Variation in the intonation of Egyptian Formal and Colloquial Arabic

DINA EL ZARKA & SAM HELLMUTH

Introduction

In an earlier paper (Hellmuth & El Zarka 2007), we explored the hypothesis that the intonation of spoken formal Arabic ('fusˤħa') would display the same intonational properties as those of the mother tongue dialect of the speaker ('ʕaammiyya'), on the basis of both quantitative and qualitative analysis of a parallel corpus of read speech data in two varieties of Arabic: Egyptian Formal Arabic (EFA, 'fusˤħa'), defined as the rendition of formal Arabic by Egyptians, and Egyptian Colloquial Arabic (ECA, 'ʕaammiyya').

Overall, the results of Hellmuth & El Zarka (2007) supported the hypothesis that EFA and ECA intonation would be similar. The quantitative study suggested no major differences between EFA and ECA in the alignment of F0 peaks and valleys with respect to the segmental string. There were however indications of qualitative differences between the two varieties which we felt required further investigation.

The present study expands on the qualitative approach (only) of Hellmuth & El Zarka (2007), in two ways: firstly, by increasing the number of speakers whose recordings are subjected to analysis (from two speakers to five); and, secondly, by investigating additional qualitative variables highlighted by the previous study as potentially fruitful areas for comparison between the two varieties.

Notably, the dataset examined in Hellmuth & El Zarka (2007) showed frequent occurrences of a previously unexplored feature of Egyptian Arabic intonation contours: the appearance of a low plateau between two successive accent peaks, instead of just a single valley point. One aim of the present study therefore was to document this pattern more fully, in order to understand it better, and to establish with greater certainty whether or not it occurs equally in both EFA and ECA. Two other qualitative variables are also explored in greater depth in the present study: the incidence of phrase breaks (which in Hellmuth & El Zarka 2007 was found to be higher in EFA than in ECA) and the incidence of secondary accents (argued in El Zarka & Hellmuth 2007 to be a feature of EFA that sets it apart from ECA).

In the next section we set out the background to the present study, including a detailed summary of the findings, both quantitative and qualitative, of our earlier study, as well as the methodology and rationale of the qualitative analysis conducted for the present study. The following section describes and illustrates the results of the analysis, highlighting the similarities and differences between EFA and ECA that were observed. A brief summary of our findings concludes the paper.

Background to the study

The hypothesis that spoken formal Arabic ('fusˤħa') will display similar intonational properties to those of the mother tongue dialect of the speaker ('ʕaammiyya') is relatively uncontroversial, given the commonly held view that, in general, the phonological (and phonetic) properties of spoken formal Arabic are transferred from the speaker's colloquial dialect. It has been claimed for example that in spoken formal Arabic speakers will apply the particular stress assignment rules of their own mother tongue dialect (Benkirane 1998). This is testable, since although most spoken Arabic dialects display quantity sensitive stress assignment, the exact rules for assigning primary stress vary from dialect to dialect (Hayes 1995), and studies of stress assignment in formal and colloquial varieties of Egyptian Arabic studies (e.g. McCarthy 1979) suggest that the claim holds true. Prior to Hellmuth & El Zarka (2007) and El Zarka & Hellmuth (2007), no studies had sought to test this hypothesis in respect to intonational properties.

The distinction between standard and spoken varieties of Arabic is notoriously complex (Badawi 1973, Ferguson 1957) and is not readily reduced to a simple two way classification. Mitchell (1984) suggests a three-way classification, with a primary distinction between formal and informal Arabic, and sub-division of informal Arabic into 'careful' and 'casual' registers. In our data collection we aimed always to elicit

the two extremes of this continuum (EFA formal vs. ECA informal casual), though we are aware that the actual data collected may not all fall neatly into these exact categories. Indeed, the findings of our 2007 study suggest that a distinction between careful vs. casual renditions can be observed in both EFA and ECA.

The study reported in Hellmuth & El Zarka (2007) used both qualitative and quantitative analysis to test for potential similarities and differences between the intonational properties of EFA and ECA. A single set of stimuli were used as input to both levels of analysis.

