyim*	*in-nounou*	*bil-layl*	*miš baaba*
	mum		puts-to-bed	the-baby	at-night	not Dad
	‘Mum		puts	the baby to bed	at night	not Dad’
3B	*mama*	*hiyya illi*	*bitnayyim*	*in-nounou*	*bil-layl*	
	mum	she who	puts-to-bed	the-baby	at-night	
	‘It’s Mum	who	puts	the baby to bed	at night’	
4B	*mama*	*hiyya illi*	*bitnayyim*	*in-nounou*	*bil-layl*	*miš baaba*
	mum	she who	puts-to-bed	the-baby	at-night	not Dad
	‘It’s Mum	who	puts	the baby to bed	at night	not Dad’

8.2.2 Data collection and analysis

The full dataset comprises 8 ‘SVO’ paragraphs (2 lexical sets x 4 focus contexts) and 12 ‘SVO+’ paragraphs (2 lexical sets x 3 syntactic combinations x 2 focus contexts). The resulting 20 paragraphs were interspersed with an equal number of filler paragraphs, then pseudo-randomised and divided into 5 sets of 8 paragraphs in such a way that each set contained two paragraphs from each lexical set.

Each set was printed over two pages and no two paragraphs from the same lexical set appeared on the same page. The sets of paragraphs were read 3 times each by 6 speakers of EA, yielding a potential 18 tokens x 18 targets (N=324) for analysis.

It would have been ideal to have subjects read each set of paragraphs on a different day, to exclude the possibility of interference between different contexts; however, this was not possible in the recording time available. In order to reduce potential interference, after the third repetition of each set the speaker performed a different style of task (reading or re-telling a narrative, a task which they found diverting and in many cases thought was the real purpose of the recording session). This rotation of tasks served to break up the pace and pattern of the recording session, in order to facilitate interpretation of each paragraph from its own internal structure, rather than in comparison with paragraphs in other sets.

Three investigations were carried out on the resulting recordings with reference to F0 and spectrogram using Praat 4.2 (Boersma & Weenink 2004):

i) a qualitative analysis, to determine the categorical presence or absence of pitch accents on target words;
ii) a quantitative analysis, to determine whether there is gradient variation in F0 excursion in trigger and/or target words;
iii) an alignment investigation to determine whether there are differences in alignment in trigger and/or target words (which might be interpreted as different pitch accent choices).

The categorical analysis was based on whether or not a local F0 maximum occurred during each target word, and thus aimed to determine whether or not target words were ever 'de-accented'. The target word in each token was labelled by hand as an interval using Praat 4.2 and the automatic pitch maximum identification function used as a guide in deciding whether a local F0 maximum occurs within (or near to) the target word. When this method is used on unaccented function words the local maximum is identified as being at the start of the word, because pitch simply falls steadily throughout the word. This was seen as being a practical and unambiguous way to determine whether a F0 maximum occurs or not, avoiding labeller bias. In every such instance, the absence of an F0 maximum would be interpreted as an instance of de-accenting.

A sub-set of the present focus data (one lexical set) were included in the corpus survey whose results are reported in chapter 3. Those results showed that there was no categorical, full, de-accenting of target words, even in post-*FOCUS* position or in sentences which used syntactic *FOCUS* strategies. It is anticipated therefore that the categorical analysis of target words here, taking in the full dataset, will yield similar results, that is, no de-accenting of target words.

The gradient analysis used F0 excursion as the dependent variable in order to determine whether there were gradient effects of focus on F0 in both target and trigger words. The position of the minimum (L) and maximum (H) F0 turning points associated with the trigger and target word in each token was labelled by hand (using the automatic pitch minima/maxima function within Praat 4.2 as a guide). The F0 value at each of these points was then extracted in semitones and F0 excursion within each word was calculated by subtraction: 'xn' = F0max - F0min.

F0 excursion was calculated in this way for the trigger word ('xxn') and target word ('yxn') in each token, and the differential in excursion between each trigger-target pair was calculated (in semitones): 'xndf' = 'xxn' - 'yxn'. The expectation is that F0 excursion would be greater in focussed words, as was observed by Norlin (1989) in EA, and by Chahal (2001) in Lebanese Arabic. The distinction made here between *focus* and *FOCUS* was designed to clarify which type(s) of focus are marked by F0 excursion in EA.

If pitch range manipulation in EA reflects *FOCUS*, with expansion of pitch range on focussed items, F0 excursion in trigger words ('xxn') will be greater in +F contexts than in -F contexts; and with pitch range compression on post-FOCUS items, F0 excursion in target words ('yxn') will be smaller in +F contexts than in -F contexts. If, instead, pitch range manipulation in EA reflects *focus*, with expansion of pitch range on +*focus* items, F0 excursion in target words ('yxn') will be greater in +f contexts than in -f contexts[201].

If however pitch range manipulation reflects *both* types of focus to some extent, then we expect a large excursion differential between target and trigger ('xndf') in [+F-f] condition, in which the trigger words is new and bears *FOCUS*, whereas the target word is given and therefore bears neither *focus* nor *FOCUS*. F0 excursion differential properties in other conditions are harder to predict, but may reveal in what ways the two types of focus are marked, if they are both marked.

The alignment analysis investigated the alignment properties of the pitch contour relative to the segmental string. Specifically, the distance of the F0 peak (H) from the consonantal onset of the stressed syllable (C0) was measured in both trigger and target words. The position of the segmental landmark, together with pitch events already

[201] There will be no effect of focus on trigger words, since all are new in context.

retrieved for investigation of F0 excursion, was labelled by hand in each target syllable as in Figure 8.1 below. Calculation of the excursion variables is illustrated in Figure 8.2

Figure 8.1 Schematised labeling diagram of the position of the C0 segmental landmark & the H F0 peak

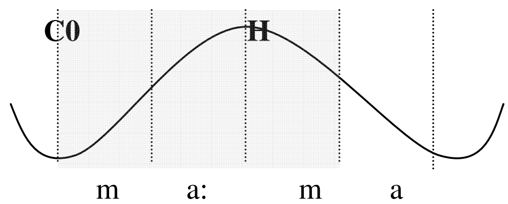
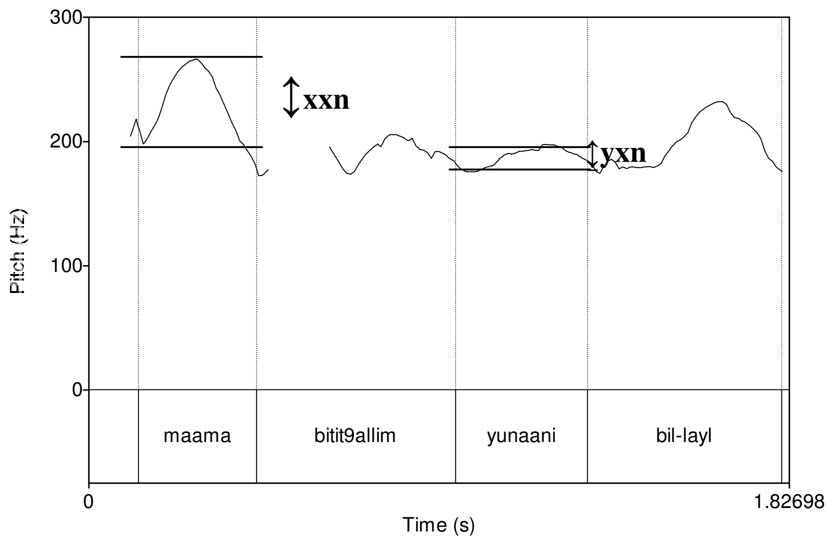


Figure 8.2. Calculation of excursion differential (xndf): xndf = xxn - yxn.



As discussed above, it has been argued that in some intonational languages the distinction between *FOCUS* and *focus* is expressible by means of a difference in pitch accent alignment, which in most analyses this difference is thought to be categorical, to the extent that distinct phonological representations are proposed for the two accents. The surface distinction between the two accent types is usually a difference in peak alignment. For example in European Portuguese +F+f nuclear falls have an earlier peak than -F+f nuclear falls (Frota 2000); in Spanish, +F pre-nuclear rising accents have an earlier peak than their +f counterparts (Face 2002).

The key dependent variable for comparison across focus conditions in the current experiment is peak delay. This was retrieved from both trigger and target words (in milliseconds): trigger peak delay (XH-XC0) and target peak delay (YH-YC0). If there is an effect of *FOCUS* on peak alignment in EA pitch accents the values of peak delay are expected to vary significantly in trigger and/or target words between ±F conditions.

8.3 Results

8.3.1 Review: results of categorical analysis (presence or absence of pitch accents)

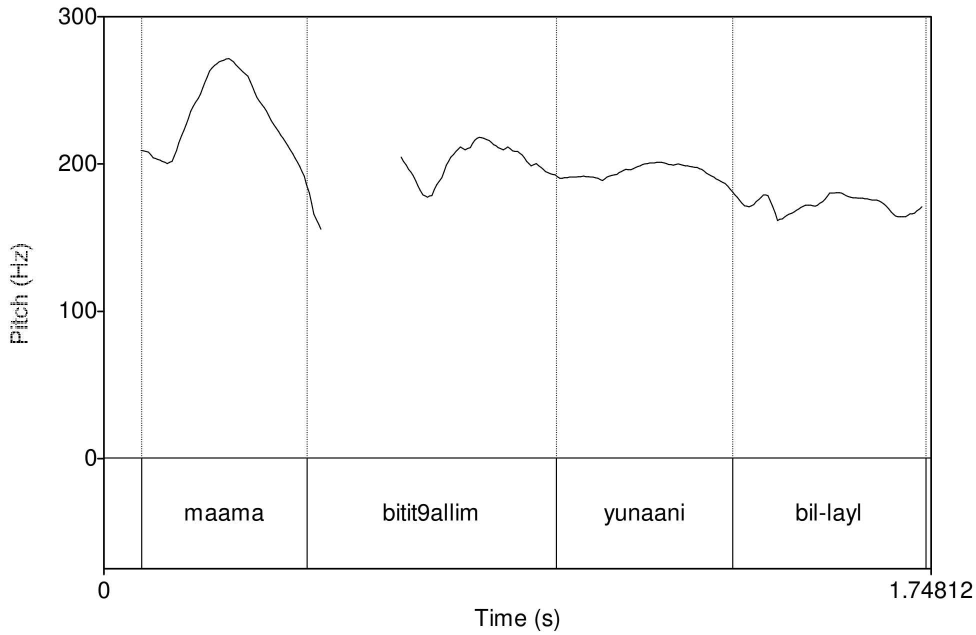
Categorical analysis of target words in the full focus dataset (144 SVO sentences + 216 SVO+ sentences) reveals that in all 360 tokens there is a local F0 maximum on or near the target word, which is taken to be a pitch movement associated with the word. This confirms the results of the corpus survey (which analysed only half of the data, just lexical set A). There is thus no categorical de-accenting of target words in the EA data, regardless of the *focus* status of the target, nor the *FOCUS* status of the trigger, nor whether *FOCUS* is expressed in-situ or by syntactic means (cleft or continuation or both). This provides strong support for the overall claim of this thesis, that in EA there is a pitch accent on every Prosodic Word, and that maintenance of rich pitch accent distribution is phonologically important.

A set of four typical pitch tracks are provided in Figure 8.3-4 below, showing target words with new information status in [+F+f] and [-F+f], and targets with given information status in [+F-f] and [-F-f], respectively. In Figure 8.4a, illustrating a given target following a contrastive focus, it is visually clear that there is a pitch movement on the target word [yunaani] ‘Greek’, but it is also clear that the degree of F0 excursion varies in the different words, suggesting that gradient manipulation of pitch range is likely to be relevant in EA.

The following sections explore the results of the gradient analysis, first in the in-situ focus sentences [SVO], in section 8.3.2, and then in the syntactic focus sentences[SVO+], in section 8.3.3.

Figure 8.3 Sample pitch tracks of 'new' target words (+f):

a) +F+f condition (*121faa1*).



b) -F+f condition (*122faa1*).

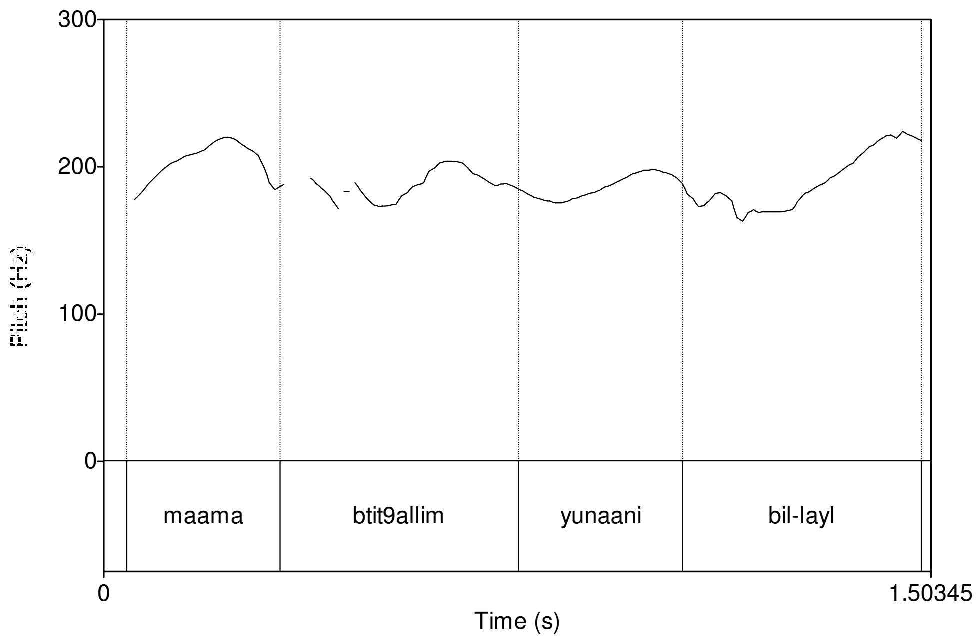
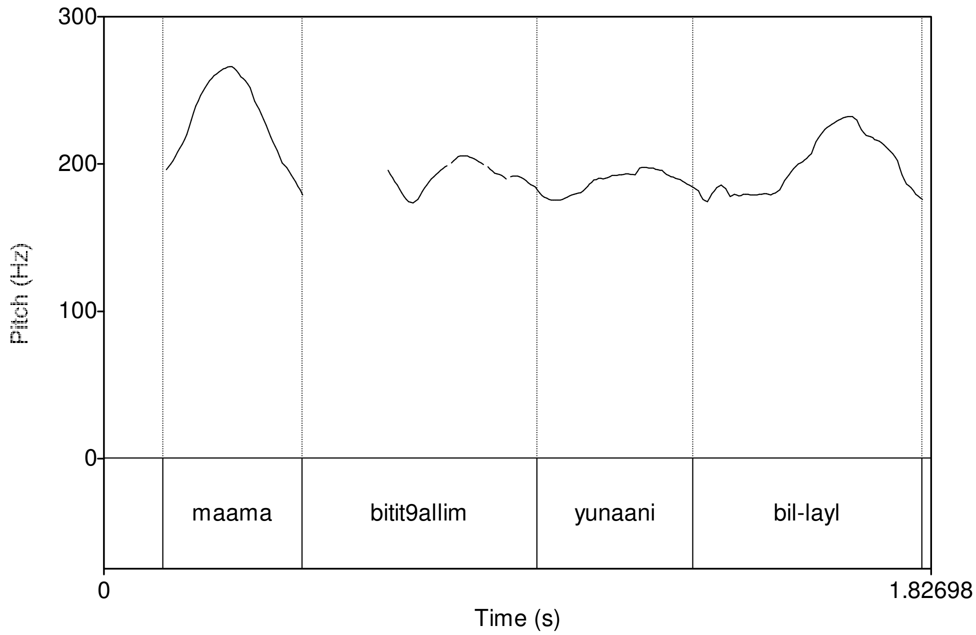
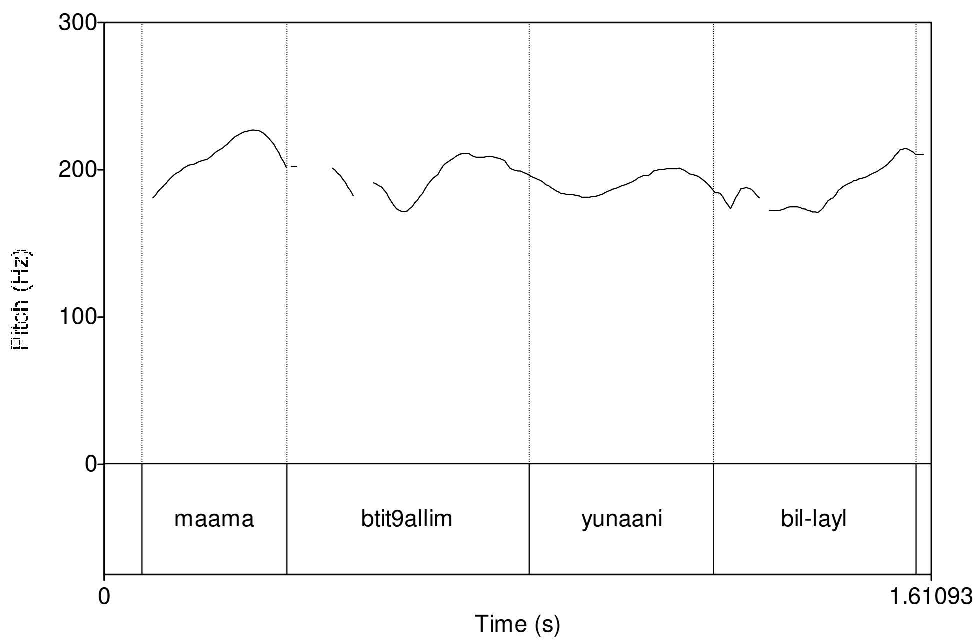


Figure 8.4 Sample pitch tracks of 'given' targets (-f):

a) +F-f condition (*123faa1*).



b) -F-f condition (*123faa1*).



8.3.2 [SVO] results: gradient effects in in-situ focus sentences

In this section F0 excursion in trigger and target words are examined separately to determine whether the gradient expansion and compression of F0 patterns observed in Norlin's study is reproduced here, and whether any such effects reflect contrastive focus or information focus status (or both).

Looking first at F0 excursion in *trigger* words (xxn), this can only be expected to vary with contrastive focus status (±F), since all of the context paragraphs were designed to elicit trigger words as new in context. To reproduce Norlin's (1989) result, F0 excursion in trigger words (xxn) should be greater in +F contexts than -F contexts.

Figure 8.5 below displays mean values of xxn by focus condition and by speaker. The patterns of F0 excursion produced by female speakers vary as expected, with greater mean F0 excursion in +F than -F contexts. The male speakers exhibit considerably more variation[202], and as a result, the degree of variation in behaviour among speakers leads to a non-significant result when mean values of xxn in +F vs -F condition are compared across all speakers (Tamhane's test N.S).

Analysis of the female speakers' data only, using a oneway ANOVA (xxn by focus condition), shows that the differences in mean values of F0 excursion in trigger words among female speakers are highly significant ($p<0.001$). A post-hoc test (Tukey HSD) shows that the observed differences in mean values are significant and that the distinctions are in the expected direction, reflecting contrastive focus status. Mean trigger F0 excursion across female speakers varies significantly between [+F+f] and [-F+f] ($p=0.046$) and also between [+F-f] and [-F-f] ($p=0.025$). (Full tables of post-hoc test results for xxn across all speakers and among female speakers only are provided in Appendix D.2 and D.3).

[202] Speaker *mns* patterns as expected in [+F+f] vs [-F+f], but unexpectedly increases F0 excursion in the (-F) trigger word in [-F-f] condition; in contrast speaker *meh* patterns as expected though to a lesser degree in [+F-f] vs. [-F-f], but unexpectedly increase F0 excursion in the -F trigger word in [-F+f] condition. The remaining male speaker (*miz*) has consistently higher F0 excursion in -F triggers than in +F triggers.

Figure 8.5 Mean *trigger* F0 excursion (xxn) by focus condition & by speaker.

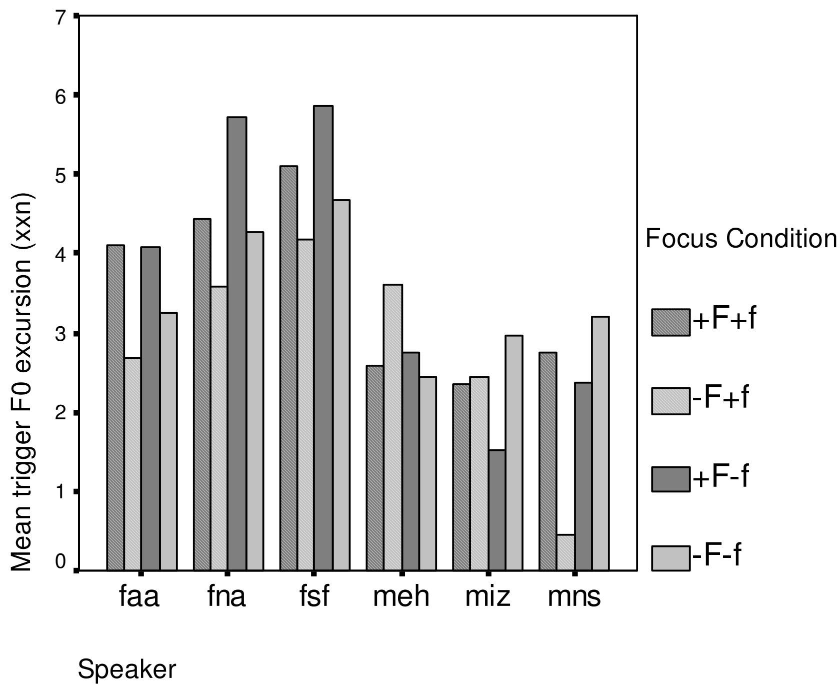
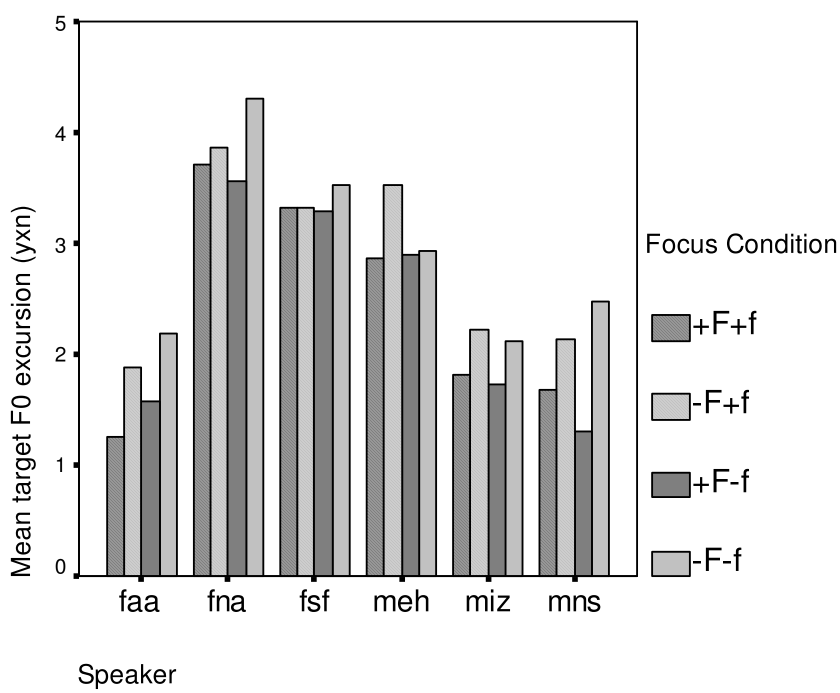


Figure 8.6 Mean *target* F0 excursion (yxn) by focus condition & by speaker.



Turning to F0 excursion in *target* words (*yxn*), this could be expected to vary either according to whether the target words follows a contrastive focus (*FOCUS*) or not, or according to the information focus (*focus*) status of the target itself (or to reflect both types of focus in some way). If the post-focal F0 compression effects reported in Norlin (1989) arise due to the fact of an item falling after a contrastive focus (±*FOCUS*), then F0 excursion in target words (yxn) will be *smaller* in +F conditions than in -F conditions. If however post-focal F0 compression effects in fact reflect the information focus (±*focus*) status of target words themselves then yxn will be *greater* in +f conditions than in -f conditions[203].

Figure 8.6 above shows mean values of *target* F0 excursion (yxn) by focus condition and by speaker. These data show more homogeneity in the general trend across all speakers than seen in trigger word F0 excursion, and the trend observed is in the direction expected if F0 compression marks the post-*FOCUS* status of the target, rather than the *focus* status of the target itself: mean target F0 excursion is generally smaller in +F conditions (indicating F0 compression) than in parallel -F conditions.

A oneway ANOVA (yxn by focus condition) shows that the differences in mean value of target F0 excursion between different focus conditions approach but do not reach significance ($p=0.073$; $\alpha = 0.05$). A less subtle but potentially revealing test is a two-way comparison of mean differences in F0 target excursion across each *type* of focus condition, rather than a fourway comparison across all focus conditions[204]. A pair of oneway ANOVAs (yxn by ±*FOCUS* status and yxn by ±*focus* status) reveals that the difference in target F0 excursion between grouped +F vs. -F conditions is highly significant ($p=0.009$), whereas target F0 excursion in grouped +f vs. -f conditions cannot be assumed to come from different populations ($p=0.898$).

In summary then, examination of F0 excursion in trigger and target words suggests quite strongly that in EA manipulation of pitch range is a reflex not of information focus (*focus*) but of contrastive focus (*FOCUS*), and further, that this is manifested both as expansion of pitch range on items bearing *FOCUS* (here, trigger words) and as compression of items which occur after a *FOCUS* (here, target words).

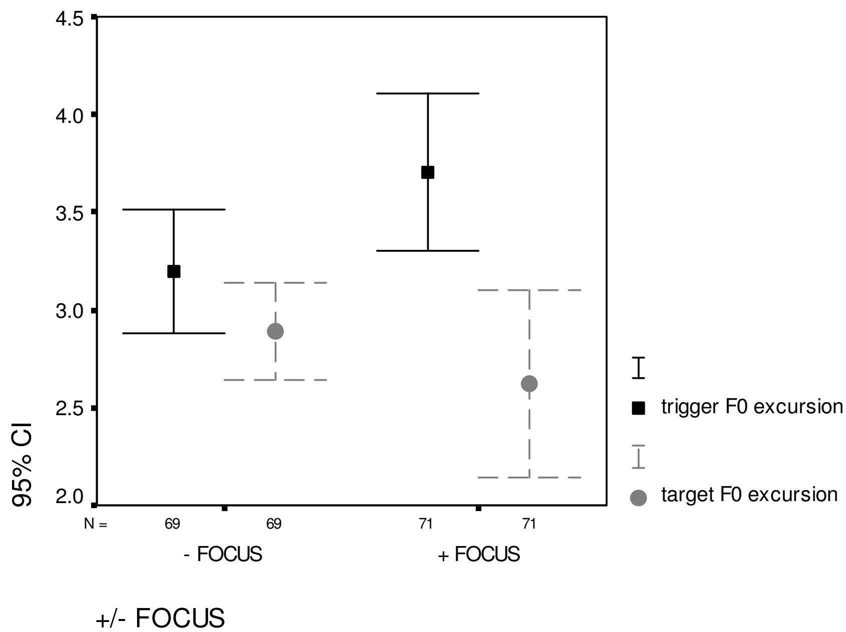
---

[203] Note that if there is some implicational relation between *focus* and *FOCUS*, whereby for example the *focus* status of a word becomes relevant only when it falls after a *FOCUS*, then one might particularly expect to see a difference between target F0 excursion in [+F-f] condition (a given target which occurs after a *FOCUS*) as compared to [-F-f] condition (a new target which occurs after a *FOCUS*).

[204] For example, to assess according to *FOCUS* status this involves grouping [+F+f] and [+F-f] values together as a '+F' set, for comparison with a '-F' set comprising [-F+f] and [-F-f] values. Similarly, to assess according to *focus* status [+F+f] and [-F+f] values are grouped together as a '+f' set, for comparison with a '-f' set comprising [+F-f] and [-F-f] values.

The results are summarised in Figure 8.7 below, which shows 95% confidence intervals around mean values of F0 excursion in trigger words (xxn, in the subject of the sentence, indicated with a solid line) and target words (yxn, in the object of the sentence, indicated with a dashed line), grouped by presence vs. absence of a contrastive focus in the sentence. F0 excursion in trigger words is larger when they bear contrastive focus (+F) than when they don't (-F), and F0 excursion in target words is smaller when they follow a contrastive focus (+F) than when they don't (-F).

Figure 8.7 95% confidence intervals around mean values of F0 excursion (xxn/yxn) in semitones, grouped by contrastive focus status of the subject.



The present findings are in line with Norlin's results, if we assume that his methodology elicited exhaustive or contrastive focus (*FOCUS*). The facts of F0 excursion in target words further confirm that compression of F0 excursion in target words is not a function of *focus* status, but purely of post-*FOCUS* position[205].

[205] The differences in mean target F0 excursion values are however too small to support a four-way distinction across focus conditions. This means that the present data do not reveal whether there is an implicational relationship between the two focus types.

8.3.3 [SVO+] results: gradient effects in syntactic focus sentences

The previous section has shown that there are clear gradient effects of in-situ contrastive focus (*FOCUS*) in EA (though not of information focus) in the form of manipulation of pitch range. These effects take the form of expansion of F0 excursion on items bearing in-situ *FOCUS* and compression of F0 excursion on items occurring after an in-situ *FOCUS*. This section explores F0 excursion properties in trigger and target words in sentences which employ syntactic *FOCUS* strategies available in EA, such as a pseudocleft or a negative continuation (or both). The results of the in-situ focus sentences have already undermined the hypothesis that gradient prosodic reflexes of focus might be contingent on a syntactic strategy (see 8.1.5 above); however, it is still plausible that there may be enhanced effects of contrastive focus with a syntactic focus, and perhaps evidence of some effects of information focus.

Since syntactic focus strategies would be infelicitous in -F contexts the only two contexts in which SVO+ sentences were elicited were [+F+f] and [+F-f]. The only comparison that can be made among focus conditions *within* the SVO+ dataset therefore are those in which information focus status (*focus*) is varied. Although there was no evidence of information focus effects on F0 excursion in the SVO data (see section 8.3.2 above), the SVO+ dataset is examined to see whether or not the 'stronger' syntactic focus strategy generates gradient effects of information focus, looking at F0 excursion in target words (yxn) within the SVO+ dataset. Comparison is also made across the full dataset of SVO and SVO+ sentences within each of the two +F conditions to see whether the effects of information focus are perhaps greater in SVO+ than SVO sentences.

Turning to contrastive focus effects in the SVO+ dataset, significant gradient contrastive focus (*FOCUS*) effects were detected in the SVO sentences, and it is possible to use this fact to determine whether there are any contrastive focus effects in the SVO+ sentences. If levels of F0 excursion in SVO+ sentences are similar to those in observed in SVO sentences in parallel +F contexts, then it is safe to assume that there are contrastive focus effects in the SVO+ sentences also. If levels of F0 excursion are significantly lower in SVO+ sentences than in SVO sentences then this could be interpreted as indication that the prosodic effects of syntactic *FOCUS* are somewhat reduced as compared to those of in-situ *FOCUS*. Some authors have argued that prosodic effects are used only to disambiguate potentially ambiguous syntactic or semantic structure (Cooper & Paccia-Cooper 1980, Straub 1997)[206]. On this line of argument the unambiguously emphatic syntactic *FOCUS* strategy (cleft or continuation or both) might reduce the need to mark FOCUS by prosodic means. Alternatively, if levels of F0 excursion are significantly greater in SVO+ sentences than in SVO sentences then this might indicate that the prosodic effects of syntactic *FOCUS* are greater than those of in-situ *FOCUS*. This scenario is expected if one of the conclusions of the pilot study reported in Hellmuth (2005) is correct, namely that de-accenting or extreme F0 compression in EA is to some extent contingent on use of syntactic focus strategies (cf. Frota 2000). This impression was gained from auditory transcription of a limited set of data from one speaker, and a major goal of collecting the SVO+ data for the present study is to confirm or clarify this earlier conclusion.

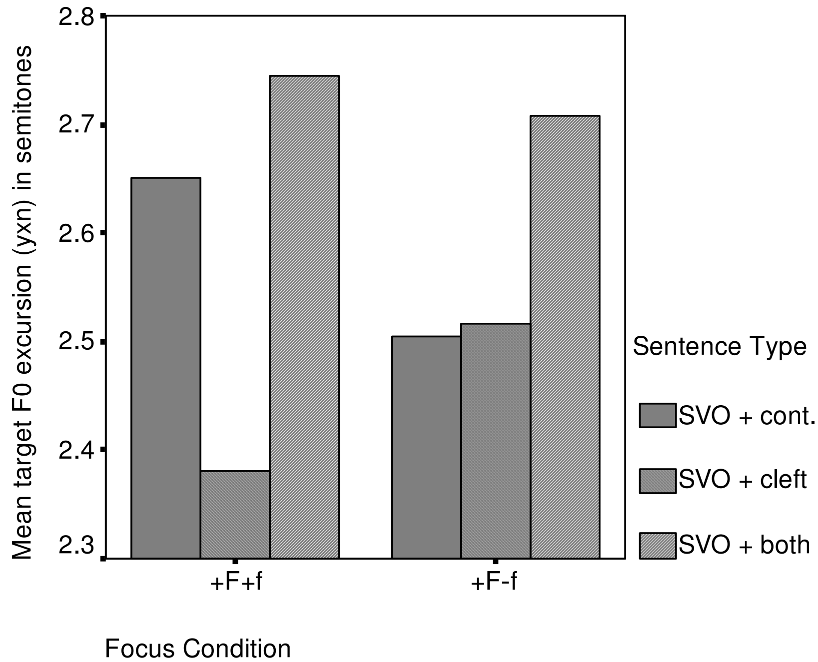
Comparison is thus made between SVO and SVO+ sentences, across sentence types and within each of the two +F focus conditions. The relevant variables to examine are F0 excursion in trigger and target words (looking for expansion in triggers (xxn) and/or compression in targets (yxn)).

8.3.3.1 Exploring the effects of information focus (*focus*) in the SVO+ sentences

Looking first at possible information focus (*focus*) effects within the SVO+ dataset only, differences in mean values of F0 excursion in target words (yxn), as shown in Figure 8.8, are very small (across a range of approx. 0.5 semitones). A pair of oneway ANOVAs (yxn by sentype and yxn by foccond; variances equal for both) confirm that the differences in mean values of yxn are not significant. This suggests that, as in the focus-in-situ SVO sentences, in the syntactic focus SVO+ sentences there are no significant variations in pitch range associated with the distinction between +f and -f information focus status.

[206] See Warren et al (1999) for experimental evidence that prosodic effects are used even in unambiguous discourse situations in semi-spontaneous speech.

Figure 8.8 Mean target F0 excursion (yxn) by focus condition & sentence type in SVO+ sentences.



Turning to comparison of possible information focus (*focus*) effects across all sentence types, these are examined within each +F condition in turn. Figure 8.9 below shows mean values in condition [+F+f] of F0 excursion in target words (yxn). These vary only very slightly across sentence types, and a oneway ANOVA confirms that the differences in mean values across different sentence types are not significant (Levene's test p=.701, ANOVA p =0.531). Within condition [+F-f], shown in Figure 8.10 below, again there are only slight differences in the mean values of yxn, and a oneway ANOVA again confirms that the differences in mean values across different sentence types are not significant (yxn by sentence type: Levene's test p=0.652, ANOVA p =0.708). The absence of significant differences in mean values of the relevant pitch range variables suggest that the absence of prosodic effects of *focus* status is consistent across the whole dataset, in both SVO and SVO+ sentences.

There are thus no apparent gradient effects of information focus (*focus*) in the SVO+ sentences. This is perhaps to be expected given the fact that information focus status does not have prosodic reflexes in the in-situ cases (the SVO set). The results from the SVO+ dataset thus strongly support the generalisation that pitch range manipulation is not used to mark information focus status in EA. In addition the results indicate that use of a syntactic focus strategy does not generate a ‘stronger’ context in which information focus effects become prosodically marked in EA.

Figure 8.9 Mean xxn and yxn in [+F+f] condition across all sentences.

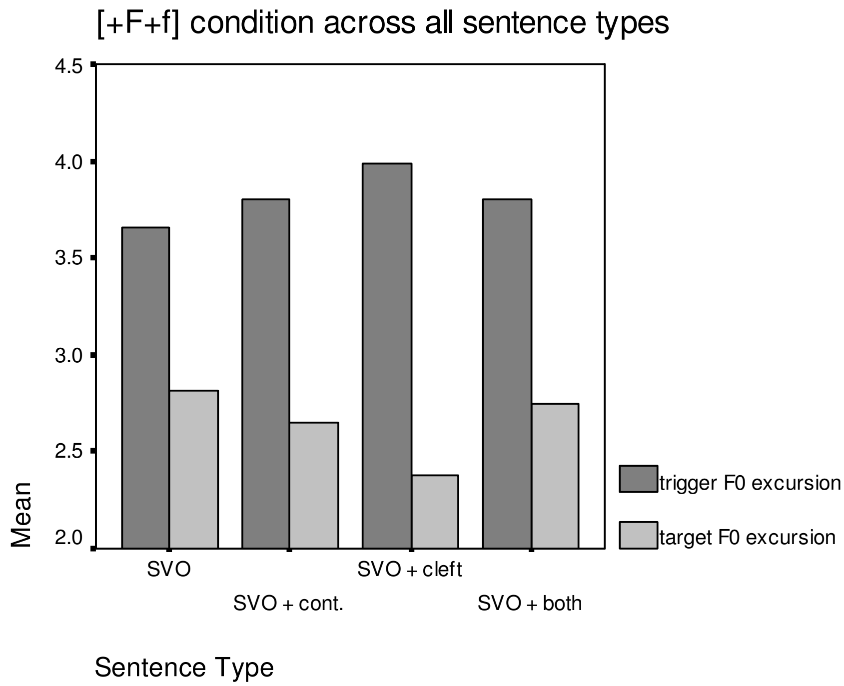
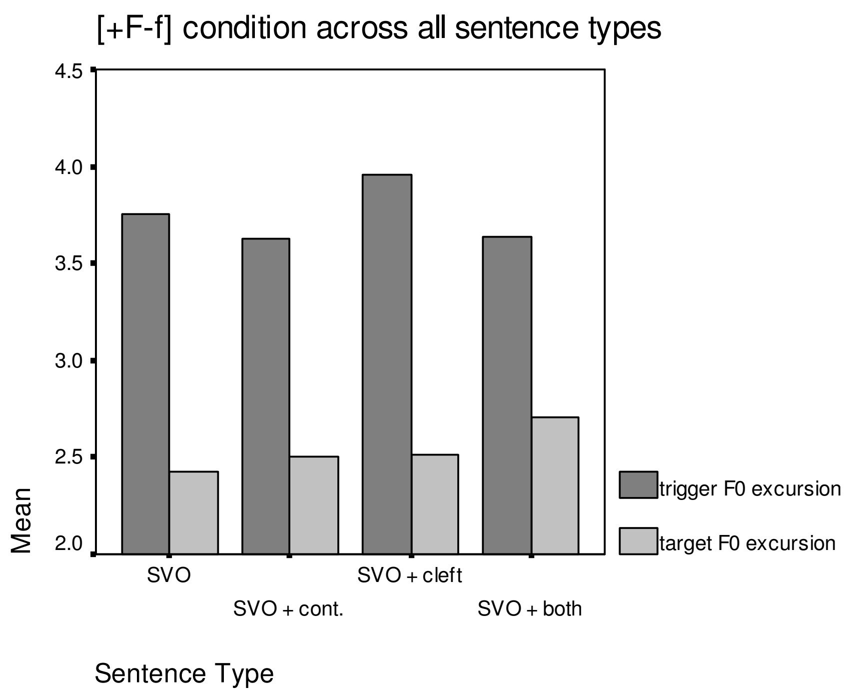


Figure 8.10 Mean xxn and yxn in [+F-f] condition across all sentences.



8.3.3.2 Exploring the effects of contrastive focus (*FOCUS*) in the SVO+ sentences

In order to investigate possible contrastive focus (*FOCUS*) effects in the SVO+ sentences, it is necessary to compare across the whole dataset (SVO and SVO+), within each of the two +F focus conditions in turn.

Figures 8.9 and 8.10 above show mean values of trigger F0 excursion (xxn) in each of the +F conditions, and these indicate that there is only small variation from one sentence type to another. A oneway ANOVA within each +F condition confirms that these variations are not significant (within [+F+f] condition: xxn by sentence type, Levene's test p=0.264, ANOVA p=0.893); within [+F-f] condition: xxn by sentence type, Levene's test p=0.146; ANOVA: p=0.891). The fact that there is no variation in F0 excursion levels, in parallel +F contexts, across the different sentences types, suggests that similar prosodic effects are associated in both in-situ and syntactic *FOCUS* strategies, and that pitch range manipulation is used in EA to mark *FOCUS*, whether expressed in-situ or by syntactic means.

This contradicts the tentative conclusions of the pilot study (Hellmuth 2005) in that the prosodic effects of *FOCUS* are not especially associated with syntactic *FOCUS* strategies. The impressionistic findings of that study were either not typical of EA in general, being perhaps over-emphatic or representing a borrowing from English[207]. Alternatively it is possible that the experimental paradigm of the present study, which uses read speech based on written prompts, has failed to elicit the full range of emphatic prosodic expression available in EA, and further investigation of the prosodic reflexes of syntactic *FOCUS* in EA might yield valuable additional information.

For the purposes of the present study however, the fact that the distinction in the prosodic reflexes of *focus* vs. *FOCUS* is consistent throughout the whole dataset of 144 SVO tokens plus 216 SVO+ tokens lends considerable support to the findings of section 8.3.2 (which were based on SVO tokens alone).

[207] The speaker who participated in this mini-pilot has an excellent command of English and therefore may have acquired some aspects of English prosody. These would not be expected to emerge in a monolingual target language experimental environment, but it was very difficult to elicit the focus scenarios and some code-switching between English and EA did occur during elicitation in an effort to disambiguate contexts.

8.3.4 Alignment and focus in EA

It has been argued that in some intonational languages, notably of the Romance family, the distinction between *FOCUS* and *focus* is expressible by means of a difference in pitch accent alignment. In most analyses this difference is thought to be categorical and distinct phonological representations are proposed for the two accents. For example in European Portuguese +F+f nuclear falls have an earlier peak than -F+f nuclear falls (Frota 2000); in Spanish, +F pre-nuclear rising accents have an earlier peak than their +f counterparts (Face 2002).

The investigation was restricted to SVO sentences only; from a potential 144 tokens (36 in each focus condition) four tokens had to be excluded due to disfluency on or near the trigger or target word, leaving 140 tokens for analysis. The key dependent variable for comparison across focus conditions in the current experiment is peak delay. This was retrieved from both trigger and target words (in milliseconds): trigger peak delay (XH-XC0) and target peak delay (YH-YC0). If there is an effect of *FOCUS* on peak alignment in EA pitch accents the values of peak delay are expected to vary significantly in trigger and/or target words between ±F conditions.

In order to determine whether or not such an alignment distinction is used in EA a small study was made of the alignment properties of the pitch contour to the segmental string in trigger and target words in the focus dataset. The trigger and target words in the focus dataset are listed in (8.20).

(8.20) Trigger and target words in the focus dataset.

	trigger		*target*	
lexset A	maama	'mum'	yunaani	'Greek'
lexset B	maama	'mum'	in-nounou	'the-baby'

As set out in section 8.2.2 above, the position of the H pitch turning point was identified by hand in each trigger and target (labelled XH and YX respectively) as well as the onset of the initial consonant of the stressed syllable in each trigger and target (X0 and Y0 respectively). The key variables for comparison across focus conditions are:

(8.21)	trigger peak delay	XH-X0
	target peak delay	YH-Y0

The labelling of Y0 (onset of the initial consonant of the stressed syllable of the target) was problematic in lexset B (in the word [in-nounou] 'the-baby'). The standard analysis of an assimilated definite article in Arabic is that it geminates to an initial sonorant consonant, and on this basis the Y0 label was placed at the mid-point of the geminate [n]. However it was observed during labelling that the L turning point on this word fell almost universally just after the onset of the [n] consonant, rather than at its mid-point. This throws some doubt on the analysis of the assimilated definite article as forming a geminate. If (as is shown in chapter 6) the L turning point of EA rising pitch accents aligns just after the onset of the stressed syllable, this suggests that the definite article does not form a geminate but instead is deleted[208]. In order to maximise the size of the dataset examined here (which was already small) onset of the initial consonant of the stressed syllable of the target was re-labelled in the lexset B tokens (N=72) at the start of the [n] segment in order to yield a realistic measure of peak delay in target words in these tokens.
Comparison of trigger peak delay values shows no significant differences across focus conditions at all (Tamhane's post-hoc test: non-significant). Recall that only the FOCUS status of trigger words varies, since all were elicited to be new (+f) in context. This results therefore suggests that there is no *FOCUS*-induced variation in pitch accent alignment on focussed items in EA.

In contrast, comparison of target peak delay across focus conditions does show variation in target peak delay values (ANOVA: F=6.029; p= 0.001). A post-hoc Tukey's HSD test divides the four focus-condition groups of target peak delay values into two homogenous subsets, with [+F+f] and [+F-f] grouped separately from [-F+f] and [+F-f]. F0 peaks are aligned earlier in target words falling after a +F than in those falling after a -F. This suggests that there is an indirect *FOCUS*-related effect on peak alignment in EA as a by-product of post-*FOCUS* F0 compression: smaller peaks are realised more quickly. There appears to be no effect of *focus* status on peak alignment in target words.

8.3.5 Summary of results

The results of the experimental investigation of the prosodic reflexes of FOCUS and focus in EA show that there is no categorical 'de-accenting' of words in any context, even if the target word is given in context (-*focus* status) and occurs after a contrastive

[208] This finding will not be pursued here further, except to note the potential benefit of establishing the intonational properties of Arabic dialects for resolving difficulties of segmental analysis.

focus (positioned after a +*FOCUS* trigger word). This latter context is thought to be the most conducive to de-accenting of words in Germanic languages such as English (cf. Selkirk 2000).

As regards gradient reflexes of focus in EA, the experimental results clearly indicate that pitch range manipulation is used in EA to enhance the contrast between a +*FOCUS* item, on which pitch range is expanded, and following items, on which pitch range is compressed. There appear to be no gradient prosodic reflexes of *focus* in EA. Both of these generalisations hold both in sentences containing *FOCUS*-in-situ and sentences containing a syntactic-*FOCUS* (expressed by means of a pseudocleft or a negative continuation, or both). These findings match those of Norlin (1989) and serve also to disambiguate them, by identifying which type of focus must have been at issue in that earlier study, namely a contrastive or exhaustive focus (that is, *FOCUS*).

Consistency of pitch accent alignment in +F vs. -F words indicates that EA does not use a choice of pitch accent type to mark focus distinctions. There are however indirect effects on the alignment of pitch peaks in words falling after a +*FOCUS* word, which have slightly earlier peak alignment than their counterparts occurring after a -*FOCUS* word.

The next two sections explore how best to understand the gradient reflexes of *FOCUS* in EA (section 8.4) and the lack of such reflexes in the case of *focus* (section 8.5).

8.4 Discussion: the prosodic reflexes of contrastive focus in EA

This study confirms that gradient pitch range manipulation is used in EA to mark items which bear contrastive focus (*FOCUS*). This is not altogether unusual since similar effects in other languages such as English have been known for some time. There has however been some debate as to whether such effects should be analysed as being a linguistic or paralinguistic effect, and if linguistic, whether categorical or gradient (Ladd 1994, Hayes 1994, Ladd 1996:269ff.).

8.4.1 Categorical vs. gradient linguistic use of extrinsic prosodic properties

Ladd (1996:269ff.) sets out a distinction between intrinsic and extrinsic phonetic factors which are relevant to the realisation of pitch. Intrinsic factors involve specification of a particular value on the pitch scale, such as H (a high tone) or L (a low tone), and these are generally, and uncontroversially, analysed as being part of the linguistic representation of pitch or tone in a language. Extrinsic factors involve modification of the scale itself, such as the difference in pitch range between young vs. old speakers, or between speech in a bored vs. enthusiastic mood. These factors are, again uncontroversially, generally agreed to be paralinguistic, and thus external to the phonological representation. Ladd argues that certain modifications of the pitch scale are however fully linguistic, and cites examples such as downstep on a H tone, which is linguistically conditioned, yet affects not only the tone itself but the range within which all subsequent H tones are realised. Such factors are thus extrinsic but linguistic.