The quantitative analysis tested for differences between EFA and ECA in the fine-grained phonetic realisation of pitch accents, using alignment of pitch targets to the segmental string as dependent variable. In order to test this a set of stimulus sentences were created, in each variety, containing pairs of target words which were segmentally parallel, and where possible mostly sonorant, in and around the stressed syllable. Target stressed syllables elicited were of three types: short open (CV), short heavy (CVC) and long heavy (CVV). All target words contained the vowel [a] as the stressed vowel. Sample target words are illustrated in Table 1 below. Twelve different target words were created per syllable type, in each of two varieties (EFA/ECA), yielding 12 x 3 x 2=72 targets in total.

Table 1: Sample target words (register/syllable type).

	CV	CVV	CVC
ECA	ˈmalik ‘king’	ˈmaːlik ‘owner’	ˈmalħa ‘salty’
EFA	ˈmalik ‘king’	ˈmaːlik ‘owner’	taˈmalmul ‘nervousness’

Table 2: Sample frame sentences.

ECA	
ˈmalik	ˈʃufna ˈmalik is-swiːdiyˈyiːn ˈlamma ˈruħna s-ˈswiːd
	saw-we king [of] DEF-Swedes when went-we Sweden “We saw the king of the Swedes when we went to Sweden.”
ˈmaːlik	ˈʃufna ˈtˤeːr ˈismu ˈmaːlik il-ħaˈziːn mawˈguːd fi ˈwaːdi n-ˈniːl
	saw-we bird name-his malik il-hazeen existing in valley [of] DEF-Nile “We saw a bird named Malik il-Hazeen that lives in the Nile valley.”
ˈmalħa	iʃ-ˈʃurba l-ˈmalħa miʃ kwayˈyisa li-sˤ-ˈsˤaħħa
	DEF-soup DEF-salty not good for-DEF-health “Salty soup is not good for your health.”
EFA	
ˈmalik	laqad raˈʔayna ˈmalikata s-swiːdiyˈyiːna ˈʕindama ˈzurna s-ˈswiːd
	PART saw-we queen [of] DEF-Swedes when visited-we Sweden “We saw the queen of the Swedes when we visited Sweden.”
ˈmaːlik	ʔinˈnana qad ˈzurna ˈmaːlikata l-maˈtaːgiri ˈlaylata l-xaˈmiːs
	PART-we PART visited-we owner.FEM [of] DEF-shops night [of] DET-Thursday “We visited the (female) owner of the shops on Thursday night.”
taˈmalmul	ˈʔinna taˈmalmula l-mudarriˈsiːna l-ˈgudud muʃˈkilatun kaˈbiːra
	PART anxiety [of] DEF-teachers DEF-new problem big “The anxiety of the new teachers is a big problem.”

Each target word was placed in its own unique frame sentence, as illustrated in Table 2 (the full set of stimulus sentences is provided in the Appendix). Each sentence was designed to be as natural as possible to facilitate elicitation of the correct register of Arabic. In many cases this resulted in a relatively longer sentence in EFA sentences than in ECA sentences, and the implications of this slight difference are discussed in the results section below. Target words were sometimes placed in more than one carrier sentence, to ensure at least one useable rendition. In addition, within each sentence, the prosodic context surrounding the target word was controlled to avoid tonal clash from a following adjacent accented

syllable. As we shall see below, the analysis in the present study shows that our efforts to avoid tonal clash created relatively long sequences of unaccented syllables, which had tonal consequences of their own.

In total the stimuli materials contained 43 ECA sentences and 42 EFA sentences. These were pseudo-randomised and interspersed with distractor sentences in the relevant register. They were presented to speakers on a computer screen one at a time typed in Arabic script. For ECA we used Egyptian lexical items and spelling, such as that used in cartoons, to elicit the correct register. For the study described in Hellmuth & El Zarka (2007) the stimuli materials were recorded by two male speakers of EA ('AM' and 'BM'), both of whom were born and raised in Cairo and had University level training in formal Arabic. Each speaker read the materials three times, first in ECA and then in EFA. The corpus was analysed with reference to F0 contour and wideband spectrogram using Praat 4.5.2 (Boersma & Weenink 2006).

The full corpus of 484 sentences was subjected to qualitative analysis, a key function of which was to identify tokens which were suitable for inclusion in the quantitative analysis. Since the prosodic effects of clash, boundary and focus on the alignment of pitch events relative to the segmental string are well documented (Prieto & Shih 1995, Prieto et al. 1995, Yeou 2004) tokens of the following kinds were tagged for exclusion from the quantitative study: tokens containing a disfluency or assignment of word stress in an unexpected position in the target syllable or in which an additional accented syllable results in less than three intervening syllables between the target syllable and the following accented syllable ('x'), tokens in which the speaker inserts a phrase boundary after the target word ('y') and tokens in which the speaker places focus on the target word itself ('z'). A count was made of the number of tokens containing a phrase boundary after the target word (the 'y' tokens), for comparison of the incidence of phrase breaks in EFA and ECA.

For the quantitative analysis, following standard practice in the literature (Prieto et al. 1995, Arvaniti et al. 1998, Ladd et al. 2000), the segmental and pitch events listed in Table 3 below were labelled by hand in each token. The following dependent variables were calculated (following Prieto et al. 1995): i) L-C0 ('L1 to onset of stressed syllable'); ii) peak delay (H-C0, 'H to onset of stressed syllable'); and iii) relative peak delay 'rpd', which is a proportional measure of peak delay relative to the duration of the stressed syllable.

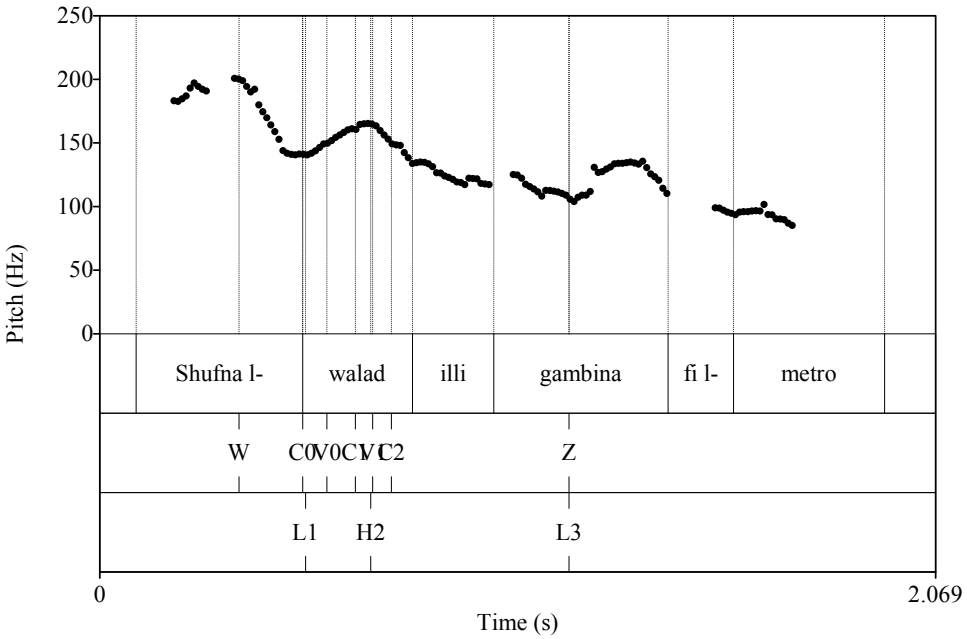
Table 3: Segmental landmarks/ pitch events labelled in each target syllable.

label	position
C0	start of initial consonant of target syllable
V0	start of vowel of target syllable
C1	start of next consonant
C2	start of coda consonant (if present)
V1	start of following vowel
W	end of the preceding accented syllable
X	left edge of the target word
Y	right edge of the target word
Z	start of the following accented syllable
H	F0 maximum turning point in target syllable
L1	F0 minimum turning point before H
L2	F0 minimum turning point after H

Qualitative results of Hellmuth & El Zarka (2007)