Ladd goes on to suggest that the difference between intrinsic and extrinsic linguistic factors can be expressed by means of a difference of representation: intrinsic linguistic effects involve objects on the tonal tier, such as H and L tones; extrinsic linguistic effects involve “abstract relations between tones and between higher-level phonological constituents” (Ladd 1996:269ff.).

Extrinsic factors which involve modification of the pitch scale are thus open to analysis as being either linguistic (and part of the phonological representation, at the level of relations between constituents) or paralinguistic (and external to the phonological representation). It has also been widely assumed that the defining feature of linguistic factors is that they are categorical are gradient, which in turn has lead to the assumption that all gradient effects must be paralinguistic (and vice versa).

One phenomenon which lay at the centre of this debate in the early 90s, and which is relevant to our present purposes, is the question of how best to analyse increased F0 excursion on certain pitch accents in English. These instances of pitch range expansion are consistently perceived by listeners as a reflex of increased emphasis or contrastive focus.

The example in (8.22) comes from Ladd (1996:281). (8.22a) is a plain imperative which could be a felicitous answer to a request for instructions as to how to dispose of some item; in (8.22b) increased F0 excursion on the H* pitch accent on the word ‘Mary’ (indicated by the upward pointing arrow [↑]) generates either a contrastive interpretation (‘Give it to Mary, not Anna.’) or perhaps impatience at having to repeat instructions previously given.

(8.22) a. H* H* L-L% b. H* ↑H* L-L%

Give it to Mary. Give it to Mary.

Ladd (1994) proposed analysis of this property of English intonation by means of a categorical feature, [raised peak], or a separate 'H^{+}' tone, to reflect the fact that in these cases the manipulation of pitch range generates a consistent linguistically distinct interpretation. Hayes (1994) argues against a categorical raised peak feature however, noting that the classification 'H+' or [raised peak] is too narrow. Emphasis may be located not only on a word bearing a H* pitch accent but on a word bearing any tonal event or events. Whatever the tonal specification of the word, under emphasis its properties are enhanced, though with the proviso that phonological relevant distinctions are preserved (such as the distinction between a word bearing an emphasised !H* and a preceding H*). He characterises this notion as 'gestural reinforcement'.

This idea is similar to the notion of 'hyperarticulation' of focussed items (based on proposals made by Lindblom 1990) which has been adopted by Chahal (2001) in her analysis of focus-related pitch range manipulation in Lebanese Arabic (LA), and by DeJong and Zawaydeh (2002) in a study on (narrow) focus effects on vowel duration and F1 values in Jordanian Arabic (JA). In the latter study the authors found that whilst word-stress effects (increased duration and F1 values) were comparable across all speakers, focus effects varied across different speakers. They suggest that this indicates that focus effects are not conventionalised and only implemented when needed (de Jong & Zawaydeh 2002:72):

> "Speakers are aware to some extent of how particular contrasts are expressed and can enhance them specifically in a way which might not be what other speakers do".

There seems to be good evidence to support the notion that focus effects are linguistic but 'optional' therefore. The outstanding question in this debate is thus, in Ladd's terms, "whether the distinction between normal and 'gesturally reinforced' is categorical" (Ladd 1994:59).

A key question to answer then, regarding the use of pitch range manipulation to mark FOCUS in EA is whether this is a linguistically categorical or gradient phenomenon.

Crucially, we need to know whether or not use of pitch range manipulation to mark *FOCUS* should be analysed as being part of the phonological grammar of EA.

8.4.2 Pitch range manipulation in EA: is it part of the grammar?

Three potential diagnostics exist in order to test whether EA pitch range manipulation is categorical (that is, conventionalised).

A first possible diagnostic would be to establish the directionality of pitch range compression before and/or after a *FOCUS* in EA. If compression is unidirectional it would be amenable to analysis as part of the phonological grammar, within a theory such as Focus Prominence Theory (Truckenbrodt 1995, Truckenbrodt 1999, Selkirk 2004a), in which focus effects are analysed by means of a constraint requiring that a focus be either left- or right- adjacent to a prosodic boundary[209]. Crucially in this analysis focus effects are predicted to be uni-directional, either leftward or rightward, but not both.

Norlin (1989) indicates clearly that in his study the effects of *FOCUS* were rightward only: pitch range compression affected all words following the *FOCUS*, but words preceding the *FOCUS* were produced in the same pitch range as their counterparts in a plain declarative. Unfortunately however it is not possible to test the directionality of pitch range compression in the current dataset, since the trigger word, in which *FOCUS* status was varied, was always sentence initial (the subject in an SVO sentence). However this is something that could be tested in future using a double-object sentences (varying *FOCUS* status of the direct object).

A second possible diagnostic, suggested by Hayes (1994:66), would be to carry out a categorical perception test to see whether listeners perceive a categorical distinction across the range of gradient pitch range variability. A test of this kind was carried out by Ladd & Morton (1997) for English with complex results: they found no evidence of the kind of classic 'S-shaped' curve which might indicate that listeners are able to discriminate 'normal' pitch range from or 'emphatic' pitch range; instead listeners were able to discern even quite small distinctions in pitch range, between utterances from any two positions along the continuum. Nonetheless there *was* evidence to suggest that

[209] This can be achieved by either insertion of a prosodic boundary or deletion of intermediate prosodic boundaries between a focus and one edge of the utterance, with the latter resulting in de-accenting.

listeners are able to classify utterances as either normal or emphatic. Ladd & Morton suggest that whilst listeners do not categorically *perceive* pitch range distinctions, they do however categorically *interpret* them (Ladd & Morton 1997:339):

> "listeners are predisposed to interpret accents or utterances as being categorically either 'normal' or 'emphatic'. A variety of acoustic and pragmatic parameters play a role in this decision... [they] 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."

A categorial perception test of this kind in EA unfortunately lies beyond the scope of the present thesis however.

A third and final potential means of disambiguating between non-linguistic and linguistic use of pitch range manipulation arises from the observations made by DeJong and Zawaydeh (2002), who interpreted variation in the reflexes of focus across speakers as an indication that the prosodic reflexes of focus are not conventionalised (compared to the reflexes, say, of word-level prominence).

Two Egyptian speakers who helped with the design of the focus experiment, one male (*maa*) and one female (*fnf*), asked to be given the opportunity to record the production stimuli, to provide a 'properly representative sample of EA speech'. These recordings were not included in the main study reported in this chapter, since the participants were not naïve as to the purpose of the task, but were nonetheless analysed alongside the other data. The male speaker (*maa*) produced a small number of tokens which did show full de-accenting of target words; however, he produced these somewhat at random, as often after a *-FOCUS* trigger as after a *+FOCUS* trigger, and they probably reflect a highly stylised mode of speech (at best), possibly influenced by his knowledge of English.

Of more interest are the utterances from the female additional speaker (*fnf*), whose tokens show a greater degree of pitch range expansion/compression than observed in the main study, but in *exactly the same* contexts, and without any instances of full de-accenting. Whilst these results are only anecdotal, the speaker *fnf* was able to enhance the degree of focus differentiation in her speech, yet maintained the same basic phonological properties, which seems to be consistent with a 'gestural reinforcement' view of pitch range manipulation, and with the DeJong & Zawaydeh's suggestion that that speakers may vary in their implementation of FOCUS because its reflexes are not fully conventionalised.

To summarise then, Norlin's results suggest that gradient *FOCUS* effects might be conventionalised, in that they appear to be directional, and thus consistent with a phonological analysis, bt we are unable to corroborate this finding in the current dataset. On the other hand, the fact that speakers are able to vary the degree to which they implement pitch range manipulation as a reflex of *FOCUS* suggest that its implementation is not fully conventionalised. The balance of evidence that can be obtained from the present dataset falls slightly on the side of analysis of EA pitch range manipulation as gradient, and probably external to the grammar.

8.4.3 Conclusion: gradient effects of contrastive focus in EA

The interim conclusion I propose is therefore that use of pitch range manipulation in EA to express contrastive focus (*FOCUS*) is gradient, and under the control of speakers, that is, not automatic, and thus probably not fully conventionalised within the phonological grammar.

Supporting evidence comes firstly from within the dataset investigated in this chapter. The core finding, that *FOCUS* induces not only F0 expansion on focussed items but also F0 compression on post-*FOCUS* items, is consistent with a 'gestural reinforcement' interpretation of gradient focus-related pitch range manipulation in EA, in which articulatory means are used to enhance the overall distinction between +F and -F items (Hayes 1994).

In addition, the focus-related alignment facts of EA reported in section 8.3.4 reveal effects on peak alignment in post-*FOCUS* items only, whilst peak alignment in +F words themselves is apparently unaffected. This is also consistent with a 'gestural reinforcement' view: the phonologically relevant alignment properties of +F items are preserved, whilst those of post-*FOCUS* items are less accurately conveyed. There is also supporting evidence from outside the dataset in that the results here parallel the focus effects observed in Lebanese Arabic (Chahal 2001), which have been analysed as the results of hyperarticulation (and were shown to be bi-directional).

Additional evidence related to EA itself, comes from a study of the semantics of metalinguistic negation in EA by Mughazy (2003), in which he describes salient but optional use of what he terms 'contrastive intonation' in specific contexts. The distinction between truth-functional negation and metalinguistic negation is not ambiguous in EA since two different forms of negation are used[210], and in the case of metalinguistic negation the negated item attracts a *FOCUS* (which Mughazy terms the 'main stress' in the utterance). An example is provided in (8.23) below, in which speaker B is correcting speaker A's use of low colloquial [marsaH] 'theatre', instead of more standard [masraH] 'theatre' (with 'main stress' indicated in italics):

(8.23)	A:	imbaareH	?ana	ruHt	el-marsaH			
		yesterday	I	went	the-theatre			
		'Yesterday, I went to the theatre.'						
	B:	?inta	meš	ruHt	el-*marsaH*	?inta	ruHt	el-masraH
		you	NEG	went	the-theatre	you	went	the-theatre
		'You didn't go the *theatre*, you went to the theatre.'						

What is of interest for our present purposes is the fact that Mughazy notes that speakers can create ambiguity in a sentence like B for ironic effect by failing to implement the contrastive intonation pattern, a strategy which Mughazy notes has also been observed in English (Chapman 1996). The fact that speakers can apply or not apply 'contrastive intonation' in this way suggests that the pattern of use of F0 expansion in EA is similar to that observed in English, which has been argued by many authors to be gradient, and that pitch range manipulation is under the pragmatic control of EA speakers.

In summary then, the conclusion of this discussion is that use of pitch range manipulation to mark *FOCUS* in EA is gradient but under the control of speakers. However further investigation is needed: production data with a sentence medial *FOCUS*, to establish the directionality of *FOCUS*-related pitch range compression, and a differentiation-based perception test to establish whether EA speakers are able to categorically perceive and/or interpret changes in pitch range.

[210] Use of the continuous form [miš] of the (usually discontinuous) negation marker [ma- -š] indicates that a metalinguistic negation interpretation is intended.

8.5 Discussion: the prosodic reflexes of information focus in EA

The results of the focus experiment reported above suggest that there is no prosodic reflex of the given/new distinction (*focus*) in EA. This section discusses the plausibility of this finding before turning to its typological implications, and potential explanations which arise from syntactic and semantic properties of EA. These facts are then related back to the claim made in chapter 6 that rich pitch accent distribution in EA is a phonological phenomenon.

8.5.1 Explaining the lack of prosodic reflexes of information focus in EA.

The focus experiment results suggest that there is no prosodic reflex of *focus* in EA. Similar findings have been observed for Italian and other Romance languages (Ladd 1996, Swerts et al 2002).

Interestingly however there is good evidence of other, non-prosodic, reflexes of *focus* in the Romance languages, which often involve word order changes. Ladd (1996:179) gives an example in Italian in which the given item is moved to the end of the sentence, a position in which it receives a 'tag' pronunciation and is produced in a low pitch range. The sentence in (8.24a) is addressed to a child whose sibling has just had their evening bath, and thus the word "bagnetto" ('bath') is given in context; in English the utterance in (8.24b) would be felicitous in this context; the equivalent in Italian, in which "bagnetto" is de-accented, is less acceptable (Ladd 1996:179 examples 5.43-45; capitals indicate the contrastively focussed item)[211].

(8.24) Right-dislocation of given (-*focus*) item in Italian

a.

Adesso	faccio	scorrere	il	TUO,	di	bagnetto.
now	I make	run	the	yours,	of	bath.dim

b. "Now I'll run YOUR bath"

c. ??Adesso faccio scorrere il TUO bagnetto.

[211] Note that this is an instance of de-accenting of a -*focus* item in post-FOCUS position. It could therefore in principle be the need to assign main prominence to the +*FOCUS* item "tuo" (your) which conditions right-dislocation, rather than the need to express the -*focus* status of the given item "bagnetto" (bath). Vallduví (1991) points out a distinction between plastic and non-plastic languages, with the position of accents in a phrase fixed in the latter, so that changes to word order are instead used to shift constituents into sentence locations where they will appear with or without accent as needed. I take Vallduví's concept of accent here to indicate the fixed position of *main prominence* in the sentence in a particular language. This Italian example could be analysed as an attempt to move "tuo" to phrase-final position where it will receive main prominence (cf. Zubizaretta 1998, Frascarelli 2000).

The question arises therefore whether EA also has non-prosodic strategies for expressing the information focus. The next section explores a syntactic strategy that has been reported for marking the given/new distinction in EA.

8.5.2 Grammaticalisation of information focus in EA

Jelinek (2002) argues that expression of the given/new distinction is at least partly grammaticalised in EA. She identifies two syntactic strategies available to express the *focus* distinction, which she expresses in terms of 'backgrounding of arguments'. The two strategies are the possibility of having a null subject (known as 'pro-drop') and the option to incorporate an object pronoun into the verb (which Jelinek describes as 'object cliticisation'[212]). In EA definite subjects can be 'pro-dropped' and are in this way "maximally backgrounded in the discourse" (Jelinek 2002:71); object clitics have a similarly backgrounded status.

In EA, an overt subject pronoun (a deictic pronoun) always introduces a discourse element which Jelinek describes as "new or contrastive in the context", which could equate to either +F or +f in the terms used in this chapter. She notes that overt deictic pronouns usually receive "added stress or a higher intonation peak" (Jelinek 2002:94). In (8.25a) below, the subject pronoun [hiyya] 'she' is optional, whereas in (8.25b), in which the pronoun bears contrastive focus the pronoun is obligatory (Jelinek 2002:94; 'higher intonation' indicated in capitals).

(8.25)	a.	(hiyya)	waSalit		
		(she)	arrive-perf.3fs		
		'She arrived.'			
	b.	HIYYA	waSalit	muš	huwwa
		she	arrive-perf.3fs	not	he
		'It was SHE who arrived, not he.'[213]			
	c.	*waSalit,	muš	huwwa	
		arrive-perf.3fs	not	he	

[212] Other authors argue that object pronouns in Arabic are not in fact clitics but are fully incorporated into the verbal complex (Shlonsky 1997).

[213] Jelinek translates this using a cleft in English, although there is no cleft or pseudocleft construction in the EA rendition.

A 'dropped' subject (a discourse anaphor, coreferent with some antecedent in the discourse) will always be "old information that is topical and maximally backgrounded" (ibid.). This latter definition parallels the criteria used here to define a *-focus* item, which is either textually or situationally given. From a syntactic point of view, the null subject in EA is licensed by person subject agreement, whilst an object clitic is argued to raise with the verb and thus fall within the left periphery of the sentence which is associated with presupposed material (see Jelinek 2002 for details).

Pro-drop and object cliticisation are however not available in all contexts. Jelinek demonstrates that the subject of a nominal sentence cannot be pro-dropped in EA[214], and that the object of a pseudoverb cannot be cliticised[215]. The subject of a sentence with an overt verb can be dropped as in (8.26b) below, but the subject of a nominal sentence cannot be dropped (as in 8.27b) (Jelinek 2002:72).

(8.26) a. 9ali fataH il-baab
Ali opened the-door
'Ali opened the door.'

b. fataH il-baab
opened the-door
'He opened the door.'

(8.27) a. il-baab maftuuH
the-door open (m.s.)
'The door is open.'

b. *maftuuH
'It is open.'

The availability of a syntactic strategy for expressing *focus*, could explain the absence of a prosodic strategy in EA (neither categorical, nor gradient). In addition one could argue further that the availability of an alternative *focus* marking strategy is causally related to the lack of prosodic marking of the concept. If so this potentially presents a problem for the analysis of EA rich pitch accent distribution advocated in this thesis: namely that it is a purely phonological phenomenon, arising due to a constraint in the phonological part of the grammar. If the key factor conditioning de-accenting in

[214] Analysed as due to the differing argument structure of a nominal sentence, with no Voice projection.
[215] 'Pseudoverbs' for Jelinek comprise possessive prepositional phrases such as [9ind-i] 'with-me', and 'psych noun' constructions such as [nifs-i] 'wish-my' and [?aSd-i] 'intent-my'. Both of these bear person subject agreement via the possessive suffix.

Germanic languages is *-focus* status[216], then one could argue that the reason why PWds are not de-accented in EA is because it is a language in which *-focus* status can be conveyed in another way (by dropping the word altogether).

This argument however fails to take into account the fact that the distribution of pitch accents in Germanic languages varies in neutral contexts also (such as in fast or informal speech), but does not vary in such contexts in EA (as demonstrated in chapter 3). Nonetheless the fact that not all subjects may be pro-dropped in EA enables us to test whether or not there is a link of some sort between the lack of de-accenting in EA and the availability of a syntactic strategy. Specifically, if there is such a link, then we would expect the subject of a nominal sentence that is given in context, and thus has *-focus* status, to be deaccented. The 'target' words in the focus experiment, in which *focus* status was varied, were all in object position in an SVO sentence, and so could optionally have been cliticised. However a survey of the entire thesis corpus reveals a small number of nominal sentences, in which the subject cannot be pro-dropped, and in which the subject is given in context.

One such example, from the spontaneous telephone conversation (in the LDC section of the corpus) is illustrated in (8.28) below. It occurs in a section of the conversation where the two interlocutors are discussing the pros and cons of living in a first floor apartment. In (8.28a) the first mention is made of [id-door il-?awwal] 'the first floor', indicated with a dotted line. In the later sentence, in (8.28b) reference is made to the first floor again, simply as [il-?awwal] 'the-first', indicated with a plain line. This second instance of [?awwal] is thus textually given in context, and is also the subject of the nominal sentence [il-?awwal 9ilwi] 'the first is great', and, crucially, is accented by the speaker. This suggests that even when a word is obligatorily overt, yet is given in context, it is accented in EA.

[216] As we have seen Ladd (1996:chapter 5) argues that this is probably too simple a generalisation.

(8.28) Nominal sentence (from LDC corpus) with accented *-focus* subject.

a.	B:	ana	ba?ullik	id-door	il-?awwal	taHtii	dakakiin	9alya	giddan
		I	say-to-you	the-floor	first	under-me	shops	high	very
		>	LH*	LH*	LH*	LH*	LH*	LH*	LH*
	A:	Mmm							
		LH*L-L%							
	B:	ya9ni	taHtii	sitta	dakakiin	taHt	il-9imaara		
		I-mean	under-me	six	shops	under	the-building		
		>	LH*	LH*	LH*	<	LH*H-H%		
	A:	Mmm							
		LH*L-L%							
b.	B:	il-?awwal	9ilwi	foo?	il-maHillaat				
		the-first	high	above	the-shops				
		LH*	LH*	LH*	LH* L-L%				

Gloss B: I tell you, on the first floor I have really great shops downstairs.
A: Hmm.
B: Well, I have six shops downstairs in the building.
A: Hmm.
B: The first (floor) is great, above the shops.

Future experimental investigation could test for categorical reflexes of *focus* in a larger corpus, and also whether gradient prosodic reflexes emerge in EA in contexts where syntactic backgrounding of arguments is not possible. Nonetheless, the fact that such a word is observed not to be de-accented yields further support for the notion that rich pitch accent distribution is a phonological phenomenon, fully independent of other aspects of the grammar. As such it is appropriate to see density of pitch accent distribution as an independent parameter of prosodic variation, and one that is amenable to analysis within the phonological component of the grammar (such as by means of Tone-Prominence Theory).

8.6 Conclusion

This chapter set out the rationale, methodology and results of an experimental investigation into the prosodic reflexes of two types of focus in EA. The distinction between *focus* (information focus) and *FOCUS* (contrastive focus) was implemented here because the distinction has been shown to be valid from syntactic evidence in MSA (Moutouakil 1989).

The results of the investigation reproduced the findings of Norlin (1989) and thus confirmed that whilst there are no categorical reflexes of either contrastive focus or information focus in EA (in the form of de-accenting), there are gradient reflexes of contrastive focus (*FOCUS*), in the form of pitch range manipulation: F0 excursion is expanded on the +*FOCUS* item, and F0 excursion is compressed on following items.

In contrast there were found to be neither categorical nor gradient effects of *focus* status on F0 excursion. An analysis of peak alignment in words bearing contrastive focus and information focus indicated similarly that peak alignment is unaffected by information focus status. Nor was there any effect on peak alignment in +*FOCUS* words; the only effect was a slight leftward shift in the peaks of word falling after a contrastive focus, which was analysed as a by-product of post-*FOCUS* F0 compression.

The gradient effects of contrastive focus on pitch range manipulation in EA were discussed in the light of arguments surrounding similar phenomena in English. On the basis of the available evidence it was suggested that use of pitch range manipulation in EA to express *FOCUS* is a gradient effect, under the control of speakers.

The apparent lack of any prosodic reflex of information focus *(focus)* in EA was discussed in the light of the syntactic properties of the language. Even though EA has a syntactic strategy to grammaticalise the ±*focus* distinction (pro-drop and object cliticisation,: Jelinek 2002), there is evidence in the thesis corpus to suggest that even in contexts when the syntactic strategy cannot be implemented due to semantic constraints, a -*focus* word is accented in EA.

Crucially, this chapter presents a final layer of evidence in support of the proposal that density of pitch accent distribution is a fully independent parameter of prosodic variation, shown here to be independent of the syntactic and semantic components of the grammar which are involved in the expression of focus.

The final chapter of the thesis (chapter 9), which follows, provides a summary of all of the evidence set out in favour of density of pitch accent distribution as an independent parameter of prosodic variation, and briefly explores the implications and potential of the proposal for future research.

9 Pitch accent distribution and the typology of prosodic variation

This final chapter provides a summary of all of the evidence set out in favour of the proposal that density of pitch accent distribution is an independent parameter of prosodic variation. Section 9.1 provides a summary of the thesis, setting out the facts EA pitch accent distribution, and the overall analysis proposed to account for those facts. In section 9.2 the main contributions of the thesis are highlighted as well as some practical implications of the thesis findings. Section 9.3 suggests potentially fruitful avenues of future research which arise as a result of the present study.

9.1 Summary of the thesis

This thesis adds EA to the range of languages for which prosodic theory must account, by increasing our knowledge of EA sentence phonology. Specifically, distributional and experimental evidence support classification of EA as a stress-accent language in which pitch accent distribution is sufficiently different from that reported in other stress accent languages as to require explanation. A new typological category is required to describe EA and other languages like it, and a more finely grained articulation of the grammatical relationship between phonological tone and prosodic structure is proposed.

Chapter 3 demonstrated the central claim of this thesis empirically. EA was shown from a corpus survey across a variety of speech styles to have very rich pitch accent distribution, with a pitch accent occurring on every content word. In addition, EA also has the property of marking each accented word with the same pitch accent *type*. This was demonstrated in a detailed survey of pitch movements localised around stressed syllables, as well as a formal model of EA intonation which proposes a single default pitch accent in the EA pitch accent inventory. The correlation between rich pitch accent distribution and use of a single pitch accent type appears not to be unique to EA, and it is therefore a useful testing ground for Jun's (2005b) suggestion that in such languages pitch may be used as a cue at the word level.

Chapter 4 therefore explored the nature of word-level prominence in EA. The widely-held assumption that EA is a stress-accent language in which pitch marks the stressed syllable of words was borne out by two experimental studies. An additional parameter of prosodic variation was proposed whereby languages may vary as to which *domain* is relevant for the realisation of pitch, regardless of the *function* of pitch in that language. Thus, among languages in which the function of pitch is lexical, we see variation in the domain within which pitch is realised: in Japanese only one lexical accent per phrase is realised. In addition we see variation among languages in which pitch is purely postlexical in the distribution of intonational pitch accents.

Chapter 5 argued for a prosodic-structure-based conception of pitch accent distribution, with the domain of pitch accent assignment varying across constituents of the Prosodic Hierarchy. Empirical evidence from prosodic phrasing in complex EA sentences indicates that EA prosodic phrases are consistently composed of more than one Prosodic Word (PWd), and thus the domain of pitch accent distribution in EA is not a phrase level constituent. Evidence from accentuation of content and function words in the corpus indicates that the correct generalisation to describe EA rich pitch accent distribution is that the domain of pitch accent distribution in EA is the PWd.

Chapter 6 explored this empirical claim in the context of two facts about the relationship between phonological tone and prosodic prominence: firstly that it is known to be a two-way relation, and secondly that the prosodic constituent whose head attracts tone may vary. The properties of a particular conception of tone-prominence relations were described - tone↔prominence theory - in which surface relations between tone and prosodic prominence result from the interaction of a pair of inherently-ranked fixed hierarchies of markedness constraints which regulate association of tone to prosodic prominence, and of prosodic prominence to tone, respectively. A formal analysis reveals that in EA it is better to insert tones than to leave PWds unaccented (PWD→T >>DEP_{TONE}) and it is better for a lexical word to lose its PWd status and go unstressed than to be a PWd and be realised without an accent (PWD→T >>LEXWD:PWD). The stringency relation between individual P→T constraints at different levels of the hierarchy was confirmed, supporting the view that T↔P constraints are in a fixed ranking.

Chapters 7 and 8 provide further empirical evidence in support of the claim that the distribution of EA pitch accents is due to a purely phonological constraint. Chapter 7 uses pitch accent alignment properties to confirm that the association properties of tone to prominence (that is, the TBU) are indeed independent of the attraction of prominence to tone (that is, the domain of pitch accent distribution). Chapter 8 showed that pitch accent distribution in EA is truly independent of information structure.

In addition, chapter 7 established the descriptive facts of EA pitch accent alignment and proposes a phonological representation of the EA pre-nuclear pitch accent: L+H*, whilst chapter 8 additionally reproduced and clarified the findings of Norlin (1989) and thus confirmed that whilst there are no categorical reflexes of either contrastive focus or information focus in EA (in the form of de-accenting), there are gradient reflexes of contrastive focus in the form of pitch range manipulation.

9.2 Contributions of the thesis

The main contribution of the thesis is the proposal of an additional parameter of prosodic variation, namely that the *size of domain* within which tone and pitch accents are distributed is independent of the *function* of that tone. This is formalised as the claim that EA pitch accent distribution is due to the effects of a purely phonological constraint on the relations between prosodic prominence and phonological tone. Thus pitch accent distribution, as a parameter of prosodic variation, is predicted to be, and is found to be, independent of other aspects of the grammar.

Formulation of the parameter in terms of two-way tone↔prominence relations, has typological advantages. Firstly it matches the observation that “postlexical prosodic pitch properties cannot be predicted from lexical pitch properties” (Jun 2005b:432). In addition the reversed direction of the fixed ranking of the two hierarchies of tone↔prominence constraints encodes the observation that the unmarked role of lexical tone is paradigmatic, distinguishing individual small units of prosodic structure; in contrast the unmarked role of postlexical tone is syntagmatic, highlighting part of a larger sized prosodic constituent (Hyman 2001).

The formal analysis also provides an explanation for the fact that languages like EA which display rich pitch accent distribution also happen to have the property of having a small pitch accent inventory, a correlation also noted in Spanish, Greek and Italian (Jun 2005b, Selkirk 2005b). If pitch accents in EA are inserted as a default tone to fulfil a prosodic requirement that every PWd bear tone, it is perhaps to be expected that the ‘epenthetic’ tone inserted should not vary greatly, by analogy with the fact that epenthetic segments do not vary greatly in segmental phonology. The distinct realisation of pitch accents in final and non-final positions in these languages is also potentially explained in terms of tone insertion vs. tone-spreading.

9.3 Future investigation

In this section I briefly suggest some potentially fruitful areas of future investigation highlighted by the thesis.

An important next research goal would be to extend the database of languages which accent every word, in order to test both the typological and theoretical claims of the thesis and also in order to determine the origin of the phenomenon. My hypothesis is that rich pitch accent distribution may arise for a variety of reasons. In the case of Egyptian Arabic (EA) I would like to explore the possibility that EA was influenced by the accentual and tonal properties of Ancient Greek, which is known to be a pitch accent language (Sauzet 1989, Steriade 1988, Hayes 1995). Since there was a historical Greek presence in Tunisia and northern regions of Egypt (including Lower Egypt, as far as Cairo)[217], this hypothesis could be tested by investigation of the pitch accent distribution properties of southern dialects of Egypt (Sa9iidi).

A second future research goal would be to explore the interrelation among rich pitch accent distribution, the availability of prosodic reflexes of information focus and the availability of syntactic focus strategies. The facts of EA in chapter 8 suggest that the lack of information focus marking in EA are more likely to be a result of the phonological requirement to mark every PWd with tone, than a result of the availability of an alternative syntactic strategy. However this question requires further investigation in a wider range of EA data.

Regarding the particular properties of EA intonation, there are questions that arise from the findings of the thesis. For example, it would be good to clarify the patterns of behaviour in phrase-final (nuclear) pitch accents, in order to determine whether they are best described as re-aligned instances of the ubiquitous pre-nuclear accent, or as a distinct pitch accent type. In addition the correlates of prosodic phrasing above the level of the word require further investigation, and in particular whether or not there are reflexes of phrasal prominence in neutral contexts, and in non-neutral sentences of different types (such as double-object constructions or negative sentences).

[217] Tunisian Arabic almost certainly has the same pitch accent distribution patterns as those observed in the (Cairene) EA data described here (p.c. Nadia Bouchhioua)

Finally, there is much to be gained from cross-dialectal comparison with other spoken dialects of Arabic. In addition to the descriptive gains to be had from adding further dialects and languages to the database of prosodic data for which intonational theory must account, the link between the availability of syntactic information structure strategies and the lack of prosodic strategies would be very easy to test - since Arabic dialects are known to be more similar in their syntactic properties than in their phonological properties (Brustad 2004). For example, the dialect of Arabic spoken in the Old City of San'aa (Yemen) allows null subjects and object cliticisation as does EA (Watson 1993), but a brief survey of its prosody reveals that it displays both sparse pitch accent distribution and instances of de-accenting[218].

[218] Thanks for Janet Watson for sharing excerpts from her San'aani data with me.

Appendix A (relates to chapter 3)

(A.1) Transcription results for align sentence 111101: [Zahar namaš 9ala gism l-walad wa-9arifna ?innul HaSba]

SPKR	*error*	zaHar	namaš	9ala	gism	il	walad	wa	9arifna	innu	-l-	Hasaba	
faa1		LH*	LH*	---	LH*	<	LH*	<	LH*	LH*	<	!LH*	L-L%
faa2		LH*	LH*	<	LH*	<	LH*	<	LH*	LH*	<	!LH*	L-L%
faa3		LH*	LH*	---	LH*	<	LH*	<	LH*	LH*	<	!LH*	L-L%
fna1	x	LH*	LH*	---	LH*	<	LH*	<	LH*	LH*	<	!LH*	L-L%
fna2	xy	LH*	LH*	---	LH*	<	LH* H-	>	LH*	LH*	<	!LH*	L-L%
fna3	y	LH*	LH*	---	LH*	<	LH* H-	>	LH*	↓LH*	<	↑LH*	L-L%
fsf1		LH*	LH*	---	LH*	<	LH*	<	LH*	LH*	<	!LH*	L-L%
fsf2		LH*	LH*	---	LH*	<	LH*	<	LH*	LH*	<	!LH*	L-L%
fsf3		LH*	LH*	---	LH*	<	LH*	<	LH*	LH*	<	!LH*	L-L%
meh1	y	LH* H-	LH*	---	LH*	<	LH* (H-?)	<	LH*	LH*	<	LH*	L-L%
meh2		LH*	LH*	---	LH*	<	LH*	<	LH*	LH*	<	LH*	L-L%
meh3		LH*	LH*	---	LH*	<	LH*	<	LH*	LH*	<	LH*	L-L%
miz1		LH*	LH*	---	LH*	<	LH*	<	LH*	LH*	<	LH*	L-L%
miz2		LH*	LH*	---	LH*	<	LH*	<	LH*	LH*	<	!LH*	L-L%
miz3		LH*	LH*	---	LH*	<	LH*	<	LH*	LH*	<	!LH*	L-L%
mns1	y	LH*	LH*	---	LH*	<	LH* H- (H%?)	>	LH*	<	<	!LH*	L-L%
mns2		LH*	LH*	<	LH*	<	LH* H- (H%?)	<	LH*	<	<	!LH*	L-L%
mns3	x	LH*	LH*	---	LH*	<	LH*	<	LH*	x	<	!LH*	L-L%

(A.2) Transcription results for align sentence 111203: [al-?asmaa wa nimar it-tilifoon bititnisi bis-sur9a]

SPKR	*error*	al	asmaa?	wa	nimar	it	tilifoon	bititnisi	bi	s	sur9a	
faa1	x	>	LH*	>	LH*	<	x	LH*	>	>	!LH*	L-L%
faa2		>	LH*	<	LH*	<	LH*	LH*	>	>	!LH*	L-L%
faa3	x	>	LH*	<	LH*	<	LH*	LH*	>	>	!LH*	L-L%
fna1		>	LH*	<	LH*	<	LH*	LH*	>	>	!LH*	L-L%
fna2	y	>	LH*	<	LH*	<	LH* H-	LH*	>	>	LH*	L-L%
fna3	y	>	LH*	<	LH*	<	LH* H-	LH*	>	>	!LH*	L-L%
fsf1		>	LH*	<	LH*	<	LH* =	LH*	>	>	!LH*	L-L%
fsf2		>	LH*	<	LH*	<	LH* =	LH*	>	>	LH*	L-L%
fsf3	y	>	LH*	<	LH*	<	LH* H-	LH*	>	>	LH*	L-L%
meh1		>	LH*	<	LH*	<	LH*	LH*	>	>	LH*	L-L%
meh2		>	LH*	<	LH*	<	LH*	LH*	>	>	↑LH*	L-L%
meh3		>	LH*	<	LH*	<	LH*	LH*	>	>	LH*	L-L%
miz1		>	LH*	<	LH*	<	LH* =	LH*	>	>	!LH*	L-L%
miz2		>	LH*	<	LH*	<	LH*	LH*	>	>	!LH*	L-L%
miz3		>	LH*	<	LH*	<	LH*	LH*	>	>	!LH*	L-L%
mns1		>	LH*	<	LH*	<	LH*	LH*	>	>	LH*	L-L%
mns2		>	LH*	<	LH*	<	LH*	LH*	>	>	LH*	L-L%
mns3		>	LH*	<	LH*	<	LH*	LH*	>	>	LH*	L-L%

(A.3) Transcription results for align sentence 112209: [HaSalit 9ala minHa min is-sifaara 9alašaan turuuH tidrus fi ?amriika]

SPKR	*error*	HaSalit	9ala	minHa	min	is	sifaara	9ala	šaan	tiruuH	tidris	fi	?amriika	
faa1	y	LH*	---	LH* H-	<	>	LH*		LH* H- ~	LH*	LH*	>	HL	L-L%
faa2		LH*	---	LH* H-	---	>	LH*	<	LH* H- ~	LH*	LH*	<	HL	L-L%
faa3		LH*	<	LH*	---	>	LH*	<	LH*	LH*	LH*	<	LH*	L-L%
fna1	x	LH*	---	LH*	---	>	LH*	<	LH* H- ~	↑ LH*	LH*	H*	! LH*	L-L%
fna2		LH*	<	LH*	<	>	LH*	---	LH* (~)	LH*	LH*	<	! LH*	L-L%
fna3		LH*	>	LH*	---	>	↑LH*	---	LH*	LH*	LH*	<	!LH*	L-L%
fsf1	x	LH*	---	LH*	<	<	LH*	---	LH* ~	↑ LH*	LH*	<	! LH*	L-L%
fsf2	x	LH*	<	LH*	<	<	↑LH*	<	LH*	LH*	LH*	<	!LH*	L-L%
fsf3		LH*	---	LH*	LH*	<	↑LH*	<	LH*	LH*	LH*	<	!LH*	L-L%
meh1		LH*	---	LH*	<	<	LH*	<	LH*	LH*	LH*	<	LH*	L-L%
meh2		LH*	---	LH* H- ~	LH*	<	LH*	<	LH*	LH*	LH*	<	LH*	L-L%
meh3	x	LH*	---	LH*	LH*	<	LH*	<	LH*	LH*	LH*	<	LH*	L-L%
miz1	x	LH*	---	LH*	---	---	LH*	<	LH* H-	LH*	LH*	<	LH*	L-L%
miz2		LH*	---	LH*	↓LH*	---	LH*	<	LH*	LH*	LH*	<	LH*	L-L%
miz3		LH*	---	LH*	---	---	LH* H- ~	>	LH*	LH*	LH*	<	HL	L-L%
mns1	x	LH*	---	LH*	<	<	LH*	<	<	LH* H- ~	LH*	<	! LH*	L-L%
mns2		LH*	---	LH*	<	<	LH*	<	LH*	LH*	LH*	<	!LH*	L-L%
mns3		LH*	<	LH* H-	<	<	LH*	<	LH*	LH*	LH* H-	<	!LH*	L-L%

(A.4) Transcription results for align sentence 112312: [miš mumkin id-duxuul taani ba9d il-xuruug]

SPKR	*error*	miš	mumkin	id	duxuul	taani	ba9d	il	xuruug	
faa1		>	LH*	<	LH*	LH*	LH*	<?	!LH*	L-L%
faa2		>	LH*	<	LH*	LH*	LH*	<	!LH*	L-L%
faa3		↓LH*	↑LH*	<	LH*	LH*	--- ?	>	!LH*	L-L%
fna1		>	LH*	<	LH*	→LH*	LH*	<	!LH*	L-L%
fna2		>	LH*	<	→LH*	↑LH*	LH*	<	!LH*	L-L%
fna3		>	LH*	<	LH*	↑LH*	LH*	<	!LH*	L-L%
fsf1		>	LH* H-	>	LH*	→LH*	LH*	<	!LH*	L-L%
fsf2		>	LH*	>	LH*	→LH*	LH*	<	!LH*	L-L%
fsf3		>	LH*	<	LH*	→LH*	LH*	<	!LH*	L-L%
meh1		>	LH*	<	LH*	→LH*	(LH*)?	<	↑LH*	L-L%
meh2		>	LH*	<	LH*	LH*	LH*	<	!LH*	L-L%
meh3		>	LH*	<	LH*	LH*	LH*	<	!LH*	L-L%
miz1		>	LH*	<	LH*	LH*	<	<	!LH*	L-L%
miz2		>	LH*	<	LH*	→LH*	<	<	!LH*	L-L%
miz3		>	LH*	<	LH*	<	<	<	!LH*	L-L%
mns1		>	LH*	<	LH*	→ LH* =	LH* (?)	<	!LH*	L-L%
mns2		>	LH*	<	LH*	LH* H- ~	LH*	<	LH*	L-L%
mns3		>	LH*	<	LH*	(↑)LH*	LH*	<	!LH*	L-L%

(A.5) Transcription results for align sentence 121114: [fii maani9 kibiir bayni wa bayni-d-diraasa l-9ulya wa huwwa l-filuus]

SPKR	*error*	fii	maani9	kibiir	bayn-	i	wa	bayn	id-	diraasa	-l-	9ulya	wa	huwwa	il	filuus	
faa1		>	LH*	LH*	LH*	<	<	LH*	<	LH*	<	LH*	xxx	LH*	<	!LH*	L-L%
faa2		>	LH*	LH*	LH*	<	<	LH*	<	LH*	<	LH*	<	LH*	<	!LH*	L-L%
faa3		>	↓LH*↓	↑LH*	LH*	<	<	LH*	<	LH*	<	LH*	<	LH*	<	!LH*	L-L%
fna1	x	>	LH*	LH*	LH*	<	<	LH*	<	LH*	<	LH* H- ~	<	LH*	<	!LH*	L-L%
fna2		>	LH*	LH*	LH*	<	<	LH*	<	LH*	<	LH* H- ~	<	LH*	<	!LH*	L-L%
fna3		>	LH*	LH*	LH*	<	<	LH*	<	LH*	<	LH* H-	<	LH*	<	!LH*	L-L%
fsf1		LH*=	LH*	LH*	LH*	<	<	LH*	<	LH*	<	LH*	<	LH*	<	HL	L-L%
fsf2		LH*=	LH*	LH*	LH*	<	<	LH*	<	LH*	<	LH*	<	LH*	<	HL	L-L%
fsf3		>	LH*	LH*	LH*	<	<	LH*	<	LH*	<	LH*	<	LH*	<	HL	L-L%
meh1	x	>	LH*	LH*	LH*	<	<	LH*	<	LH* H- ~	<	LH* H-	<	LH*	<	!LH*	L-L%
meh2		LH*	LH*	LH*	LH*	<	<	LH*	<	LH*	<	LH*	<	LH*	<	!LH*	L-L%
meh3		>	LH*	LH*	LH*	<	<	LH*	<	LH*	<	LH*	<	LH*	<	LH*	L-L%
miz1		>	LH*	LH*	LH*	<	<	<	<	LH*	<	LH*	<	LH*	<	LH*	L-L%
miz2	x	>	LH*	LH*	xxx	x	<	LH*	<	LH*	<	LH*	<	LH*	<	LH*	L-L%
miz3		>	LH*	LH*	LH*	<	<	<	<	LH*	<	LH*	<	LH*	<	LH*	L-L%
mns1	x	>	LH*	LH*	LH*	<	<	LH*	<	LH*	<	!LH* L- ~	>	LH*	<	!LH*	L-L%
mns2		>	LH*	LH*	LH*	<	<	LH*	<	LH*	<	!LH* L- ~	>	LH*	<	!LH*	L-L%
mns3		>	LH*	LH*	LH*	<	<	LH*	<	LH*	<	!LH* H- ~	>	LH*	<	!LH*	L-L%

(A.6) Transcription results for align sentence 121317: [šufna nuunu SuGayyar ?awwi fil-mustašfa]

SPKR	*error*	šufna	nuunu	SuGayyar	?awwi	fi	-l-	mustašfa	
faa1		LH*	LH*	LH*	LH*	<	<	!LH*	L-L%
faa2		LH*	LH*	LH*	LH*	<	<	!LH*	L-L%
faa3		LH*	LH*	LH*	LH*	<	<	!LH*	L-L%
fna1		LH*	LH*	↑LH*	→LH* =	<	<	!LH*	L-L%
fna2		LH*	LH*	LH*	↑LH*	<	<	!LH*	L-L%
fna3		LH*	LH*	LH*	LH* H- ~	LH*	<	!LH*	L-L%
fsf1		LH*	LH*	LH*	LH*	<	<	!LH*	L-L%
fsf2		LH*	LH*	LH*	LH*	<	<	LH*	L-L%
fsf3		LH*	LH*	LH*	LH*	<	<	!LH*	L-L%
meh1	x	LH*	LH*	LH* H-	LH*	<	<	!LH*	L-L%
meh2		LH*	LH*	LH*	LH*	<	<	!LH*	L-L%
meh3		LH*	LH*	LH*	LH*	<	<	!LH*	L-L%
miz1		LH*	LH*	LH*	LH*	<	<	!LH*	L-L%
miz2		LH*	LH*	LH*	LH*	<	<	!LH*	L-L%
miz3		LH*	LH*	LH*	LH*	<	<	!LH*	L-L%
mns1		LH*	LH*	LH*	LH* H- ~	LH*	<	!LH*	L-L%
mns2		LH*	LH*	LH*	LH*	<	<	!LH*	L-L%
mns3		LH*	LH*	LH*	LH*	<	<	!LH*	L-L%

(A.7) Transcription results for align sentence 212120: [il-walad da minamrad xaaliS ma šuftiš zayyuh]

SPKR	*error*	il	walad	da	minamrad	xaaliS	ma	šuft	-š	zayy-uh	
faa1		>	LH*	LH*	LH*	LH*	>	LH*	< (?)	LH*	L-L%
faa2		>	LH*	LH*	LH*	LH*	>	LH*	<	LH*	L-L%
faa3		>	LH*	LH*	LH*	LH*	<	LH*	<	HL ?	L-L%
fna1		>	LH*	LH*	LH*	LH*	>	↑↑LH*	<	LH*	L-L%
fna2		>	LH*	↑LH*	LH*	LH* H-	>	LH*	<	LH*	L-L%
fna3		>	LH*	↓LH*	↑LH*	LH* H-	>	LH*	<	LH*	L-L%
fsf1		>	LH*	LH*	(LH*)?	LH* ~	>	↑↑LH*	<	LH*	L-L%
fsf2		>	LH*	LH*	(LH*)?	LH*	>	↑LH*	<	LH*	L-L%
fsf3		>	LH*	LH*	LH*	LH* H- ~	>	LH*	<	LH*	L-L%
meh1		>	LH*	LH* H- ~	LH*	LH* H-	>	LH*	<	LH*	L-L%
meh2		>	LH*	LH* (H-)? =	LH*	LH*	>	LH*	<	LH*	L-L%
meh3		>	LH*	LH* H-	LH*	→LH*	>	LH*	<	LH*	L-L%
miz1		>	LH*	LH*	→LH*	LH*	< ?	LH*	<	HL?	L-L%
miz2		>	LH*	LH* =	LH*	LH*	>	LH*	<	LH*	L-L%
miz3	x	>	LH*	< ?	LH*	LH*	>	LH*	<	LH*	L-L%
mns1	x	>	LH*	>	LH*	LH* H-	>	LH*	< ~~	↓LH*	L-L%
mns2		>	LH*	>	LH*	LH* H-	>	LH*	<	LH*	H-L%
mns3		>	LH*	>	→LH*	LH* H-	>	LH*	<	LH*	L-L%

(A.8) Transcription results for align sentence 212121: [9ammi mimangih nafsuh ?awwi ba9d ma gayy min barra]

SPKR	*error*	9ammi	mimangih	nafsuh	?awwi	ba9d	ma	gayy	min	barra	
faa1		LH*	LH*	LH*	LH* H- =	---	<	LH*	<	!LH*	L-L%
faa2	x	LH*	LH*	→LH*	→LH* ~~	→LH*	<	→LH*	<	LH*	L-L%
faa3		LH*	LH*	LH*	↑LH*	LH*	<	LH*	<	HL ?	L-L%
fna1		LH*	→LH*	LH*	LH* H- ~	LH*	<	→LH*	<	!LH*	L-L%
fna2		LH*	→LH*	→LH*	LH* H- ~	LH*	<	→LH*	<	!LH*	L-L%
fna3	x	LH*	LH*	→LH*	↑LH* H-	LH*	<	→LH*	<	!LH*	L-L%
fsf1		LH*	LH*	→LH*	→LH* =	→LH*	<	LH*	<	!LH*	L-L%
fsf2		LH*	LH*	→LH*	→LH*	LH*	<	LH*	<	!LH*	L-L%
fsf3		LH*	LH*	→LH*	LH* H- =	→LH*	<	LH*	<	!LH*	L-L%
meh1	x	LH*	→LH*	→LH* H-	→LH*	LH*	<	↑LH*	<	LH*	L-L%
meh2	x	LH* ~	LH*	LH*	LH*	< ?	<	LH*	<	LH*	L-L%
meh3		LH* H-?	LH*	LH*	LH* =	<	<	LH*	<	LH*	L-L%
miz1		LH*	→LH*	→LH*	→LH* H-	LH*	<	LH* H-	>	LH*	L-L%
miz2		LH*	LH*	→LH*	→LH* H-	--- ?	<	LH*	<	LH*	L-L%
miz3		LH*	→LH*	→LH*	LH*	<	<	→LH*	<	LH*	L-L%
mns1	x	LH* H- ~	LH*	LH*	LH* =	LH*	<	→LH*	<	LH*	L-L%
mns2		LH* H-H%	LH*	LH*	LH* =	↓LH* (?<)	<	LH*	<	LH*	L-L%
mns3		LH* H-H%	LH*	LH*	LH*	<	<	LH*	<	LH*	L-L%

(A.9) Transcription results for focus sentence 121: [maama bitit9allim yunaani bil-layl]