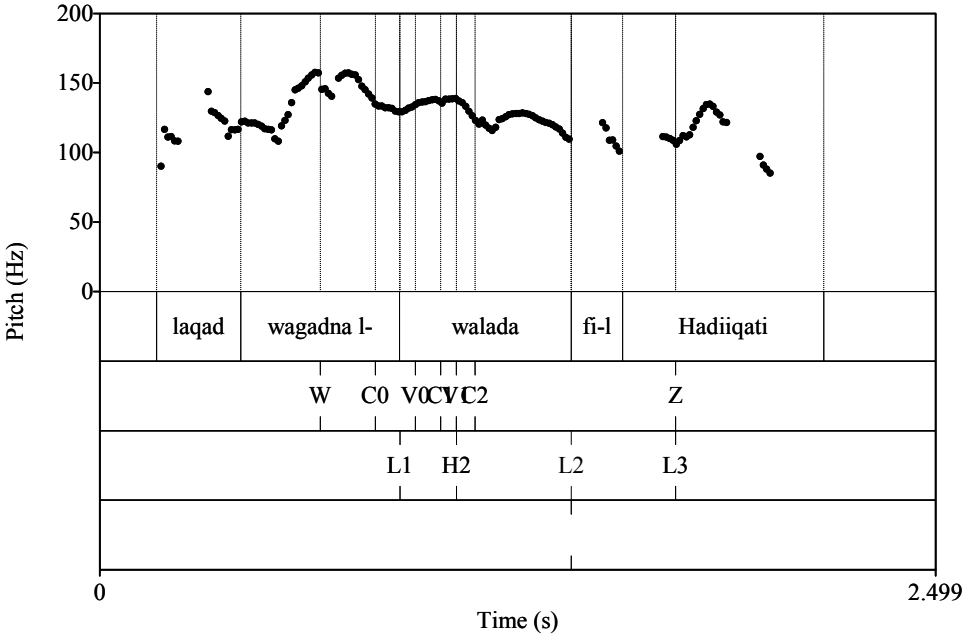
Qualitative analysis of the corpus confirmed that EFA and ECA share a key intonational property, which is that both varieties displayed highly populated pitch accent distribution, with an accent realised on almost every content word (cf. for ECA (Hellmuth 2006) and for EFA (Rastegar-El Zarka 1997, Rifaat 1991). Also, in the majority of cases, all accents in non-phrasefinal ('pre-nuclear') position were of roughly the same shape, namely an accent whose most salient property is an H peak realised towards the end of the stressed syllable. These properties can be observed in the examples in Figures 2 and 3.

Figure 2: ECA token (speaker AM).
The 1st annotation tier (at the top) shows word boundaries and gloss; the 2nd tier shows segmental landmark labels; the 3rd tier (at the bottom) shows tonal event labels.



ʃufna-l walad illi gambina fi-l metru
we-saw-the boy that next-us in-the metro
"We saw the boy next to us in the metro."

Figure 3: EFA token (speaker AM).
The additional annotation tier (at the bottom) shows phrase boundaries (if any).



laqad wagadna-l walada fi-l ħadiiqati
PART we-found-the boy in-the garden
"We found the boy in the garden."

In autosegmental-metrical terms (Ladd 1996) an accent whose most salient property is an H peak can be analysed as an H tone associated with the stressed syllable (thus, H*). In fact all prior phonological analyses of EA pre-nuclear pitch accents agreed on this aspect of the basic EA pitch accent, that is all assigned 'starred tone' status to the H tone. The accent has been analysed variously for EFA as H*+L (Rastegar-El Zarka 1997) and as a H* with flanking fall and rise (Rifaat 2003), and in ECA as L+H* (Hellmuth 2006). We suggested in our earlier study that this apparent variation may thus reflect differences of theoretical analysis of what is in fact the same phonological object in EFA and ECA: an H peak with a dependent L tone. Examination of the positioning of L tone valleys either side of the H peak suggested that the prior L tone was consistently present, and the quantitative study showed that it was also consistently aligned, at the onset of the stressed syllable. We did however note the frequent occurrence in the corpus of a low plateau between successive H peaks (suggesting two L tones between peaks, rather than a single L tone). Since the incidence of L plateaus seemed to be roughly the same in EFA and ECA, further investigation was set aside for future study, hence the present paper which permits a more thorough investigation of the incidence of 'L2' tones, in the two varieties.

Overall then, the qualitative analysis suggested that EFA and ECA showed that they share key intonational properties, in the type and distribution of pitch accents. However, one clear difference between the two registers was found: a greater proportion of sentences containing an internal phrase boundary were observed in EFA than in ECA. Recall that this observation took into account only phrase boundaries occurring after a target word; that is, the count was made among those tokens excluded from the quantitative study (the 'y' cases). Nevertheless, even only counting boundaries at this particular position in the sentence, there was a clear difference between the two registers for both speakers, and most clearly so for speaker BM. The number of 'y' tokens, as a proportion of total fluent tokens was as follows: speaker AM, ECA 24.5%, EFA 54.2%; speaker BM, ECA 9.0%, EFA 37.1%. A sample EFA token containing a boundary is provided in Figure 3 above. A more thorough investigation of the incidence of phrase boundaries was therefore seen as important for inclusion in the present study.