SPKR	*error*	maama	bitit9alilm	yunaani	bi-	-l-	layl	
faa1		↑LH* =	LH*	LH*	<	<	LH*	L-L%
faa2		↑LH* =	LH*	LH*	<	<	LH*	L-L%
faa3		↑LH*	LH*	LH*	<	<	LH*	H-H%
fna1		LH*	LH*	LH*	<	<	LH*	H-H%
fna2		LH*	→LH*	→LH*	<	<	LH*	L-L%
fna3	x	LH*	→LH*	LH*	<	<	!LH*	L-L%
fsf1		LH*	LH*	LH*	<	<	LH*	L-L%
fsf2		LH*	LH*	→LH*	<	<	→LH*	L-L%
fsf3		LH*	LH*	→LH*	<	<	LH*	H-H%
meh1		LH*	LH*	LH*	<	<	H*	L-L%
meh2		LH*	LH*	LH*	<	<	→H*	L-L%
meh3		LH*	LH*	→LH*	<	<	→LH*	L-L%
miz1		LH*	LH*	LH*	<	<	LH*	H-
miz2		LH*	LH*	→LH*	<	<	LH*	H-H%
miz3		LH*	LH*	→LH*	<	<	LH*	H-H%
mns1		LH*	LH*	LH*	<	<	LH*	H-H%
mns2		LH*	↓LH*	LH*	<	<	!LH*	L-L%
mns3		LH*	LH* H- = ~	LH*	<	<	!LH*	L-L%

(A.10) Transcription results for focus sentence 122: [maama bitit9allim yunaani bil-layl]

SPKR	*error*	maama	bitit9alilm	yunaani	bi-	-l-	layl	
faa1		LH*	LH*	LH*	<	<	LH*	H-H%
faa2		LH*	LH*	LH*	<	<	LH*	H-H%
faa3		LH*	LH*	LH*	<	<	LH*	H-H%
fna1		LH*	LH*	LH*	<	<	LH*	H-H%
fna2		LH*	LH*	LH*	<	<	LH*	L-L%
fna3		LH*	LH*	LH*	<	<	LH*	H-H%
fsf1		LH*	LH*	LH*	<	<	LH*	H-H%
fsf2		LH*	LH*	LH*	<	<	LH*	H-L%
fsf3		LH*	LH*	LH*	<	<	LH*	H-H%
meh1		LH*	→LH*	LH*	<	<	LH*	L-L%
meh2		LH*	LH*	→LH*	<	<	LH*	H-H%
meh3		LH*	LH*	LH*	<	<	LH*	H-H%
miz1		LH*	LH*	LH*	<	<	LH*	H-H%
miz2		LH* H-	→LH*	→LH*	<	<	LH*	H-H%
miz3		LH*	LH*	LH*	<	<	LH*	H-H%
mns1		LH*	→LH*	LH*	<	<	LH*	H-H%
mns2		LH*	→LH*	→LH*	<	<	LH*	H-H%
mns3		LH*	→LH*	→LH*	<	<	LH*	H-H%

(A.11) Transcription results for focus sentence 123: [maama bitit9allim yunaani bil-layl]

SPKR	*error*	maama	bitit9alilm	yunaani	bi-	-l-	layl	
faa1		LH*	LH*	LH*	<	<	LH*	L-L%
faa2		LH*	LH*	LH*	<	<	LH*	L-L%
faa3		LH*	LH*	LH*	<	<	LH*	L-L%
fna1		LH*	LH*	LH*	<	<	LH*	L-L%
fna2		LH*	LH*	LH*	<	<	LH*	L-L%
fna3		LH*	LH*	LH*	<	<	LH*	L-L%
fsf1		LH*	LH*	LH*	<	<	LH*	L-L%
fsf2		LH*	LH*	LH*	<	<	LH*	L-L%
fsf3		LH*	LH*	LH*	<	<	LH*	L-L%
meh1		LH*	LH*	LH*	<	<	LH*	L-L%
meh2		LH*	LH*	LH*	<	<	LH*	L-L%
meh3		LH*	LH*	LH*	<	<	LH*	L-L%
miz1		LH*	LH*	LH*	<	<	LH*	L-L%
miz2		LH*	LH*	LH*	<	<	LH*	L-L%
miz3		LH*	LH*	LH*	<	<	LH*	L-L%
mns1		LH*	LH*	LH*	<	<	LH*	L-L%
mns2		LH*	LH*	LH*	<	<	LH*	L-L%
mns3		LH*	LH*	LH*	<	<	LH*	L-L%

(A.12) Transcription results for focus sentence 124: [maama bitit9allim yunaani bil-layl]

SPKR	*error*	maama	bitit9alilm	yunaani	bi-	-l-	layl	
faa1		LH*	LH*	LH*	<	<	LH*	L-L%
faa2		LH*	LH*	LH*	<	<	LH*	L-L%
faa3		LH*	LH*	LH*	<	<	LH*	L-L%
fna1		LH*	LH*	LH*	<	<	LH*	L-L%
fna2		LH*	LH*	LH*	<	<	LH*	L-L%
fna3		LH*	LH*	LH*	<	<	LH*	L-L%
fsf1		LH*	LH*	LH*	<	<	LH*	L-L%
fsf2		LH*	LH*	LH*	<	<	LH*	L-L%
fsf3		LH*	LH*	LH*	<	<	LH*	L-L%
meh1		LH*	LH*	LH*	<	<	LH*	L-L%
meh2		LH*	LH*	LH*	<	<	LH*	L-L%
meh3		LH*	LH*	LH*	<	<	LH*	L-L%
miz1		LH*	LH*	LH*	<	<	LH*	L-L%
miz2		LH*	LH*	LH*	<	<	LH*	L-L%
miz3		LH*	LH*	LH*	<	<	LH*	L-L%
mns1		LH*	LH*	LH*	<	<	LH*	L-L%
mns2		LH*	LH*	LH*	<	<	LH*	L-L%
mns3		LH*	LH*	LH*	<	<	LH*	L-L%

(A.13) Transcription of the narrative folk tale "Guha and the banana seller" (Abdel-Massih 1975).

1	guHa	kaan	Tuul	9umr-uh	9aayiš	fi-l-?ariyaaf
	Guha	was	all	life-his	living	in-the-country

'Guha had lived all his life in the countryside.'

2	wa	marra	min-l-	marraat	fakkar	yinzil	maSr	?aSHaab-uh	?aaluu-l-uh
	and	time	from-the	times	thought	go-down	Cairo	friends-his	said-to-him

'And one time he thought he might go to Cairo. His friends said to him:'

3	xalli	baal-ik	ya	guHa	min-il	bayyaa9iin	bitu9	maSr	dool	waHšiin	?awwi
	take	care-your	oh	Guha	from-the	sellers	belonging	Cairo	those	bad	very

'Be careful Guha, of the traders in Cairo. They're up to no good.'

4	?awwal-ma	yisma9u	laHghat-ak	wa	ya9rifu	?inn-ak	min-il-	?ariyaaf	yiGallu	9alay-k	il-Haaga
	first-that	they-hear	accent-your	and	they-know	that-you	from-the	country	they-increase	to-you	the-thing

'As soon as they hear your accent, and realise you are from the country, they'll put the prices up.'

5	ya9ni	il-Haaga	illi	taman-ha	9ašara	SaaG	yi?uluu-lak	9alay-ha	bi-9išriin	?irš
	it-means	the-thing	that	price-its	ten	piastres	they-say-you	on-it	with-twenty	piastres

'Like, if something is ten piastres, they will tell you its twenty piastres.'

6	fa-?inta	?iza	Habbayt	tištiri	Haaga	wa-?aalu-lak	9ala	tamanha	?uluhum	in-nuSS	9ala	Tuul
	so-you	if	you-wanted	you-buy	thing	and-they-say you	on	price-its	say-to-them	the-half	at	once

'So if you want to buy something and tell you a price tell them half of that price at once.'

7	wa	law	waaHid	gayy	yiddiilak	Haaga	?ul-luh	iD-Da9af
	and	if	one	came	he-gives-you	thing	say-him	the-double

'And if someone gives you something ask him for double.'

8	ya9ni	masalan	law-ruHt	tištiri	Haaga	wa-l-bayaa9	?aal-lak	bi-9ašara	SaaG	?ul-luh	la?	bi-xamsa	SaaG
	it-means	for-example	if you-went	you-buy	thing	and-the seller	said-to-you	for-ten	piastres	say-to-him	no	for-five	piastres

'Like, for example, if you go to buy something and the trader says ten piastres, tell him five!'

9	wa-law	?ult	li-waaHid	bayaa9	?iddiini	xamsa	SaaG	burtu?aan	wa-?addaalak	xamas	burtu?unaat	?ul-luh la?	9ašara
	and-if	you-said	to-one	seller	give-me	five	piastres	of-oranges	and-gave-you	five	oranges	say-to-him no	ten!

'And if you ask a trader to give you fives piastres' worth of oranges, and he gives you five oranges, tell him you want ten!'

10	guHa	?aal	la-hum	mataxafuuš	9alayya	wa	raaH	guHa	9ala	maSr
	Guha	said	to-them	don't worry	about-me	and	went	Guha	to	Cairo

'Guha said to them: "Don't you worry about me!". And Guha went to Cairo.'

11	wa	?a9ad	9ala	?ahwa	min	?ahawii-ha	yišrab	šayy
	and	sat	in	café	from	her-cafés	he-drinks	tea

'He was sitting in one of the cafés, drinking tea...'

12	faat	waaHid	bayyaa9	il-mooz	biynaadi	9ala-l-mooz	wa	yi?uul	il-mooz	il-Helw	il-mooz	il-Helw
	passed	one	seller	of-bananas	he calls	of-the bananas	and	he-says	the-bananas	the-nice	the-bananas	the-nice

'when a banana seller passed by, calling out about his bananas: "Lovely bananas! Lovely bananas!"'

13	nadaa	-luh	guHa	wa	?aal	luh	ta9aala	hina	bi-kam	kiilu	l-mooz
	called	to-him	Guha	and	said-	to-him	come	here	how-much	kilo	of-bananas

'Guha called out to him and said: "Come over! How much is a kilo of bananas?"'

14	il-bayyaa9	?aal-luh	bi-	tnaašir	?irš
	the-seller	said-to-him	for	twelve	piastres

'The trader said: "Twelve piastres"'

15	wa	guHa	?aal-luh	la?	sitta	SaaG
	and	Guha	said-to-him	no	six	piastres

‘And Guha said to him: “No! Six piastres!”’

16	ir-raagil	?aal-luh	?a?ul-lak	?eh	?ana	Habi9-lak	kiilu	bi-sitta	SaaG	bass	9ašaan	xaTrak
	the-man	said-to-him	I-say-to -you	what	I	will-sell-to-you	kilo	for-six	piastres	just	on-account	your-sake

‘The man said to him: “I tell you what, I will sell you a kilo for six piastres, just for you.”’

17	fa	guHa	?aal-luh	9ala	Tuul	la?	bi-talaata	SaaG
	so	Guha	said-to-him	at	once	no	for-three	piastres

‘So Guha immediately said to him: “No! Three piastres!”’

18	il	bayyaa9	?aal-luh	ya	xuya	?inta	?ult	sitta	SaaG	min	da?ii?a
	the	seller	said-to-him	oh	brother-my	you	said	six	piastres	from	a-minute

‘The trader said to him: “Mate, you said six piastres a minute ago.”’

19	guHa	?aal-luh	talaata	SaaG	wa	ma-fiiš	Gayr	kida
	Guha	said-to-him	three	piastres	and	not-there-is	except	that

‘Guha said to him: “Three piastres and not a penny more!”’

20	bayyaa9	l-mooz	?aal-luh	?inta	fakir-ni	la?ay	fi-š-šaari9	walla	sar?-uh
	seller	the-bananas	said-to-him	you	think-me	finding	in-the-street	or	stealing-it

‘The banana seller said to him: “Do you think I found them in the street, or stole them?”’

21	?awzin	lak	kiilu	bi-sitta	SaaG	yalla
	I-will-weigh	for-you	kilo	for-six	piastres	come-on

‘“I will weigh you out a kilo for six piastres, come on now.”’

22	guHa	?aal-luh	(laa?)	bi-talaata	SaaG
	guha	said-to-him	no	for-three	piastres

‘Guha said to him: “No! Three piastres!”’

23	il-bayyaa9	buSS	li-guHa	kidda	wa	?aal-luh
	the-seller	looked	to-Guha	like-that	and	said-to-him

‘The trader looked at him for at moment and said:’

24	?ey	ra?y-ak	?inta	raagil	Tayyib	wa	?ana	Habbayt-ak	min	?awwal-ma	šuft-ak
	what	your-opinion	you	man	honest	and	I	liked-you	from	first-that	I-saw-you

‘“What do you think? You are an honest man and I liked you from the moment I saw you.”’

25	?ana	9awz	?awzin-lak	kiilu	bi-balaaš
	I	want	I-weigh-for-you	kilo	for-free

‘“I will weigh you out a kilo for free!”’

26	guHa	naTT	min-il	kursi	illi	(huwwa)	kaan	?aa9id	9alayh	wa	?aal-luh
	Guha	jumped	from-the	chair	that	he	was	sitting	in-it	and	said-to-him

‘Guha jumped up from where he was sitting and said to him:’

27	itniin	kiilu	itniin	kiilu	?aw9i	tiftikir	?innak	Ha-tiDHak	9alayya
	two	kilos	two	kilos	or	you-think	that-you	will-laugh	at-me

‘“Two kilos! Two kilos! Don’t you dare think you can make fun of me!”’

(A.14) Transcription results for read narrative

1	guHa	kaan	Tuul	9umr	-uh	9aayiš	fi	-l-	?ariyaaf
fna2	LH*	---	LH*	LH*	<	LH*	<	<	!LH*
fsf2	LH*	<	<	LH*	<	LH*	<	<	!LH* L-L%
meh2	LH*	<	(LH*)	LH*	<	LH*	<	<	LH* H-
miz2	LH*	<	LH*	LH*	<	LH*	<	<	LH* H-
mns2	LH*	<	<	LH*	<	LH*	<	<	LH* L-

2	wa	marra	min	-l-	marraat	fakkar	yinzil	maSr	?aSHaab	-uh	?aaluu	-l-	-uh
fna2	>	LH*	<	<	LH*	< ?	LH*	!LH* L-L%	↑↑LH*	<	LH*	<	<
fsf2	>	LH*	<	<	LH*	LH*	LH*	!LH* L-L%	LH*	<	LH*	<	< H-H%
meh2	>	LH*	LH*	<	LH*	LH*	LH*	H* L-L%	LH*	<	LH*	<	< L-L%
miz2	>	LH*	<	<	LH*	LH*	LH*	!LH* L-L%	LH*	<	LH*	<	< H-H%
mns2	>	LH*	<	<	LH*	LH*	LH*	!LH* L-L%	LH*	<	LH*	<	<

3	xalli	baalik	ya	guHa	min	-il	bayyaa9iin	bitu9	maSr	dool	waHšiin	?awwi
fna2	>	LH*	<	↑LH*	<	<	LH*	<	↑LH* H-	---	→LH*	→LH* L-L%
fsf2	LH*	LH*	<	LH*	<	<	LH*	<	↑LH* H-	>	LH*	!LH* L-L%
meh2	LH* =	LH*	<	LH* L-	LH*	<	LH*	LH*	(L)H* L-L%	LH*	LH*	!LH* L-L%
miz2	>	LH*	<	LH*	<	<	LH*	<	LH* L-	>	LH*	LH* L-L%
mns2	<	LH*	<	LH* L-	>	>	LH*	<	!LH* L-	>	↑LH*	!LH* L-L%

4	?awwal	ma	yisma9u	laHghat-	-ak	wa	ya9rifu	?innak	min	il	?ariyaaf	yiGallu	9alayk	il-	Haaga
fna2	↑LH*	<	LH*	LH*	<	<	LH*	< (↓LH*)	<	<	LH* H-	LH*	LH*	<	!LH* L-L%
fsf2	LH*	<	LH*	LH*	< H-	>	LH*	LH*	<	<	LH*	LH*	LH*	<	!LH* L-L%
meh2	LH*	<	LH*	LH*	<	<	LH*	LH*	LH*	<	LH* H-	LH*	LH*	<	!LH* L-L%
miz2	LH*	<	LH*	LH*	<	<	LH*	<	<	<	LH*	LH*	LH*	<	!LH* L-L%
mns2	LH*	<	LH*	LH*	<	<	LH*	<	<	<	LH* H-	LH*	LH*	<	LH* H-H%

(A.14) Transcription results for read narrative (continued)

5	ya9ni	il	Haaga	illi	taman	-ha	9ašara	SaaG	yi?uluu-	-lak	9alay	-ha	bi	-9išriin	?irš
fna2	↑↑LH*	<	→LH*	<	LH*	< ~	↑ LH*	↑LH* H-	→LH*	<	xxx	xxx	<	LH*	!LH* L-L%
fsf2	>	>	LH*	<	LH*	<	→LH*	LH* H-	LH*	<	→LH*	<	<	LH*	!LH* L-L%
meh2	LH*	<	↑LH*	<	LH*	<	LH* ~	LH* H-	LH*	<	LH*	<	<	LH*	!LH* L-L%
miz2	---	>	LH*	<	LH*	<	LH*	LH*	LH*	xxx	LH*	<	<	LH*	LH* L-L%
mns2	>	>	LH*	<	LH*	<	LH*	LH* H-	LH*	< H-	LH*	<	<	< H-H%	LH* L-L%

6	fa	?inta	?iza	(giit)	Habbayt	tištiri	Haaga	wa	?aalu	lak	9ala	tamanha	?uluhum	in-nuSS	9ala	Tuul
fna2	>	LH*	LH* =	LH*	xxx	LH*	↑↑LH* LH* [ayyi Haaga]	>	LH*	<	↓LH*	LH*	↑LH*	!LH* L-L%	xxx	xxx
fsf2	>	LH*	<	xxx	LH*	<	LH* H-	>	LH*	<	↓LH*	LH* H-	LH*	LH*	<	!LH* L-L%
meh2	>	LH*	LH*	LH* H-	LH*	LH*	LH* H-	>	LH*	<	↓LH*	LH* H-	LH*	LH*	<	!LH* L-L%
miz2	>	LH*	↓LH* =	xxx	LH*	LH*	LH*	<	LH*	<	<	LH* H-	LH*	LH*	---	!LH* L-L%
mns2	>	LH*	<	xxx	LH*	<	LH* H-	<	LH*	<	<	LH* H- ~	LH*	LH*	---	!LH* L-L%

7	wa	law	waaHid	gayy	yiddiilak	Haaga	?ul	luh	(la?)	iD-Da9af
fna2	xxx	>	LH* [er]	LH*	LH*	LH*	LH*	<	xxx	H* L-L%
fsf2	xxx	>	LH*	LH*	LH* [yiddiik]	LH* H-	LH*	<	xxx	!LH* L-L%
meh2	>	>	LH*	LH*	LH*	LH* H-	LH*	<	LH*	LH* L-L%
miz2	>	>	LH*	LH*	LH*	LH*	LH*	<	LH*	LH* L-L%
mns2	>	>	LH*	LH*	LH* [yiddiik]	LH* H-	LH*	<	xxx	!LH* L-L%

8	ya9ni	masalan	law	ruHt	tištiri	Haaga	wa	-l-	bayaa9	?aal-lak	bi-	9ašara	SaaG	?ul-luh	(la?)	bi-xamsa	SaaG
fna2	>	LH* H-	>	LH*	LH*	LH*	<	<	LH*	LH*H- ~~	>	↑LH*	!LH*L-L%	LH* H-	xxx	LH*	!LH* L-L%
fsf2	---	LH*	<	LH*	LH*	LH* H-	>	>	LH*	LH*	<	LH*	!LH*L-L%	↑↑LH*	xxx	LH*	!LH* L-L%
meh2	---	LH*	<	LH*	LH*	LH*	<	<	LH*	LH*H- ~	>	LH*	LH* H-	LH*	LH* L-	LH*	!LH* L-L%
miz2	>	LH*	<	LH*	(LH*)	LH*	<	<	LH*	LH*	<	LH*	LH*	? <	? <	LH*	LH*
mns2	LH*	!LH* L-L%	>	LH*	(LH*)	LH*	< [min]	<	LH*	LH*	<	LH*	!LH*L-L%	LH* L- =~	xxx	LH*	!LH* L-L%

(A.14) Transcription results for read narrative (continued)

9	wa	law	?ult	li-	waaHid	bayaa9	?iddiini	xamsa	SaaG	burtu?aan	wa-	?addaalak	xamas	burtu?unaat	?ul-luh	la?	9ašara
fna2	LH*	<	LH* ~~	>	LH*	LH* [biybii9]	LH*	→H*	→H*	LH*	>	LH* [?addaak]	H*	LH* H-	LH*	<	!LH* L-L%
fsf2	>	>	LH*	<	LH* ~	LH*	LH*	LH*	LH*	LH* H- H%	>	LH*	LH*	LH* H-	LH*	!LH* L-	LH* L-L%
meh2	xxx	>	LH*	<	LH*	LH*	LH*	LH* H- ~	LH* H- ~	LH* H-	>	LH*	LH*	LH* H-	LH*	!LH* L-	LH* L-L%
miz2	---	---	LH*	<	LH*	LH*	LH*	LH*	LH* H- ~~	LH*	<	LH*	(LH*)?	LH* H-	LH*	<	!LH* L-L%
mns2	>	>	LH*	<	LH* H-	---	LH*	<	<	LH*	<	LH*	LH*	LH* H-	LH*	!LH* L-	!LH* L-L%

10	guHa	?aal	la-	hum	mataxafuuš	9alayya	wa	raaH	guHa	9ala	maSr
fna2	LH*	LH*	<	<	LH*	!LH* L-L%	>	LH*	LH*	<	LH* L-L%
fsf2	LH*	LH*	<	<	LH*	!LH* L-L%	>	LH*	LH*	<	!LH* L-L%
meh2	LH*	LH*	<	< H-	LH*	LH* L-L%		LH*	LH*	<	LH* L-L%
miz2	LH*	LH*	<	<	LH*	LH*	<	LH*	LH*	<	LH* H-(L%)
mns2	LH*	LH* H- ~	LH*	< L-	LH* H-	LH* L-L%	<	LH*	LH*	<	LH* H-

11	wa	?a9ad	9ala	?ahwa	min	?ahawii-ha	yišrab	šayy
fna2	>	LH*	<	LH*	<	LH*	LH*	!LH* L-L%
fsf2	>	LH*	<	LH*	<	LH*	LH*	!LH* L-L%
meh2	>	LH*	↓LH*	LH* L-	LH*	LH*	LH*	LH* L-L%
miz2	>	LH*	<	LH*	<	LH*	LH*	LH* H-
mns2	>	LH*	<	LH* H-	>	LH*	<	LH* L-L%

12	faat	waaHid	bayyaa9	mooz	biynaadi	9ala-l-	mooz	wa	yi?uul	il-	mooz	il	Helw	il-	mooz	il	Helw
fna2	%H ---	---	LH*	LH*	LH*	>	LH*	<	!LH* L-L%	>	↑↑LH*	<	→LH* L-L%	>	→LH*	<	LH* L-L%
fsf2	LH*	LH*	LH*	LH*	LH*	<	LH*	<	LH* H-	>	LH*	<	!LH* L-L%	>	LH*	<	!LH* L-L%
meh2	LH*	↑LH*	LH*	LH* L-	LH*	<	LH*	<	LH* H-	>	LH*	<	↑LH* L-L%	>	LH*	<	LH* L-L%
miz2	>	LH*	LH*	LH*	LH*	<	LH*	<	LH*	>	LH*	<	LH*	>	LH*	<	!LH* L-L%
mns2	>	LH*	LH*	LH* L-	LH*	<	LH*	<	LH*	>	LH*	xxx	xxx	xxx	xxx	xxx	xxx

(A.14) Transcription results for read narrative (continued)

13	nadaa	-luh	guHa	wa	?aal	luh	ta9aala	hina	bikam	kiilu	l-	mooz
fna2	LH*	<	H*	<	LH*	< H-	↑LH*	!LH* L-L%	↑LH*	LH*	<	!LH* H-L%
fsf2	LH*	<	LH*	<	LH*	< H-	LH*	LH* L-	LH*	<	<	!LH* H-L%
meh2	LH*	<	LH*	<	LH*	< L-	LH*	LH*	LH*	LH*	<	!LH* H-L%
miz2	LH*	<	LH*	<	LH*	< H-	LH*	LH*	LH*	LH*	<	!LH* L-L%
mns2	xxx	xxx	xxx	xxx	xxx	xxx	xxx	xxx	LH*	LH*	<	< H-L%

14	il-	bayyaa9	?aal	luh	bi-	tnaašir	?irš
fna2	>	LH*	LH*	<	<	LH*	LH* H-L%
fsf2	>	LH*	LH*	< H-	>	LH*	!LH* L-L%
meh2	>	LH*	LH*	< H-	<	LH*	LH* H-L%
miz2	>	LH*	LH*	<	< ~	LH*	!LH* L-L%
mns2	>	LH*	LH*	<	<	LH*	!LH* L-L%

15		wa	guHa	?aal (luh)	la?	sitta	SaaG	wa	ma	fiiš	Gayr	kida
fna2		>	LH*	LH* =	LH*L-	↑LH*	LH* L-L%	xxx	>	LH*	<	!LH* L-L%
fsf2		>	LH*	LH* H-	xxx	LH*	LH*	<	<	LH*	<	!LH* L-L%
meh2		>	LH*	LH*	LH*L- =	LH*	LH* L-L%	>	>	LH*	<	!LH* L-L%
miz2		>	LH*	LH*	xxx	LH*	LH* L-L%	>	>	LH*	<	!LH* L-L%
mns2		>	LH*	LH* H-	xxx	LH*	LH* L-L%	>	>	LH*	<	!LH* L-L%

16	ir	-raagil	?aal	luh	?a?ul	lak	(?ey)	?ana	Habi9	lak	kiilu	bi	sitta	SaaG	bass	9ašaan	xaTrak
fna2	xxx	xxx	xxx	xxx	xxx	xxx	xxx	>	LH*	<	LH*	<	LH*	LH* H-	> (?)	LH*	!LH* L-L%
fsf2	>	LH*	!LH*	H-L%	LH*	H-	xxx	>	LH*	< H-	LH*	<	LH*	!LH* L-L%	LH*	<	LH* L-L%
meh2	xxx	LH*	LH*	< L-	LH*	<	!LH* L-L%	LH* L- = ~	LH*	LH* (repair)	LH*	<	LH*	LH* L-	LH*	LH* H-	LH* L-L%
miz2	>	LH*	LH*	< ~	LH*	H-	xxx	>	LH*	< H- ~	LH*	<	LH*	LH*	> (?)	(LH*)	!LH* L-L%
mns2	>	LH*	LH*	<H-	LH*	<	xxx	>	LH*	<	LH* H-	<	LH*	LH* H-	LH* L-	LH*	LH* L-L%

(A.14) Transcription results for read narrative (continued)

17	fa	guHa	?aal-luh	9ala	Tuul	la?	bi-	talaata	SaaG
fna2	>	LH*	LH*	---	LH* H-	LH* L-	>	LH*	!LH* L-L%
	fa	guHa	9ala	Tuul	?aal-luh	la?	bi-	talaata	SaaG
fsf2	>	LH*	<	LH*	LH* H-	xxx	>	LH*	!LH* L-L%
meh2	>	LH*	---	LH*	LH*	LH* L-	>	LH*	!LH* L-L%
miz2	>	LH*	<	LH*	LH*	LH*	>	LH*	LH* L-
mns2	>	LH*	<	LH* H-	LH* H-	xxx	>	LH*	!LH* L-L%

18	il	bayyaa9	?aal-luh	ya	xuya	?inta	?ult	sitta	SaaG	min	da?ii?a
fna2	>	LH*	LH* H-	>	LH*	LH*	LH*	LH*	LH*	>	↑LH* L-L%
fsf2	>	LH*	LH* H-	>	LH*	LH*	LH*	LH*	LH*	<	LH* L-L%
meh2	>	LH*	LH* H-	>	LH*	LH*	LH*	LH*	LH*	↓LH*	↑LH* L-L%
miz2	>	LH*	LH* H- ~	>	LH*	LH*	LH*	LH*	LH*	<	LH* L-L%
mns2	>	LH*	LH*	<	LH*	LH* H-	LH*	LH*	LH*	<	LH* H-

19	guHa	?aal-luh	talaata	SaaG	wa	ma	fiiš	Gayr	kida
fna2	LH*	LH* H-H%	LH*	LH* L-L%	xxx	>	LH*	<	LH* L-L%
fsf2	LH*	LH* H-H%	LH*	LH*	<	<	LH*	<	LH* L-L%
meh2	LH*	LH* H-	LH*	LH* L-L%	>	>	LH*	---	LH* L-L%
miz2	LH*	LH*	LH*	LH* L-L%	xxx	>	LH*	<	LH* L-
mns2	LH*	LH* H-H%	LH*	LH* L-	>	>	LH*	<	LH* L-L%

20	bayyaa9	l-	mooz	?aal-luh	?inta	faakir	-ni	la?ay	fi-š	šaari9	walla	sar?-uh
fna2	LH*	---	LH*	LH* H- H%	>	↑LH*	<	LH* [laa?i]	<	LH*	<	LH* H-L%
fsf2	LH*	<	LH*	LH* H- H%	>	LH*	<	LH*	<	LH*	<	LH* L-L%
meh2	LH*	<	LH*	LH* H-	>	LH*	<	LH*	<	LH* H- ~	LH*	↑LH* L-L%
miz2	LH*	<	LH*	LH*	<	LH*	<	LH*	<	LH* L-	>	↑LH* H-H%
mns2	LH*	<	↓LH*	LH* H-	<	LH*	<	LH*=	<	LH*	<	LH* H-L%

(A.14) Transcription results for read narrative (continued)

21	?awzin	lak	kiilu	bi	sitta	SaaG	yalla	*yalla*
fna2	LH*	<	LH*	<	LH*	LH*	↑LH*	LH* L- L%
fsf2	LH*	<	---	<	(L)H*	(L)H*	LH* L-L%	
meh2	LH*	<	LH*	<	LH*	LH* H-	LH* L-	
miz2	LH*	<	LH*	<	LH*	LH*	LH* L-L%	
mns2	LH*	<	LH*	<	(LH*)	LH* L-	LH* L-L%	

22	guHa	?aal-luh	(laa?)	bi-	talaata	SaaG
fna2	LH*	LH* H-	xxx	>	LH*	LH* L-L%
fsf2	LH*	LH* H-	xxx	>	LH*	LH* L-L%
meh2	LH*	LH*	LH* L-	>	LH*	LH* L-L%
miz2	LH*	LH*	LH*	<	LH*	!LH* L-L%
mns2	LH*	LH*	xxx	>	LH*	!LH* L-L%

23	il	bayyaa9	buSS	li-	guHa	kidda	wa	?aal-luh
fna2	>	LH*	LH*	<	LH* ~~	LH*	<	LH* H-
fsf2	>	LH*	LH*	< ~	LH*	LH*	<	LH* H-
meh2	>	LH*	LH*	<	LH* ~	LH*	<	LH* H-
miz2	>	LH*	LH*	<	LH*	LH*	<	LH* H-
mns2	>	(LH*)	LH*	<	LH*	LH*	<	LH* H-

24	?ey	ra?y	-ak	?inta	raagil	Tayyib	wa	?ana	Habbayt-ak	min	?awwal	ma	šuft-ak
fna2	LH*	<	<	>	LH*	LH*	<	LH*	LH* L-	>	↑LH*	>	LH* L-L%
fsf2	LH*	<	<	<	LH*	LH* H-	>	>	LH*	<	LH*	<	!LH* L-L%
meh2	LH*	LH*	< L-L%	>	LH*	!LH* L-	>	LH* H- ~	LH*	<	LH*	<	LH* L-L%
miz2	>	LH*	<	<	LH*	LH*	<	<	LH* ~	<	LH*	<	!LH* L-L%
mns2	LH*	LH*	< L-L%	<	LH*	LH* H-	<	LH*	LH*	<	LH*	<	!LH* L-L%

(A.14) Transcription results for read narrative (continued)

25	?ana	9awz	?awzin	lak	kiilu	bi	balaaš
fna2	>	↑LH*	LH*	<	LH*	<	!LH* L-L%
fsf2	>	LH*	LH*	<	LH*	<	!LH* L-L%
meh2	>	LH*	LH*	<	LH*	<	!LH* L-L%
miz2	>	LH*	LH*	<	LH* =	<	LH* L-L%
mns2	>	LH*	LH*	<	LH*	<	!LH* L-L%

26	guHa	naTT	min	il-	kursi	illi	(huwwa)	kaan	?aa9id	9alayh	wa	?aal-luh
fna2	LH*	LH*	<	xxx	LH*	---		LH*	LH*	LH*	<	LH* H-H%
fsf2	LH*	LH*	<	<	LH*	<	LH*	<	LH*	LH*	<	LH* H-
meh2	(LH*)?	LH*	<	<	LH* H-	<	LH*	<	LH*	LH*	<	LH* H-
miz2	LH*	LH*	<	<	LH*	<	LH*	<	LH*	LH*	<	LH* H-
mns2	LH*	LH*	<	<	LH* H-	<	LH*	<	LH*	LH* H-	>	LH* H-H%

27	itniin	kiilu	itniin	kiilu	?aw9i	tiftikir	?innak	Ha-	tiDHak	9alayya
fna2	↑LH*	→LH*	→LH*	→LH*	→LH*	xxx	xxx	xxx	LH*	!LH* L-L%
fsf2	LH*	LH* L-	LH*	LH* L-	LH*	LH*	<	<	LH*	!LH* L-L%
meh2	LH*	LH* L-	LH*	LH* L- L%	LH*	LH*	<	<	LH*	!LH* L-L%
miz2	LH*	LH*	LH*	LH*	xxx	LH*	LH*	>	LH*	LH* L-L%
mns2	LH*	LH*	LH*	!LH* L-	LH*	LH*	<	<	LH*	!LH* L-L%

(A.15) Sample transcription results for retold narrative by speaker *fna*

guHa	Tuul	9umr	-uh	9aayiš	fi	-l-	?ariyaaf
LH*	<	LH*	<	LH*	<	<	!LH* L-L%

wa	fi	marra	min	-l-	marraat	fakkar	yinzil	maSr
>	>	LH*	<	<	LH*	<	LH*	!LH* L-L%

laakin	?aSHaab	-uh	?awwal	ma	9arifuu	kida	xawwifuuh	wa	?aaluu-luh
LH*	LH*	<	LH*	<	LH*	LH* H-	LH*	<	!LH* L-L%

xalli	baalik	ya	guHa	in-naas	il-	bayyaa9iin	bitu9	maSr	waHšiin	?awwi
>	LH*	<	LH* L_	LH*	<	LH*	<	LH*	LH*	LH* L-H% (??)

?awwal	ma	yisma9u	l-laHga	bita9tak	HayGallu	9alayk	it-taman	bi-n-nuSS
↑LH*	<	LH*	LH*	↑LH* H-	↑LH*	LH*	LH*	!LH* L-L%

ya9ni	masalan	law	?inta	9aayiz		law	?inta	giit	tištiri	Haaga	min	bayyaa9
LH*	LH*	<	LH*	LH* ~		>	:H*	LH*	LH*	<	<	LH* H-

wa	?aal-lak	la?	da	bi- itnaašir	?irš	?ul-luh	9ala	Tuul	bi	sitta	SaaG
>	LH*	LH*	--- =	LH*	!LH* L-L%	↑LH*	>	LH*	<	LH*	!LH* L-L%

li-?anna	humma	biyGallu	il-	Haaga	iD-Da9af
>	LH*+	LH*	<	LH*	↑H* L-L%

da	wa	law	?inta	Habbayt	kamaan	taaxud	Haaga	min	Hadd	?ul-luh	iD-Da9f
>	>	>	LH*	LH*	LH*	(LH*)	LH*	<	↑LH* H-	↑LH*	!LH* L-L%

(A.15) Sample transcription results for retold narrative by speaker *fna* (continued)

raaH	guHa	?aal-lu-hum	mataxafuuš	9alayya		raaH	naazil	maSr
LH*	LH*	LH*	LH*	!LH* L-L%		LH*	LH* ~	!LH* L-L%

wa	?a9ad	9ala	?ahwa	min	il-?ahaawi	?ašaan	yišrab	šayy
>	↑LH*	<	LH*	<	LH*	<	LH*	LH* L-L%

fii	waaHid	bitaa9	mooz	kaan	maaddi
>	↑LH*	<	LH*	<	!LH* L-L%

9aayiz	yi?uul	il	mooz	il	Helw	il-	mooz	il	Helw
LH*	LH*	>	LH*	<	↑LH*L-	>	LH*	<	LH*L-L%

raaH	guHa	nada	9alayh	wa	?aal-luh		bikam	kiilu	l-	mooz
LH*	↑LH*	LH*	LH*	>	!LH* L-L%		LH*	→LH*	<	LH* L-L%

?aal	luh	bi-	9ašara	SaaG		?aal-luh	la?	bi	xamsa
LH*	<	<	LH*	!LH* L-L%		LH*	LH*	<	!LH* L-L%

?aal-luh	?ašaan	xaTr-ak	?inta	?ana	Ha-	adda-hu-	lak	bi	xamsa
↑LH*	<	↑LH*	↑LH*	<	<	LH*	<	<	!LH* L-L%

raaH	?aal-luh	la?	bi-	?itneen	wa	nuSS
>	LH*	<	>	LH*	<	!LH* L-L%

?aal-luh	wa	raaH	il	bayyaa9	?aal-luh	manta	lissa	?aayil	min	da?ii?a	waHda	bi	xamsa	SaaG
↑LH* =	>	LH*	<	↑LH*	LH*	LH*	↑LH*	↑LH*	<	↑LH*	↑LH* L-L%	>	↑LH*	↑LH* L-L%

(A.15) Sample transcription results for retold narrative by speaker *fna* (continued)

?aal-luh	la?	huwwa	bi	itneen	nuSS
LH*	<	↑LH*	<	LH*	!LH* L-L%

raaH	il	bayyaa9	?aal-luh	lii	wa	?ana	saari?-uh	walla	la?iy	fi-š	šaari9
LH*	<	LH*	LH* L-	LH*	<	LH*	LH*	<	LH*	<	LH* H-H%

?aal-luh	huwwa	itneen	wa	nuSS		?in	kaan	9agb-ak
LH*	↑LH*	LH*	<	!LH* L-L%		>	LH*	!LH* L-L%

raaH	?aal-luh	?inta	šakl-ak	raagil	Tayyib	wa	Galbaan	?ana	Haddii-k	kiilu	bi	balaaš
LH*	LH*	LH* =	LH*	LH*	LH*	<	LH* H-	---	LH* =	LH*	<	!LH* L-L%

?aal-luh	la?	itneen	kiilu	?inta	ma	(fa)	ma	tiDHAk	-š	9alayya
LH*	<	LH*	!LH* L-L%	>	>	>	>	LH*	<	!LH* L-L%

Appendix B (relates to chapter 4)

(B.1) Full test sentences containing target words with word-medial stress.

TargetID

212119 il-mudarris **mimalmil** min iT-Talaba
the-teacher nervous from the-students
"The teacher is nervous of his students"

212120 il-walad da **minamrad** xaaliS ma-šuft-iš zayy-uh
the-boy that rebellious completely NEG-I saw-NEG like-him
"That boy is completely uncontrollable, I've never seen anything like it"

212121 9amm-i **mimangih** nafsuh ?awwi ba9d-ma gay min barra
my-uncle boastful himself very after-that he-came from outside
"My uncle is a bit full of himself since he came back from abroad".

212122 ibn-i **mitmanZar** xaaliS il-yoomayn dool
my-son naughty completely the-two-days these
"My son has been a real handful these last few days"

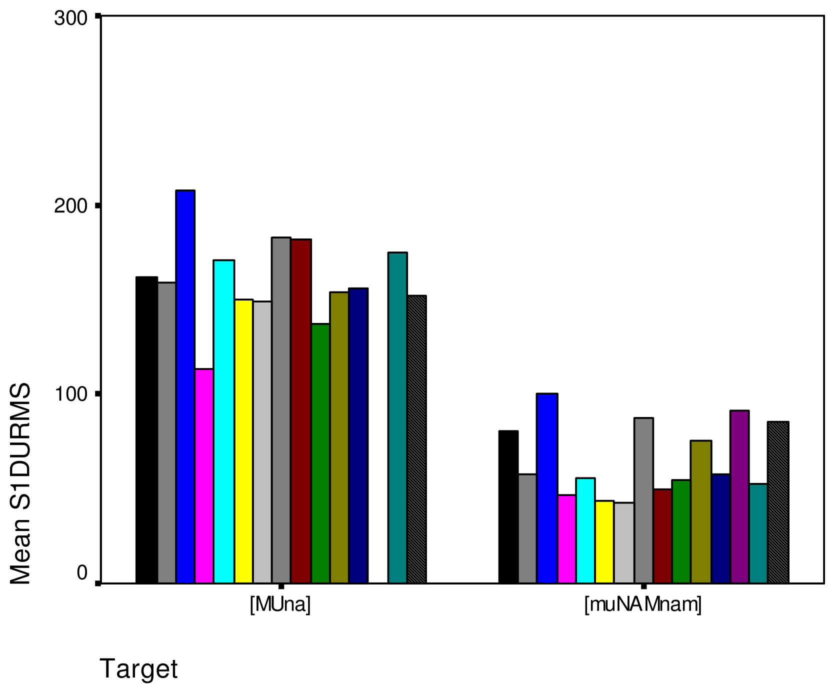
212123 šuf-na nuunu **mitnamnim** xaaliS fil-mustašfa
we-saw baby tiny/cute completely in-the-hospital
"We saw a really tiny baby in the hospital."

212224 bu?? muna **munamnim** xaaliS wa ša9riha Tawiil
mouth Muna tiny/cute completely and hair-her long
"Muna's mouth is tiny and her hair is long".

(Target word is in bold type; stressed syllables are underlined)

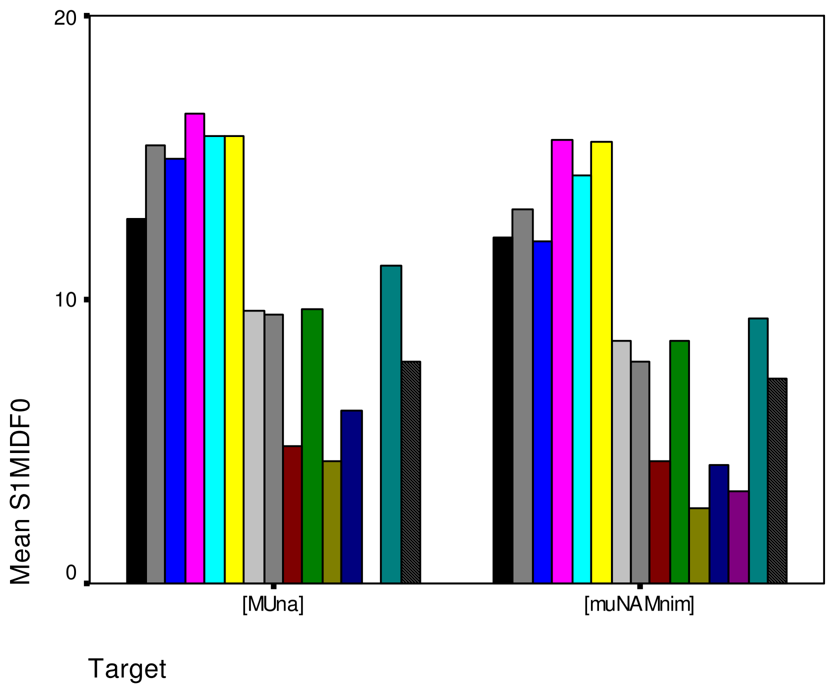
(B.2) **Phonetic correlates of word-level prominence by speaker.**

Bar chart: mean **duration** (in milliseconds) in stressed vs. unstressed S1 by speaker.



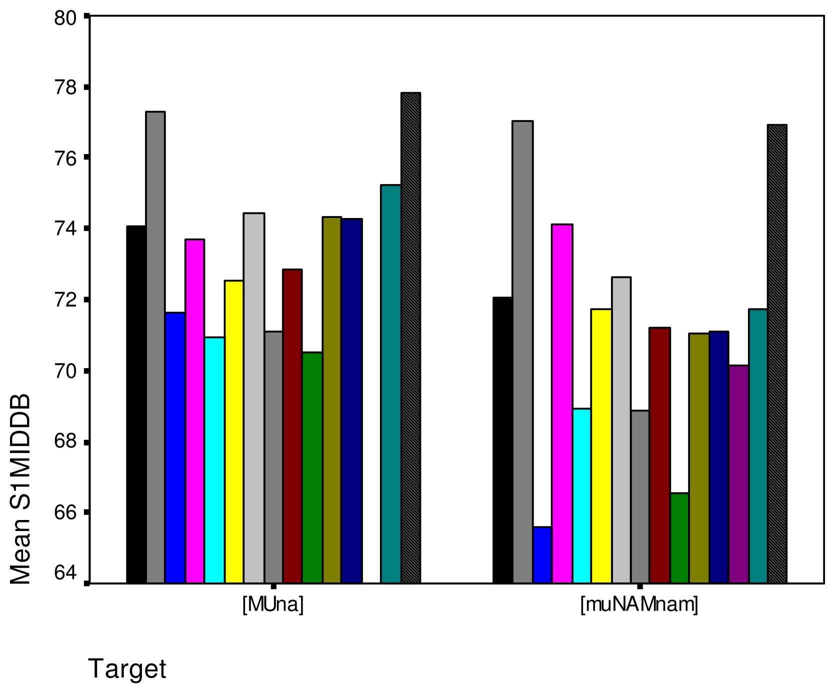
(B.3) **Phonetic correlates of word-level prominence by speaker.**

Bar chart: mean **F0** (in semitones) in stressed vs. unstressed S1 by speaker.



(B.4) Phonetic correlates of word-level prominence by speaker.

Bar chart: mean **intensity** (in decibels) in stressed vs. unstressed S1 by speaker.



Appendix C (relates to chapter 7)

C.1 Comparison of descriptive statistics for various measures of alignment.