Quantitative results of Hellmuth & El Zarka (2007)

The purpose of the quantitative analysis was to determine whether there are fine-grained differences between EFA and ECA, in the phonetic realisation of what was shown in the qualitative analysis to be a single phonological category. This is because, although research on pitch target alignment indicates broad cross-linguistic similarities, such as a tendency towards more consistent alignment to the segmental string at the beginning of the syllable than at the end (Prieto & Torreira 2006), extensive cross-linguistic variation in the fine detail of alignment patterns has been found (Arvaniti et al. 1998, Ladd et al. 2000).

Following Lickley et al. (2005), when reporting distances from a pitch event which is relatively close to a segmental landmark an absolute measure is used (e.g. 'L-C0'), in order to minimise variance; in contrast, a proportional variable 'relative peak delay' (rpd) is used to express the distance of the H peak from the syllable onset (H-C0) as a proportion of the stressed syllable duration. A total of 227 tokens were included in the quantitative analysis (for speaker AM, 91 tokens: ECA 58 + EFA 33; for speaker BM, 136 tokens: ECA 71 + EFA 65).

As already noted, the position of L1 (the low turning point before the H peak) was observed during qualitative analysis to fall routinely at the onset of the stressed syllable. Comparison of mean values (across all included tokens, for both speakers and registers, N=227) of L1-C0 (distance of L1 from the start of the stressed syllable onset) vs. L1-X (distance of L1 from the start of the target word) confirms that L1 is more closely aligned with the start of the stressed syllable than with the start of the word: mean L1-C0 = 5.59ms.; mean L1-X = 30.51 ms. C0 is thus the closest segmental landmark to L1. Figure 4 displays median and interquartile values of L1-C0 by speaker and by register. L1 is positioned just after the onset of the accented syllable in both registers, for both speakers. A oneway ANOVA by register, with L1-C0 as the dependent variable (performed for each speaker independently), indicated that the position of L1 does not vary significantly between registers for speaker AM ($F(1,89) = 0.300$; $p = .585$; $\alpha = 0.05$); for speaker BM the variation in the distribution of L1-C0 values was too great to permit a statistical test to be performed. We thus took the position of L1 to be parallel across registers for both speakers.

Figure 4: Median and interquartile values of the distance from L1 to the syllable onset (L1-C0) in ms. The vertical line indicates the *beginning* of the accented syllable; EFA plain box, ECA striped box.

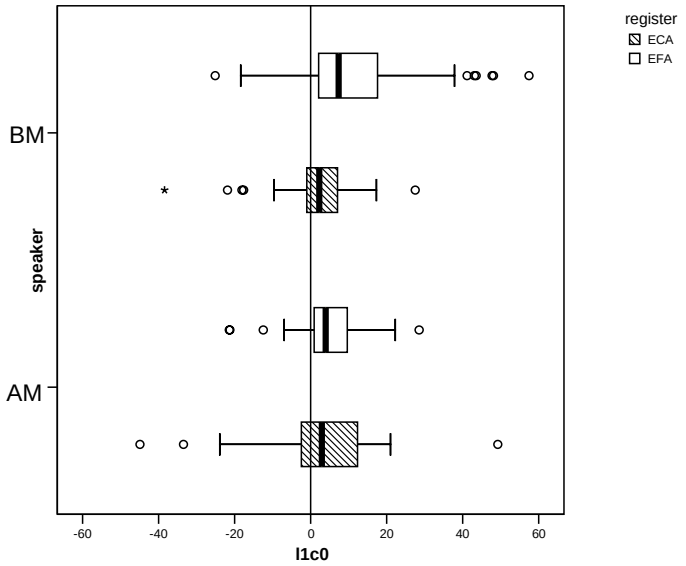
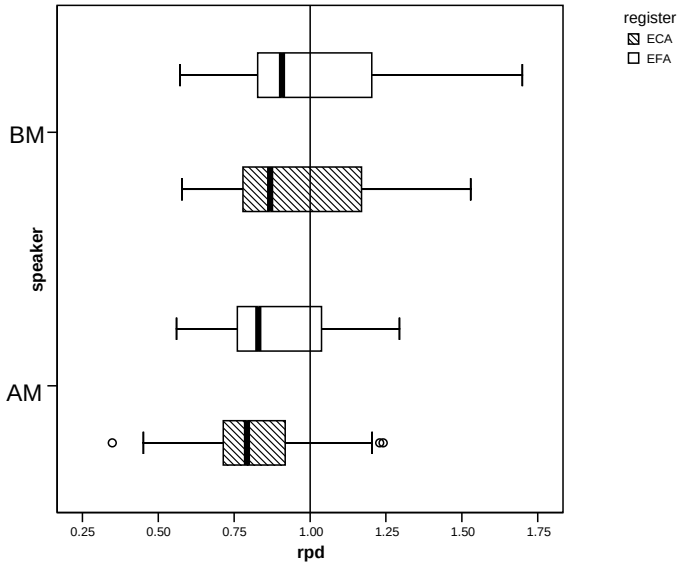