	L1-C0	L1-V0	H-C1	H-V1	peak delay	rise duration	sylldur#1	sylldur#2	ftdur	pwddur	RPD#1	RPD#2	RPD#3	RPD#4	F0 change (sem)
N	709	709	709	709	709	709	709	709	709	709	709	709	709	709	709
Mean	11.486	-53.4	23.50	-58.4	169.845	158.35	171.25	186.04	217.88	324.56	1.042	.932	.826	.536	3.215
Median	3.665	-55.8	22.31	-62.21	162.117	150.80	164.84	182.93	204.29	320.49	.943	.888	.788	.511	2.990
Skewness	.948	.473	.247	-.147	1.163	.971	.695	.645	.828	.605	.546	.833	1.101	.542	.931

Definitions:

peak delay H- C0
rise duration H-L

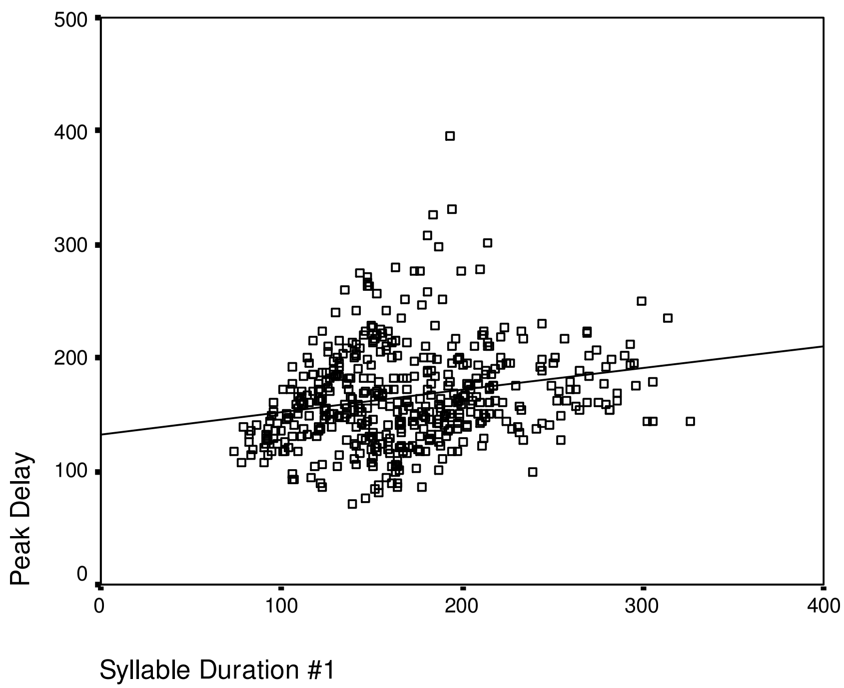
sylldur#1 Syllable Duration #1 (canonical syllable)
sylldur#2 Syllable Duration #2 (‘ambisyllabic’ syllable)
ftdur Foot Duration
pwddur Prosodic Word Duration

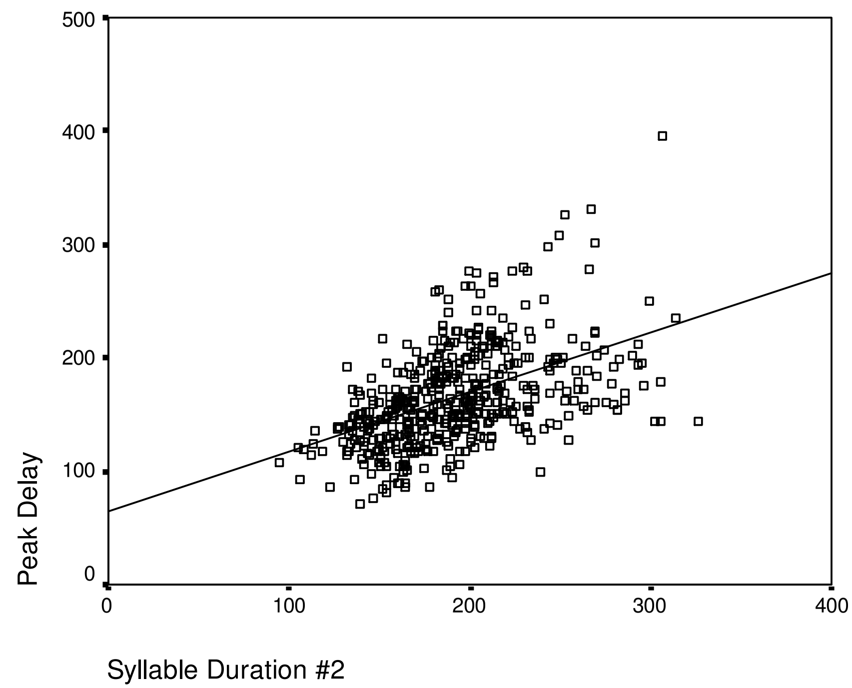
RPD#1 peak delay / Syllable Duration #1
RPD#2 peak delay / Syllable Duration #2
RPD#3 peak delay / Foot Duration
RPD#4 peak delay / PWd Duration

F0 change F0 at H - F0 at L (semitones)

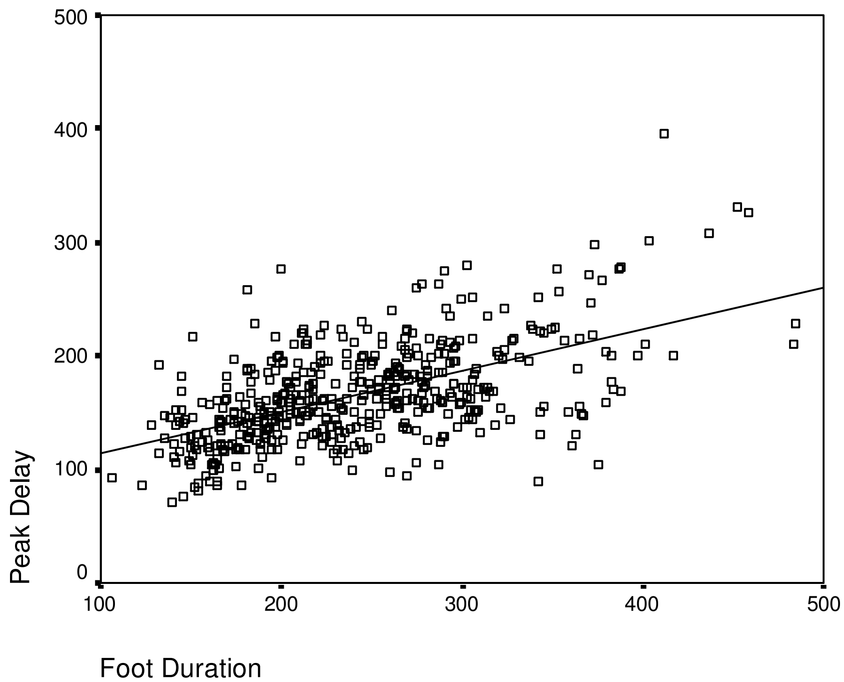
C.2 Scatter plots: peak delay x four different structural domain durations.

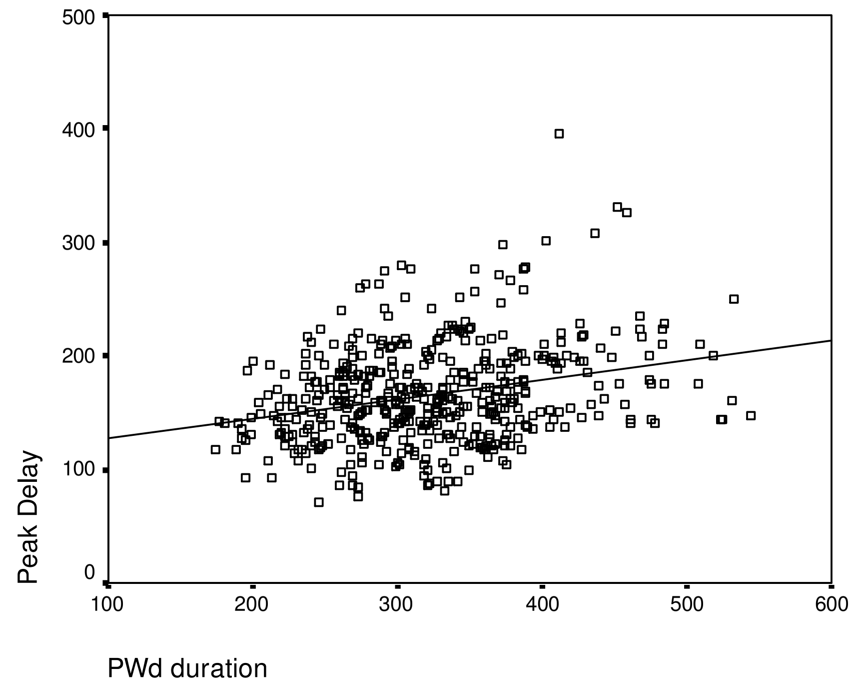
The plots show correlation between the two variables, in all four measures of structural domain, which is *compatible* with the ‘structural anchoring hypothesis’. Sylldur#1 = 'canonical syllable; Sylldur#2 = ambisyllabic syllable.





C.2 Scatter plots: peak delay x four different structural domain durations (continued).





Appendix D (relates to chapter 8)

(D.1) Context paragraphs used in the focus experiment - lexical set A (target sentences in bold type).

A1	زميلي قالي انه سمع ان بابايا بيروح الجامعة بالليل لكن انا قولتله: لا، **ماما بتتعلّم يوناني بالليل**. بابا بيقعد في البيت وبيتفرج على التليفزيون.
[+F+f]	zamiil-i ?aal-li inn-uh simi9 ?in babaaya biyruuH il-gama9a bil-layl laakin ?ana ?ulti-luh la?. **maama bitit9allim yunaani bil-layl**. baaba biyu?9ud fil-bayt yitfarrag 9ala t-tilivizyoon
	my-colleague said-me that-he he-heard that my-dad he-goes the-university at-night but I said-him no. **Mum she-learns Greek at-night**. Dad he-sits in-the-house he-watches-at the-television
	'My colleague said they heard my dad went to university in the evenings but I told him no. **Mum is learning Greek in the evenings**. Dad sits at home and watches TV.'
A2	أمي بتحب تتعلم حاجات جديدة. **ماما بتتعلّم يوناني بالليل**، و تدرس تاريخ كمان.
[-F+f]	?umm-i bitHebb tit9allim Haagaat gidiida. **maama bitit9allim yunaani bil-layl**. wa tidris tariix kamaan.
	my-mother she-loves she-learns things new-pl. **Mum she-learns Greek at-night**. And she-studies history also.
	'My mother loves learning new things. **Mum is learning Greek in the evenings** and she also studies history.'
A3	زميلي قالي انه سمع ان بابايا بيتعلّم يوناني بالليل لكن انا قولتله: لا، **ماما بتتعلّم يوناني بالليل**. بابا بيقعد في البيت وبيتفرج على التليفزيون.
[+F-f]	zamiil-i ?aal-li inn-uh simi9 ?in babaaya biyit9allim yunaani bil-layl laakin ?ana ?ulti-luh la?. **maama bitit9allim yunaani bil-layl**. baaba biyu?9ud fil-bayt yitfarrag 9ala t-tilivizyoon.
	my-colleague said-me that-he heard that my-dad he-learns Greek at-night but I said-him no. **Mum she-learns Greek at-night**. Dad he-sits in-the-house he-watches-at the-television
	'My colleague said they heard my dad was learning Greek in the evenings but I told him no. **Mum is learning Greek in the evenings**. Dad sits at home and watches TV.'
A4	أمي بتحب اليوناني. ماما بتتعلّم يوناني بالليل، و بتحب تتفرج على افلام للتاريخ اليوناني.
[-F-f]	?umm-i bitHebb il-yunaani. **maama bitit9allim yunaani bil-layl**. wa bitHebb titfarrag 9ala ?aflaam it-tariix il-yunaani.
	my-mother she-loves the-Greek.. **Mum she-learns Greek at-night**. And she-loves she-watches-at films the-history the-Greek.
	'My mother loves Greek. *Mum is learning Greek in the evenings* and she likes to watch films on Greek history.'

(D.2) Comparison of mean trigger F0 excursion (xxn): post-hoc test results ALL SPEAKERS.

						95% Confidence Interval	
	(I) Focus Condition	(J) Focus Condition	Mean Difference (I-J)	Std. Error	Sig.	Lower Bound	Upper Bound
Tamhane	+F+f	-F+f	.658319	.3481269	.322	-.284646	1.601284
		+F-f	-.203899	.4004916	.997	-1.290376	.882578
		-F-f	.042212	.3155137	1.000	-.814180	.898604
	-F+f	+F+f	-.658319	.3481269	.322	-1.601284	.284646
		+F-f	-.862218	.3946476	.180	-1.933666	.209229
		-F-f	-.616107	.3080619	.264	-1.452421	.220206
	+F-f	+F+f	.203899	.4004916	.997	-.882578	1.290376
		-F+f	.862218	.3946476	.180	-.209229	1.933666
		-F-f	.246111	.3662011	.985	-.752498	1.244719
	-F-f	+F+f	-.042212	.3155137	1.000	-.898604	.814180
		-F+f	.616107	.3080619	.264	-.220206	1.452421
		+F-f	-.246111	.3662011	.985	-1.244719	.752498

(D.3) Comparison of mean trigger F0 excursion (xxn): post-hoc test results FEMALE SPEAKERS ONLY. Multiple Comparisons

Dependent Variable: trigger F0 excursion

	(I) Focus Condition	(J) Focus Condition	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Tukey HSD	+F+f	-F+f	1.057484(*)	.3967654	.046	.012517	2.102452
		+F-f	-.677406	.3967654	.328	-1.722373	.367562
		-F-f	.472483	.3967654	.635	-.572485	1.517450
	-F+f	+F+f	-1.057484(*)	.3967654	.046	-2.102452	-.012517
		+F-f	-1.734890(*)	.3967654	.000	-2.779858	-.689922
		-F-f	-.585002	.3967654	.458	-1.629970	.459966
	+F-f	+F+f	.677406	.3967654	.328	-.367562	1.722373
		-F+f	1.734890(*)	.3967654	.000	.689922	2.779858
		-F-f	1.149888(*)	.3967654	.025	.104920	2.194856
	-F-f	+F+f	-.472483	.3967654	.635	-1.517450	.572485
		-F+f	.585002	.3967654	.458	-.459966	1.629970
		+F-f	-1.149888(*)	.3967654	.025	-2.194856	-.104920

* The mean difference is significant at the .05 level.

References

Abdalla, A.G. 1960. An instrumental study of the intonation of Egyptian Colloquial Arabic. University of Michigan. Thesis/Dissertation

Abdel-Massih, E.T. 1975. An introduction to Egyptian Arabic. Ann Arbor, University of Michigan.

Abu-Salim, I.M. 1983. A re-analysis of some aspects of Arabic phonology: a metrical approach. Ann Arbor, U.M.I.

Al-Ani, S. 1970. Arabic phonology: an acoustical and physiological investigation. The Hague, Mouton.

Al-Ani, S. 1992. Stress variation of the construct phrase in Arabic: a spectrographic analysis. Anthropological Linguistics, 34, 256-276.

Al-Harbi, L. 1991. Formal analysis of intonation: the case of the Kuwaiti dialect of Arabic. Herriot-Watt University. Thesis/Dissertation

Al-Khalifa, F. 1984. An intonational study of Kuwaiti Bedouin Arabic. SOAS, University of London. Thesis/Dissertation

Al-Nassir, E.S. 1993. Sibawayh the phonologist. London, Kegan Paul.

Alhawary, M.T. 2003. Elicitation techniques and considerations in data collection in Early Arabic Grammatical Tradition. Journal of Arabic Linguistics Tradition, 1, 1-24.

Anttila, A. & Bodomo, A.B. 2000. Tonal polarity in Dagaare. *In* Trends in African Linguistics 4: Advances in African linguistics. V.Carstens and F.Parkinsons, eds. pp. 119-134. Trenton,N.J., African World Press.

Arvaniti, A. & Baltazani, M. 2005. Intonational analysis and prosodic annotation of Greek spoken corpora. *In* Prosodic transcription and typology: a unified approach. S.-A.Jun, ed. pp. 84-117. Oxford, OUP.

Arvaniti, A. & Ladd, D.R. 1995 Tonal alignment and the representation of accentual targets. Elenius, K. and Branderad, R. 4, 220-223. Proceedings of the 13th International Congress of Phonetic Sciences.

Arvaniti, A., Ladd, D.R., & Mennen, I. 1998. Stability of tonal alignment: the case of Greek prenuclear accents. Journal of Phonetics, 26, 3-25.

Arvaniti, A., Ladd, D.R., & Mennen, I. 2000. What is a starred tone? Evidence from Greek. *In* Papers in Laboratory Phonology V. M.Broe and J.Pierrehumbert, eds. Cambridge, CUP.

Ashby, M.G. 1978. A study of two English nuclear tones. Language and Speech, 21, 326-336.

Atterer, M. & Ladd, D.R. 2004. On the phonetics and phonology of "segmental anchoring" of F0: evidence from German. Journal of Phonetics, 32, 177-197.

Badawi, E.S. 1965. An intonational study of colloquial Riyadhi Arabic. UCL. Thesis/Dissertation

Beckman, M. 1986. Stress and non-stress accent. Dordrecht, Foris.

Beckman, M. 2004 Tone inventories and tune-text alignments (paper presented at the ESF Workshop on Typology of Tone and Intonation, Cascais, Portugal.. April 2004).

Beckman, M. & Edwards, J. 1994. Articulatory evidence for differentiating stress categories. *In* Phonological structure and phonetic form: paper in Laboratory Phonology III. P.Keating, ed. pp. 7-33. Cambridge, CUP.

Beckman, M. & Elam, G.A. 1993 Guidelines for TOBI Labelling (version 3.0 1997). The Ohio State University Research Foundation.

Beckman, M. & Pierrehumbert, J. 1986. Intonational structure in Japanese and English. Phonology Yearbook, 3, 255-309.

Benkirane, T. 1998. Intonation in Western Arabic (Morocco). *In* Intonation systems: a survey of twenty languages. D.Hirst and A.Di Cristo, eds. pp. 345-359. Cambridge, MA, CUP.

Benmamoun, E. 2000. The feature structure of functional categories. Oxford, OUP.

Boersma, P. & Weenink, D. 2004 Praat: doing phonetics by computer (Version 4.2).

Bolinger, D. 1972. Accent is predictable (if you're a mind-reader). Language, 48, 633-644.

Borer, H. 1996. The construct in review. *In* Studies in Afroasiatic Grammar. Lecarme, J.Lowenstamm, and U.Shlonsky, eds. pp. 30-61. The Hague, HAG.

Borowsky, T., Itô, J., & Mester, R.-A. 1984. The formal representation of ambisyllabicity: evidence from Danish. NELS, 14, 34-48.

Boudlal, A. 2001. Constraint interaction in the phonology and morphology of Casablanca Moroccan Arabic. Université Mohammed V, Rabat, Morocco. Thesis/Dissertation

Broselow, E. 1976. The phonology of Egyptian Arabic. UMass (Amherst). Thesis/Dissertation

Brown, G. 1983. Intonation, the categories given/new and other sorts of knowledge. *In* Prosody: models and measurements. A.Cutler and D.R.Ladd, eds. Berlin, Springer-Verlag.

Bruce, G. 1977. Swedish word accents in sentence perspective. Lund, Gleerup.

Bruce, G. 1982. Developing the Swedish intonation model. Working Papers, Department of Linguistics and Phonetics, University of Lund, 22, 51-116.

Brustad, K. 2004. The syntax of spoken Arabic: a comparative study of Moroccan, Egyptian, Syrian and Kuwaiti dialects. Washington, Georgetown University Press.

Chafe, W. 1974. Language and consciousness. Language, 50, 111-133.

Chahal, D. 2001. Modeling the intonation of Lebanese Arabic using the autosegmental-metrical framework: a comparison with English. University of Melbourne. Thesis/Dissertation

Chapman, S. 1996. Some observations on metalinguistic negation. Linguistics: An Interdisciplinary Journal of the Language Sciences, 32, 387-402.

Chen, M. 1990. What must phonology know about syntax? *In* The phonology-syntax connection. S.Inkelas and D.Zec, eds. pp. 19-46. Chicago, University of Chicago Press.

Cho, Y.-M.Y. 1990. Syntax and phrasing in Korean. *In* The phonology-syntax connection. S.Inkelas and D.Zec, eds. pp. 47-62. Chicago, CSLI.

Chomsky, N. 2001. On derivation by phase. *In* Ken Hale: a life in language. M.Kenstowicz, ed. pp. 1-52. Cambridge, Mass., MIT Press.

Chomsky, N. & Halle, M. 1968. The sound pattern of English. New York, Harper & Row.

Clements, G.N. 1978. Tone and syntax in Ewe. *In* Elements of tone, stress and intonation. D.J.Napoli, ed. pp. 21-99. Washington D.C., Georgetown University Press.

Cooper, W. & Paccia-Cooper, J. 1980. Syntax and Speech. Cambridge, MA, Harvard University Press.

Cruttenden, A. 1997. Intonation. Cambridge, CUP.

Cruttenden, A. 2006. The de-accenting of old information: a cognitive universal. *In* Pragmatic organization of discourse in the languages of Europe. G.Bernini and M.L.Schwartz, eds. Berlin, Mouton de Gruyter.

Crystal, D. 1969. Prosodic systems and intonation in English. Cambridge, CUP.

D'Imperio, M. 1997. Narrow focus and focal accent in the Neapolitan variety of Italian. *In* Intonation: theory, models and applications. proceedings of an ESCA workshop (Athens). A.Botinis, G.Kouroupetroglou, and G.Carayannis, eds. pp. 87-90. Athens, ESCA/University of Athens.

de Jong, K. & Zawaydeh, B. 2002. Comparing stress, lexical focus and segmental focus: patterns of variation in Arabic vowel duration. Journal of Phonetics, 30, 53-75.

de Jong, K. & Zawaydeh, B.A. 1999. Stress, duration, and intonation in Arabic word-level prosody. Journal of Phonetics, 27, 3-22.

deLacy, P. 1999 Tone and prominence. [ROA-333].

deLacy, P. 2002. The interaction of tone and stress in Optimality Theory. Phonology, 19, 1-32.

deLacy, P. 2004. Markedness conflation in Optimality Theory. Phonology, 21, 145-199.

Downing, L. 2004. What African languages tell us about accent typology. ZAS Papers in Linguistics, 37, 101-136.

Downing, L. 2006. Canonical Forms in Prosodic Morphology. Oxford., OUP.

Dresher, E. 1994. The prosodic basis of Tiberian Hebrew system of accents. Language, 70, 1-52.

El Zarka, D. 1997. Prosodische Phonologie des Arabischen. Karl-Franzens-Universität Graz. Thesis/Dissertation

Elenbaas, N. & Kager, R. 1999. Ternary rhythm and the lapse constraint. Phonology, 16, 273-330.

Elordieta, G., Frota, S., Prieto, P., & Vigario, M. 2003 Effects of constituent weight and syntactic branching on phrasing decisions in Romance. 487-490. Proceedings of the 15th International Congress of Phonetic Sciences.

Emonds, J. 1976. A transformational approach to English syntax. New York, Academic Press.

Face, T. 2002. Intonational marking of contrastive focus in Madrid Spanish. Lincom Europa.

Face, T. 2003. Intonation in Spanish declaratives: differences between lab speech and spontaneous speech. Catalan Journal of Linguistics, 2, 115-131.

Frascarelli, M. 2000. The syntax-phonology interface in focus and topic constructions in Italian. Dordrecht, Kluwer.

Frota, S. 2000. Prosody and focus in European Portuguese: phonological phrasing and intonation. New York, Garland.

Fujisaki, H. & Sudo, H. 1971. A generative model for the prosody of connected speech in Japanese. Annual Report of the Engineering Research Institute, University of Tokyo, 30, 75-80.

Gary, J.O. & Gamal-Eldin, S. 1981. Cairene Egyptian colloquial Arabic. London, Croom Helm.

Ghalib, G. 1977. The intonation of colloquial Iraqi spoken Arabic. University of Leeds. Thesis/Dissertation

Ghazali, S., Hamdi, R., & Barkat, M. 2002 Speech rhythm variation in Arabic dialects. Aix-en-Provence. Proceedings of Prosody 2002.

Ghini, M. 1993. Phonological phrase formation in Italian: a new proposal. Toronto Working Papers in Lingusitics, 12, 41-77.

Göksel, A. & Özsoy, A.S. 2000. Is there a focus position in Turkish? *In* Studies on Turkish and Turkic Languages. A.Göksel and C.Kerslake, eds. pp. 219-228. Wiesbaden, Harrasowitz.

Göksel, A. & Özsoy, A.S. 2003. da: a focus/topic associated clitic in Turkish. Lingua: International Review of General Linguistics, 113, 1143-1167.

Goldsmith, J. 1976. Autosegmental Phonology. Bloomington, Indiana University Linguistics Club.

Goldsmith, J. 1987. Tone and accent and getting the two together. BLS, 13, 88-104.

Grabe, E., Nolan, F., & Farrar, K. 1998. IViE - a comparative transcription system for intonational variation in English. Proceedings of the 5th International Conference on Spoken Language Processing (Sydney, Australia)., 1259-1262.

Grabe, E. & Low, E.L. 2004. Durational variability in speech and the rhythm class hypothesis. *In* Papers in Laboratory Phonology 7. Cambridge, CUP.

Grice, M. 1995a. Leading tones and downstep in English. Phonology, 12, 183-233.

Grice, M. 1995b. The intonation of interrogation in Palermo Italian: implications for intonation theory. Tubingen, Niemeyer.

Grice, M., D'Imperio, M., Savino, M., & Avesani, C. 2005. Strategies for intonation labelling across varieties of Italian. *In* Prosodic typology: the phonology of intonation and phrasing. S.-A.Jun, ed. pp. 362-389. Oxford, OUP.

Grice, M., Ladd, D.R., & Arvaniti, A. 2000. On the place of phrase accents in intonational phonology. Phonology, 17, 143-185.

Grønnum, N. 1983. Two issues in the prosody of Standard Danish. *In* Prosody - Models and Measurements. A.Cutler and D.R.Ladd, eds. pp. 27-38. Berlin, Springer.

Guindy, A.-K. 1988. On the stress in the 'madrasa' word structure in Cairene Colloquial Arabic. Zeitschrift für arabischen Linguistik, 18, 33-58.

Gussenhoven, C. 1980. Focus, mode and the nucleus. Journal of Linguistics, 19, 377-417.

Gussenhoven, C. 1983. On the grammar and semantics of sentence accents. Dordrecht, Foris.

Gussenhoven, C. 2000. The lexical tone contrast of Roermond Dutch in Optimality Theory. *In* Prosody, Theory and Experiment: studies presented to Gösta Bruce. M.Horne, ed. pp. 129-167. Dordrecht, Kluwer.

Gussenhoven, C. 2004. The phonology of tone and intonation. Cambridge, CUP.

Gussenhoven, C. 2005. Transcription of Dutch Intonation. *In* Prosodic Typology. S.-A.Jun, ed. pp. 118-145.

Gussenhoven, C. & Rietveld, A.C.M. 1992. Intonation Contours, Prosodic Structure and Preboundary Lengthening. Journal of Phonetics, 20, 283-303.

Gussenhoven, C. & Rietveld, T. 2000. The behavior of H* and L* under variations in pitch range in Dutch rising contours. Language and Speech, 43, 183-203.

Hale, K. & Selkirk, E.O. 1987. Government and tonal phrasing in Papago. Phonology Yearbook, 4, 151-183.

Hall, T.A. 1989. Lexical Phonology and the distribution of German [ç] and [x]. Phonology, 6, 1-17.

Halle, M. & Vergnaud, J.-R. 1987. An essay on stress. Cambridge, Mass., MIT Press.

Halliday, M.A.K. 1967. Intonation and grammar in British English. The Hague, Mouton.

Halliday, M.A.K. 1970. A course in spoken English: intonation. Oxford, OUP.

Hammoumi, A. 1988. Syllabation, accentuation, éffacement et abrègement dans le parler de Fès et d'autres parlers marocains. Faculté de Lettres, Rabat. Thesis/Dissertation

Harrell, R. 1957. The phonology of colloquial Egyptian Arabic. American Council of Learned Societies.

Harris, J. 2004. Release the captive coda: the foot as domain of phonetic interpretation. *In* Phonetic interpretation: papers in laboratory phonology VI. J.Local, R.Ogden, and R.Temple, eds. pp. 103-129. Cambridge., CUP.

Hayes, B. 1981. A metrical theory of stress rules. New York, Garland Press.

Hayes, B. 1989. The prosodic hierarchy in meter. *In* Rhythm and meter. P.Kiparsky and G.Youmans, eds. pp. 201-260. Orlando, Academic Press.

Hayes, B. 1994. "Gesture" in prosody: comments on the paper by Ladd. *In* Phonological structure and phonetic form: Papers in Laboratory Phonology III. P.Keating, ed. pp. 64-75. Cambridge, CUP.

Hayes, B. 1995. Metrical stress theory: principles and case studies. Chicago, University of Chicagoe Press.

Hayes, B. & Lahiri, A. 1991. Bengali intonational phonology. Natural Language & Linguistic Theory, 9, 47-96.

Heliel, M. 1977. The rhythm of Egyptian collquial Arabic: an experimental study. UCL. Thesis/Dissertation

Hellmuth, S. 2002a A preliminary investigation into the intonation of Cairo Arabic: focus strategies in declarative sentences. [paper submitted for credit LOT Summer School 2002 "Intonational Structure"].

Hellmuth, S. 2002b Sentence accent in spoken Arabic dialects: revealing (the lack of) deaccentuation in Cairene Arabic. (m.s.).

Hellmuth, S. 2004. Prosodic weight and phonological phrasing in Cairene Arabic. Proceedings of Annual Meeting of Chicago Linguistic Society, 40.

Hellmuth, S. 2005. No De-accenting in (or of) Phrases: Evidence from Arabic for cross-linguistic and cross-dialectal prosodic variation. *In* Prosodies (with special reference to Iberian languages). S.Frota, M.Vigario, and M.J.Freitas, eds. pp. 99-121. Mouton de Gruyter.

Holes, C. 1995. Modern Arabic: structures, functions and varieties. Longman.

Hualde, J.I. 2002. Intonation in Spanish and the other Ibero-Romance languages: overview and status quaestionis. *In* Romance Phonology and Variation. C.Wiltshire and J.Camps, eds. pp. 101-115. Amsterdam, John Benjamins.

Hyman, L. 2001. Tone systems. *In* Language typology and language universals: an international handbook. Vol. 2. M.Haspelmath, E.König, H.E.Wiegand, and H.Steger, eds. pp. 1367-1380. Berlin, Mouton de Gruyter.

Ibn-Jinni. 1986. Al-Khaºaa'iº. Cairo, Al-Hay'ah Al-Miºriyya Al-'Amma li-al-Kitaab.

Ibrahim, O.A.G., El-Ramly, S.H., & Abdel-Kader, N.S. 2001 A model of F0 contour for Arabic affirmative and interrogative sentences. 517-524. Mansoura University, Egypt. 18th National Radio Science Conference.

Ingham, B. 1994. Najdi Arabic: Central Arabian. Amsterdam, John Benjamins.

Inkelas, S. & Zec, D. 1990. The phonology-syntax connection. Chicago, University of Chicago Press.

Inkelas, S. & Zec, D. 1995. Syntax-phonology interface. *In* The handbook of phonological theory. J.Goldsmith, ed. pp. 535-549. Oxford, Blackwell.

Ishihara, T. 2003. A phonological effect on tonal alignment in Tokyo Japanese. Proceedings of the 15th ICPhS, 615-618.

Jelinek, E. 2002. Agreement, clitics and focus in Egyptian Arabic. *In* Themes in Arabic and Hebrew Syntax. J.Ouhalla and U.Shlonsky, eds. pp. 71-105.

Jensen, J.T. 2000. Against ambisyllabicity. Phonology, 17, 187-235.

Jun, S.-A. 1993. The phonetics and phonology of Korean prosody. Ohio State University. Thesis/Dissertation

Jun, S.-A. 1996. The phonetics & phonology of Korean prosody: intonational phonology and prosodic structure. New York, Garland.

Jun, S.-A. 2003. The effect of phrase length and speech rate on prosodic phrasing. Proceedings of the 15th ICPhS, 483-486.

Jun, S.-A. 2004 The Intermediate Phrase in Korean intonation: Evidence from sentence processing data. (Paper presented at Tone & Intonation in Europe, Santorini, 9-11 Sept 2004).

Jun, S.-A. 2005a. Korean intonational phonology and prosodic transcription. *In* Prosodic Typology. S.-A.Jun, ed. pp. 201-229. Oxford, OUP.

Jun, S.-A. 2005b. Prosodic typology. *In* Prosodic transcription and typology: a unified approach. S.-A.Jun, ed. pp. 430-458. Oxford, OUP.

Jun, S.-A. 2005c. Prosodic Typology: The Phonlgy of Intonation and Phrasing. Oxford, OUP.

Karins, K., Liberman, M., McLemore, C., & Rowson, E. 2002 CallHome Egyptian Arabic Speech Supplement LDC2002S37. Philadelphia, Linguistic Data Consortium.

Keane, E. 2004 Prominence in Tamil.

Kenstowicz, M. 1980. Notes on Cairene Arabic syncope. Studies in the Linguistic Sciences, 10, 39-53.

Kiss, K. 1998. Identificational focus versus informational focus. Language, 74, 245-273.

Kubozono, H. 1993. The organization of Japanese prosody. Tokyo, Kurosio.

Ladd, D.R. 1980. The structure of intonational meaning: evidence from English. Bloomington, Indiana University Press.

Ladd, D.R. 1994. Constraints on the gradient variability of pitch range, or, pitch level 4 lives! *In* Phonological structure and phonetic form: Papers in Laboratory Phonology III. P.Keating, ed. pp. 43-63. Cambridge, CUP.

Ladd, D.R. 1996. Intonational phonology. Cambridge, CUP.

Ladd, D.R. 2003. Phonological conditioning of F0 target alignment. Proceedings of the 15th ICPhS, 249-252.

Ladd, D.R. 2004. Segmental anchoring of pitch movements: autosegmental phonology or speech production? *In* On Speech and Language: Studies for Sieb G. Noteboom. H.Quene and V.J.van Heuven, eds. pp. 123-131. Utrecht, LOT.

Ladd, D.R. 2005 Coordination of pitch movements and segments: what does it tell us about phonology? (Paper presented at PAPI 2005, Barcelona).

Ladd, D.R., Faulkner, D., Faulkner, H., & Schepman, A. 1999. Constant "segmental anchoring" of F0 movements under changes in speech rate. Journal of the Acoustical Society of America, 106, 1543-1554.

Ladd, D.R., Mennen, I., & Schepman, A. 2000. Phonological conditioning of peak alignment in rising pitch accents in Dutch. Journal of the Acoustical Society of America, 107, 2685-2696.

Ladd, D.R. & Morton, R. 1997. The perception of intonational emphasis: continuous or categorical? Journal of Phonetics, 25, 313-342.

Langendoen, D.T. 1968. The London School of Linguistics: a study of the linguistic theories of B. Malinowski and J.R. Firth. Cambridge, Mass., MIT Press.

Liberman, M. 1975. The intonational sytem of English.   MIT. Thesis/Dissertation

Liberman, M. & Pierrehumbert, J. 1984. Intonational variance under changes in pitch range and length. *In* Language Sound Structure: Studies in Phonology present to Morris Hallé. M.Aronoff and R.T.Oehrle, eds. pp. 157-233. Cambridge, M.A., MIT Press.

Liberman, M. & Prince, A. 1977. On stress and linguistic rhythm. Linguistic Inquiry, 8, 249-336.

Lickley, R., Schepman, A., & Ladd, D.R. 2006. Alignment of 'phrase accent' lows in Dutch falling-rising questions: theoretical and methodological implications. (in press). Language and Speech.

Lindblom, B.E.F. 1990. Explaining phonetic variation: a sketch of the H & H Theory. *In* Speech production and speech modeling, NATO ASI series D: Behavioural and social sciences. H.J.Hardcastle and A.Marchal, eds. pp. 403-439. Dordrecht, Kluwer.

McCarthy, J. 1979. On stress and syllabification. Linguistic Inquiry, 10, 443-466.

McCarthy, J. 1993. A case of surface constraint violation. Canadian Journal of Linguistics/Revue Canadienne de Linguistique, 38, 169-195.

McCarthy, J. 2002. A thematic guide to Optimality Theory. Cambridge, CUP.

McCarthy, J. 2003. OT constraints are categorical. Phonology, 20, 75-138.

McCarthy, J. & Prince, A. 1990. Foot and word in Prosodic Morphology: the Arabic broken plural. Natural Language & Linguistic Theory, 8, 209-286.

McCarthy, J. & Prince, A. 1993. Generalised Alignment. *In* Yearbook of Morphology. G.Booij and J.van Marle, eds. pp. 79-153. Dordrecht, Kluwer.

McCarthy, J. & Prince, A. 1995. Faithfulness and reduplicative identity. *In* Papers in Optimality Theory. J.Beckman, L.Dickey, and S.Urbanczyk, eds. pp. 249-384. Amherst, MA, GLSA.

McCarthy, J. & Prince, A. 1996. Prosodic Morphology. *In* Essential Readings in Phonology. J.Goldsmith, ed. pp. 102-136. Oxford, Blackwell.

McCawley, J.D. 1978. What is a tone language? *In* Tone: a linguistic survey. V.Fromkin, ed. pp. 113-131.

Mitchell, T.F. 1952. Prominence and syllabification in Arabic. Bulletin of the School of Oriental and African Studies, 23, 269.

Mitchell, T.F. 1960. Prominence and syllabification in Arabic. [Reprinted in Mitchell (1975:75-98)]. Bulletin of the School of Oriental and African Studies, 23, 269.

Mitchell, T.F. 1962. Colloquial Arabic: the living language of Egypt. London, English Universities Press.

Mitchell, T.F. 1969. Review of 'Elements of general Phonetics. Journal of Linguistics, 5, 153-164.

Mitchell, T.F. 1975. Principles of Firthian Linguistics. London, Longman.

Mitchell, T.F. 1990. Pronouncing Arabic 1. Oxford, Clarendon Press.

Mitchell, T.F. 1993. Pronouncing Arabic 2. Oxford, Clarendon Press.

Morgan, J. & Demuth, K. 1996. Signal to syntax: bootstrapping from speech to grammar in early acquisition. Lawrence Erlbaum.

Moutouakil, A. 1989. Pragmatic functions in a functional grammar of Arabic. Dordrecht, Foris.

Mughazy, M. 2003. Metalinguistic negation and truth functions: the case of Egyptian Arabic. Journal of Pragmatics: An Interdisciplinary Bi-monthly of Language Studies, 35, 1143-1160.

Myers, S. 1997. OCP effects in Optimality Theory. Natural Language & Linguistic Theory, 15, 847-892.

Nejmi, H. 1993. Contribution à une étude des processus accentuels de l'arabe marocain de Casablanca: approche phonétique et phonologique (accent de mot). Paris III: Universitéde la Sorbonne Nouvelle. Thesis/Dissertation

Nespor, M. & Vogel, I. 1986. Prosodic phonology. Dordrecht, Foris.

Nespor, M. & Vogel, I. 1989. On clashes and lapses. Linguistic Inquiry, 10, 467-482.

Nibert, H. 2000. Phonetic and phonological evidence for intermediate phrasing in Spanish intonation. University of Illinois at Urbana Champaign. Thesis/Dissertation

Norlin, K. 1989. A preliminary description of Cairo Arabic intonation of statements and questions. Speech Transmission Quarterly Progress and Status Report, 1, 47-49.

O'Connor, J.D. & Arnold, G.F. 1961. The intonation of colloquial English. London, Longman.

Ohala, J. & Jaeger, J. 1986. Introduction. *In* Experimental Phonology. J.Ohala and J.Jaeger, eds. pp. 1-12. Orlando., Academic Press.

Ouhalla, J. 1994. Focus in Standard Arabic. Linguistics in Potsdam, 1, 65-92.

Ouhalla, J. 1997. Remarks on focus in Standard Arabic. *In* Perspectives on Arabic Linguistics X. M.Eid and R.R.Ratcliffe, eds. pp. 9-45. Amsterdam, John Benjamins.

Ouhalla, J. 1999. Focus and arabic clefts. *In* The grammar of focus. G.Rebuschi and L.Tuller, eds. pp. 335-360. Amsterdam, John Benjamins.

Peters, J. 2005 The tone-bearing unit in Limburgian Dutch (paper presented at the Manchester Phonology Meeting).

Pierrehumbert, J. 1980. The phonology and phonetics of English intonation. MIT. Thesis/Dissertation

Pierrehumbert, J. & Beckman, M. 1988. Japanese tone structure. Cambridge, MA, MIT Press.

Poser, W.J. 1984. The phonetics and phonology of tone and intonation in Japanese. MIT. Thesis/Dissertation

Post, B. 2000. Tonal and phrasal structures in French intonation. Katholieke Universiteit Nijmegen. PhD. Diss. (LOT series, Netherlands Graduate School of Linguistics, 34. Thesus: The Hague). Thesis/Dissertation

Potts, C. 2002. The syntax and semantics of *as*-parentheticals. Natural Language & Linguistic Theory, 20, 623-689.

Potts, C. 2003. The logic of conventional implicatures. University of California at Santa Cruz. Thesis/Dissertation

Prieto, P. 2003 Coarticulation and stability effects in tonal clash contexts in Catalan.

Prieto, P. 2005a. Stability effects in tonal clash contexts in Catalan. Journal of Phonetics, In Press, Corrected Proof.

Prieto, P. 2005b. Syntactic and eurhythmic constraints on phrasing decisions in Catalan. Studia Linguistica.

Prieto, P., D'Imperio, M., & Gili Fivela, B. 2005. Pitch accent alignment in Romance: primary and secondary associations with metrical structure. *In* Intonation in language varieties. P.Warren, ed.

Prieto, P. & Torreira, F. 2004 Effects of syllable structure and speech rate on H prenuclear peak placement in Spanish. (ms.).

Prieto, P., van Santen, J., & Hirschberg, J. 1995. Tonal alignment patterns in Spanish. Journal of Phonetics, 23, 429-451.

Prince, A. 1983. Relating to the grid. Linguistic Inquiry, 14, 19-100.

Prince, A. & Smolensky, P. 1993 Constraint interaction in generative grammar. New Brunswick, NJ, Rutgers University Center for Cognitive Science.

Quirk, R., Duckworth, A.P., Rusiecki, J.P.L., & Colin, A.J.T. 1964. Studies in the correspondence of prosodic to grammatical featuresin English. *In* Proceedings of the Ninth International Congress of Linguistics. H.G.Lunt, ed. The Hague, Mouton.

Rammuny, R.M. 1989. Instrumental and auditory analysis of colloquial Jordanian Arabic: an update. Journal of Arabic Linguistics, 20, 23-42.

Ramus, F. 2002 Acoustic correlates of linguistic rhythm: perspectives. Aix-en-Provence, France. Proceedings of Speech Prosody 2002.

Rifaat, K. 1991. The intonation of Arabic: an experimental study. University of Alexandria. Thesis/Dissertation

Rifaat, K. 2004 The structure of Arabic intonation: a preliminary investigation.

Rifaat, K. 2005 The phonology of Colloquial Egyptian Arabic intonation. (Abstract submitted to PAPI 2005, Barcelona.).

Roach, P. 1982. On the distinction between 'stress-timed' and 'syllable-timed' languages. *In* Linguistic Controversies. D.Crystal, ed. pp. 73-79. London, Edward Arnold.

Romaine, S. 2002. Multilingualism. *In* The Handbook of Linguistics. Aronoff and Rees-Miller, eds. pp. 512-532. Oxford, Blackwell.

Rooth, M. 1996. Focus. *In* The Handbook of Contemporary Semantic Theory. S.Lappin, ed. pp. 271-297. Oxford, Blackwell.

Salib, M. 1981. Spoken Arabic of Cairo. Cairo, AUC Press.

Sandalo, F. & Truckenbrodt, H. 2002. Some notes on phonological phrasing in Brazilian Portuguese. MIT Working Papers in Linguistics, 42, 285-310.

Sauzet, P. 1989. L'accent du grèc ancien et les relations entre structure métrique et représentation autoegmentale. Langages, 95, 81-113.

Schafer, A. & Jun, S.-A. 2002. Effects of accentual phrasing on adjective interpretation in Korean. *In* East Asian Language Processing. M.Nakayama, ed. pp. 223-255. Stanford, CSLI.

Schwarzschild, R. 1999. Givenness, avoid F, and other constraints on the placement of accents. Natural Language Semantics, 7, 141-177.

Selkirk, E.O. 1981a. Epenthesis and degenerate syllables in Cairene Arabic. *In* Theoretical issues in the grammar of Semitic languages. H.Borer and J.Aoun, eds. pp. 209-232. Cambridge, Mass., Department of Linguistics, MIT.

Selkirk, E.O. 1981b. On prosodic structure and its relation to syntactic structure. *In* Nordic Prosody II. T.Fretheim, ed. pp. 111-140. Trondheim, TAPIR.

Selkirk, E.O. 1984. Phonology and syntax: the relation between sound and structure. Cambridge, Mass., MIT Press.

Selkirk, E.O. 1986. On derived domains in sentence phonology. Phonology Yearbook, 3, 371-405.

Selkirk, E.O. 1995. Sentence prosody: intonation, stress and phrasing. *In* The Handbook of Phonological Theory. J.Goldsmith, ed. pp. 550-569. Cambridge,Mass., Blackwell.

Selkirk, E.O. 1996. The prosodic structure of function words. *In* Signal to syntax: bootstrapping from speech to grammar in early acquisition. J.Morgan and K.Demuth, eds. Lawrence Erlbaum.

Selkirk, E.O. 2000. The interaction of constraints on prosodic phrasing. *In* Prosody: theory and experiment. M.Horne, ed. Dordrecht, Kluwer.

Selkirk, E.O. 2002 Contrastive *FOCUS* vs, presentational *focus*: prosodic evidence from right node rasining in English. Aix-en-Provence, France. Proceedings of Speech Prosody 2002.

Selkirk, E.O. 2004a. Bengali intonation revisited. *In* Topic and focus: a cross-linguistic perspective. Lee, Gordon, and D.Büring, eds. pp. 217-246. Kluwer Academic Publishers.

Selkirk, E.O. 2004b From tonal typology to focus prosody: the role for metrical prominence.

Selkirk, E.O. 2005a. Comments on intonational phrasing in English. *In* Prosodies. (Selected papers from the Phonetics and Phonology in Iberia Conference, 2003). Phonetics and Phonology Series. S.Frota, M.Vigario, and M.J.Freitas, eds. Mouton de Gruyter.

Selkirk, E.O. 2005b Seminar in Tone and intonation. Class handouts and notes. (UMass Amherst).

Selkirk, E.O. & Kratzer, A. 2004 Seminar in Meaning and intonation. Class handouts and notes. (UMass Amherst).

Selkirk, E.O., Shinya, T., & Kawahara, S. 2004. Phonological and phonetic effects of Minor Phrase length on F0 in Japanese. Speech Prosody 2004 (Nagahara, Japan).

Selkirk, E.O., Shinya, T., & Sugahara, M. 2003. Degree of initial lowering in Japanese as a reflex of prosodic structure organization. Proceedings of the 15th ICPhS.

Selkirk, E.O. & Tateishi, K. 1988 Constraints on minor phrase formation in Japanese. Proceedings of Annual Meeting of Chicago Linguistic Society 24. Chicago, CLS.

Selkirk, E.O. & Tateishi, K. 1991. Syntax and downstep in Japanese. *In* Interdisciplinary approaches to language: essays in honor of S-Y Yuroda. Dordrecht, Kluwer.

Shinya, T., Selkirk, E.O., & Kawahara, S. 2004 Rhythmic boost and recursive Minor Phrase in Japanese.

Shlonsky, U. 1997. Clause structure and word order in Hebrew and Arabic. New York, OUP.

Sibawayhi. 1990. Al-Kitaab. Beirut, Mu'assassat Al-'A'lami Li-Al-Matbuu'aat.

Siemund, R., Heuft, B., Choukri, K., Emam, O., Maragoudakis, E., Tropf, H., Gedge, O., Shammass, S., Moreno, A., Nogueiras Rodriguez, A., Zitouni, I., & Iskra, D. 2002 OrienTel - Arabic speech resources for the IT market. LREC 2002 (Arabic workshop).

Silverman, K.E.A. & Pierrehumbert, J. 1990. The timing of prenuclear high accents in English. *In* Papers in Laboratory Phonology. J.Kingston and M.Beckman, eds. pp. 72-106. Cambridge, CUP.

Sneed, E. 2004. The role of prosody in Hungarian information structure. Proceedings of the Annual Meeting of Chicago Linguistic Society, 40.

Steriade, D. 1988. Greek accent: a case for preserving structure. Linguistic Inquiry, 19, 271-314.

Straub, K.A. 1997. The Production of Prosodic Cues and their Role in the Comprehension of Syntactically Ambiguous Sentences. University of Rochester, N.Y. Thesis/Dissertation

Suleiman, Y. 1999. The Arabic Grammatical Tradition: A Study in Ta'liil. Edinburgh, Edinburgh University Press.

Swerts, M., Krahmer, E., & Avesani, C. 2002. Prosodic marking of information status in Dutch and Italian: a comparative analysis. Journal of Phonetics, 30, 629-654.

Truckenbrodt, H. 1999. On the relation between syntactic phrases and phonological phrases. Linguistic Inquiry, 30, 219-255.

Truckenbrodt, H. 1995. Phonological phrases: their relation to syntax, focus and prominence. MIT. Thesis/Dissertation

Vallduví, E. 1991. The role of plasticity in the association of focus and prominence. Proceedings of the Eastern States Conference on Linguistics (ESCOL), 7, 295-306.

Vanderslice, R. & Ladefoged, P. 1972. Binary suprasegmental features and transformational word-accentuation rules. Language, 48, 819-838.

Vigario, M. & Frota, S. 2003. The intonation of Standard and Northern European Portuguese: a comparative intonational phonology approach. Journal of Portuguese Linguistics, 2, 115-137.

Warren, P., Schafer, A., Speer, S.R., & White, S.D. 1999. Prosodic resolution of prepositional phrase ambiguity in ambiguous and unambiguous situations. UCLA Working Papers in Phonetics, 99, 5-33.

Watson, J.C.E. 1993. A syntax of San'ani Arabic. Wiesbaden, Harrossowitz Verlag.

Watson, J.C.E. 2002. The phonology and morphology of Arabic. Oxford, OUP.

Welden, A. 1980. Stress in Cairo Arabic. Studies in the Linguistic Sciences, 10, 99-120.

Wightman, C.W., Shattuck-Hufnagel, S., Ostendorf, M., & Price, P.J. 1992. Segmental durations in the vicinity of prosodic phrase boundaries. Journal of the Acoustical Society of America, 91, 1707-1717.

Xu, Y. 2002 Articulatory constraints and tonal alignment. Proceedings of Speech Prosody 2002. Aix en Provence, France.

Xu, Y. & Liu, F. 2005 What can tonal alignment tell us about segmental alignment?

Xu, Y. & Xu, C.X. 2005. Phonetic realization of focus in English declarative intonation. Journal of Phonetics, 33, 159-197.

Yeou, M. 2004 Effects of focus, position and syllable structure on F0 alignment patterns in Arabic. JEP-TALN 2004.

Yip, M. 2000. The complex interaction of tones and prominence. Proceedings of NELS, 31.

Yip, M. 2002. Tone. Cambridge, CUP.

Zec, D. 2002. On the prosodic status of function words. Working Papers of the Cornell Phonetics Laboratory, 14.

Zubizaretta, M.L. 1998. Prosody, focus and word order. Cambridge, Mass., MIT Press.