Figure 5: Median and interquartile values of relative peak delay (H-C0/syllable duration) in ms. The vertical line indicates the *end* of the accented syllable; EFA plain box, ECA striped box.



To investigate the position of the H peak we used a proportional measure of the position of H relative to the duration of the whole accented syllable (relative peak delay). Figure 5 displays the median and interquartile ranges of relative peak delay (rpd) by speaker and by register, across all syllable types. Note that values of rpd < 1 indicate alignment of H within the accented syllable. For both speakers, H is aligned within the syllable, with H aligned slightly earlier overall for speaker AM. A one-way ANOVA by register, with rpd as the dependent variable (performed for each speaker independently), indicates that the position of H varies significantly between registers for speaker AM only: speaker AM ($F(1,89) = 4.525$; $p = 0.036$; $\alpha = 0.05$); speaker BM ($F(1,134) = 1.676$; $p = 0.198$; $\alpha = 0.05$). Thus, H was shown to be aligned differently in ECA and EFA for speaker AM only, whilst for speaker BM H is aligned in parallel fashion in ECA and EFA.

Closer examination of H alignment in AM's data showed that the significant result was due to early alignment of H in CV target syllables, and, contrary to expectations, in ECA only. The pattern observed in speaker AM's CV syllables was similar to that observed in the corpus of broadcast EFA in (Rastegar-

El Zarka 1997). Hellmuth & El Zarka (2007) thus suggested that this marginal early H alignment may be a hallmark of 'professional speech' in both ECA and EFA, rather than a hallmark of EFA itself; this was because speaker AM's productions in both registers were in general more deliberate (or planned) than those of speaker BM. However, for this speaker (AM), due to the large number of tokens produced with a phrase boundary after the target word, the number of EFA CV tokens which could be included in the quantitative analysis was relatively small: just 13 tokens (cf. 22 in ECA). We suspect that in a larger dataset the apparent difference between EFA and ECA in the alignment of H in CV syllables, even for this speaker, might disappear.

Methodology of the present study

Overall then, based on the variables considered, the conclusions of Hellmuth & El Zarka (2007) were that the same basic intonational patterns are observed in EFA and ECA. The only clear difference found was in greater use of phrasing boundaries in EFA.

The present study uses qualitative analysis to investigate three intonational features which may potentially differ between EFA and ECA. The first variable is, again, the frequency of utterance-internal phrase boundaries, which is here investigated based on a count of all phrase breaks (and not only of those occurring immediately after what was the target word for the quantitative study). The second variable is the incidence of secondary accents, whereby a rhythmically stressed syllable (not bearing primary stress) is optionally marked also with an H* pitch accent; this variable is included because El Zarka & Hellmuth (2007) suggest that secondary accents are more prevalent in EFA than in ECA. The third variable included in the present study is the incidence of 'L2' turning points, which was noted in Hellmuth & El Zarka (2007) to occur in tokens displaying a low plateau between successive H peaks, rather than a 'single' valley low turning point; there appeared to be as many plateau cases in EFA as in ECA, however this was not formally investigated and thus the variable is included here.

The dataset examined consists of the dataset from the earlier study together with additional data collected for the present study with three additional speakers of EA, using the same stimulus materials. The full stimulus set described above was read three times in EFA and then, after a break, three times in ECA. There were two speakers for the earlier study (AM and BM) and three additional speakers for the present study (CM, DF and EF). Three speakers were male and two female (note that the second part of each subject identifier denotes gender (M = male, F = female); four were aged under 30 and one was in his forties (speaker CM). All speakers were born and raised in Cairo, except EF who was born and raised in Alexandria. All are mother tongue speakers of ECA and have university level formal training in the grammar of standard Arabic. Digital recordings were made at 44.1KHz 16bit directly to digital format in a quiet room (resampled to 22.05KHz 16bit prior to analysis).

Since the original stimulus materials consisted of 85 sentences (42 EFA sentences and 43 ECA sentences), five speakers reading the material three times each yields a corpus for the present study which consists of 630 EFA sentences (42 x 3 x 5 = 630) and 645 ECA sentences (43 x 3 x 5 = 645), that is, 1275 sentences in total. The decision was made to exclude as few sound files as possible from analysis; of the total 1275 sentences, 99 were discarded (due to technical recording problems, or a long pause after the target word or focus on the target word) and 1176 were submitted to analysis (610 EFA sentences and 566 ECA sentences).

The expanded corpus of data were subjected to qualitative analysis by the first author, with reference to the F0 pitch track and spectrogram extracted using Praat 5.1, and a count was made of the incidence of each variable. The count of phrase breaks and secondary accents includes a phrase break or secondary accent occurring anywhere in each whole frame sentence. The count of L2 occurrences includes only those occurrences of an L2 target appearing after a target word (as designated in the original study). Since this is the first study of the incidence of 'L2' in EA intonation a study of the potential factors governing its (non-)occurrence was also carried out, in a subset of the data (from the male speakers only). Specifically, a record was kept of the behaviour of L2 in relation to a number of factors, such as phrase breaks, number of unstressed syllables and nearby tonal configurations. The overall results of the qualitative analysis are summarised in Table 4 below.

Comparing EFA and ECA intonation

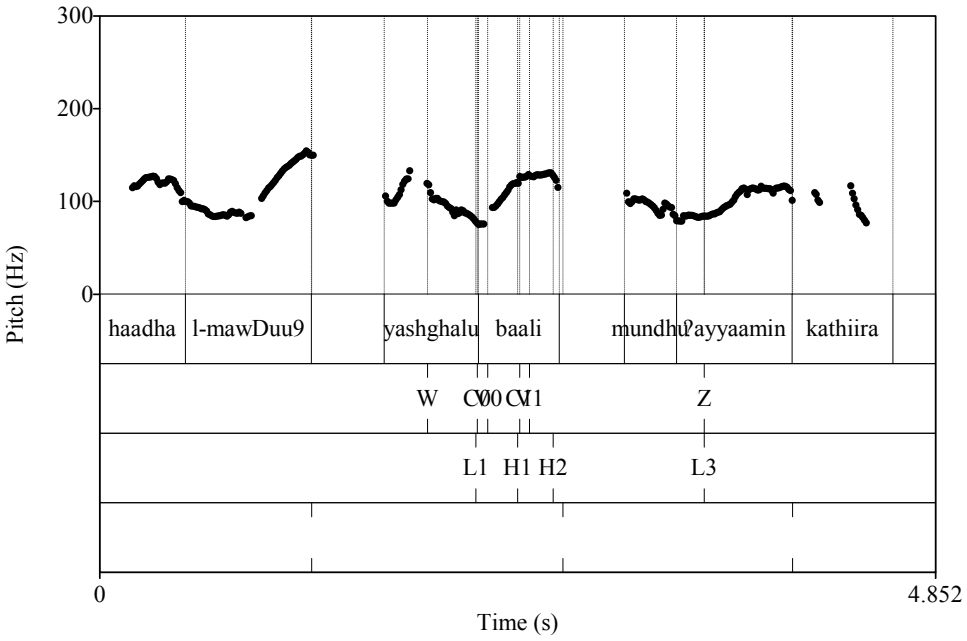
The incidence of phrase breaks and secondary accents in EFA and ECA

The data show a difference between EFA and ECA in the incidence of utterance-internal phrase breaks or boundaries. As shown in Table 4 below, the overall number of phrase breaks in ECA is 326 (occurring in 566 sentences), whereas in EFA it is almost twice that number (721 phrase breaks in 610 tokens). Figure 6 below shows an example of a token containing phrase boundaries (compare also Figure 3 above).

Table 4: Qualitative analysis summary.
Count of phrase breaks (PB), secondary accents (SA) and L2 occurrences in the corpus of all five speakers.

Colloquial					**Formal**				
speaker	**PB**	**SA**	**L2**	**TOTAL**	**speaker**	**PB**	**SA**	**L2**	**TOTAL**
AM	88	10	57	87	**AM**	197	46	73	118
BM	49	11	52	105	**BM**	184	38	65	129
CM	89	6	76	118	**CM**	157	26	49	112
DF	67	2	73	130	**DF**	90	4	75	124
EF	33	12	63	126	**EF**	93	11	85	127
	326	**41**	**321**	**566**		**721**	**125**	**347**	**610**

Figure 6: EFA token, containing phrase boundaries (speaker CM).
The 1st annotation tier (at the top) shows word boundaries and gloss; the 2nd tier shows segmental landmark labels; the 3rd tier shows tonal event labels; the 4th tier (at the bottom) shows phrase boundaries (if any).



haaða-l	mawdˤuuˤ	yaʃɣalu	baal-i	munðu	ʔayyaamin	kaθiira
this-DEF	subject	it-troubles	mind-my	since	days	many

"This subject has worried me for many days."

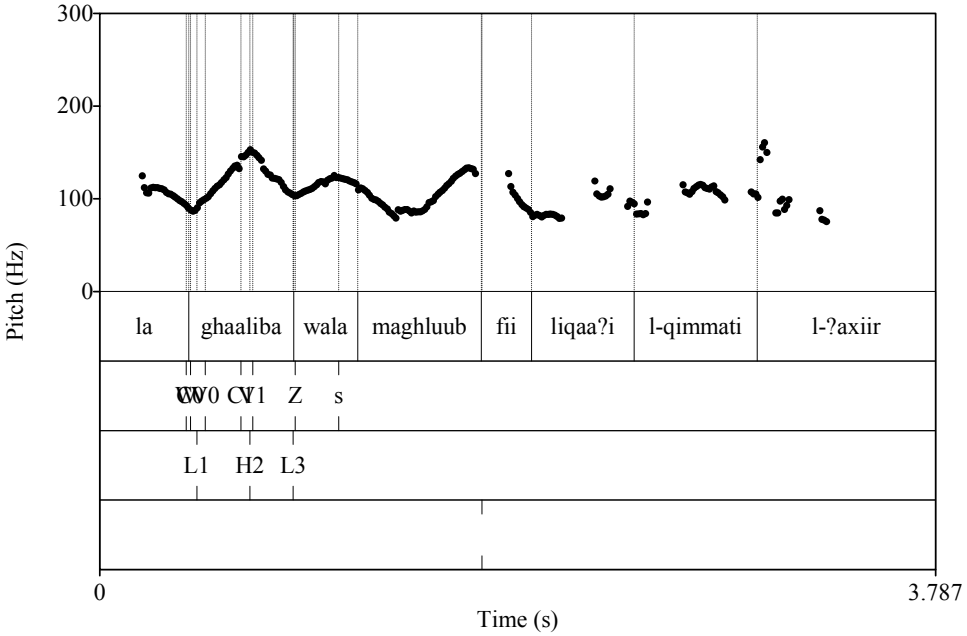
The tendency to insert more phrase breaks in EFA than in ECA holds for all speakers, but is particularly noticeable for the male speakers. The male speakers' renderings were often at a slower speech rate than those of the female speakers, and their EFA articulation could also be termed more formal and more typical of 'fusˤħa' then that of the female speakers. As already noted above, the EFA sentences were in general slightly longer than the ECA sentences, but even when this is not the case, the tendency to break up larger units into smaller ones seems to hold.

Turning to the incidence of secondary accents, this was found to be a feature more typical of EFA than of ECA. In EFA 125 secondary accents were observed, across 610 tokens; an incidence of 20.5%. In contrast, in ECA, only 41 secondary accents were observed, across 566 tokens; an incidence of only 7.2%. There is again a difference between the male and female speakers, in that secondary accents were much more likely to be observed in the speech of the male speakers than the female speakers. For the purposes of the count, a secondary accent was defined as an accent occurring on a function word or grammatical morpheme, or occurring in addition to the primary accent, on some other syllable, in a content word. Figures 7 and 8 below show tokens containing a secondary accent on a function word and on a non-primary stressed syllable, respectively.

It seems that "real" secondary accents like for example in 'Tàalibatáyni' "two female students" or 'mànSibáyn' occur almost exclusively in the formal register. Their function may be to break up longer stretches of unstressed syllables and to increase the number of pitch accents, rendering the utterances more formal and less natural. This has been claimed to be a deliberate choice to make 'fusʿħa'-utterances sound more important (Rastegar-El Zarka 1997).

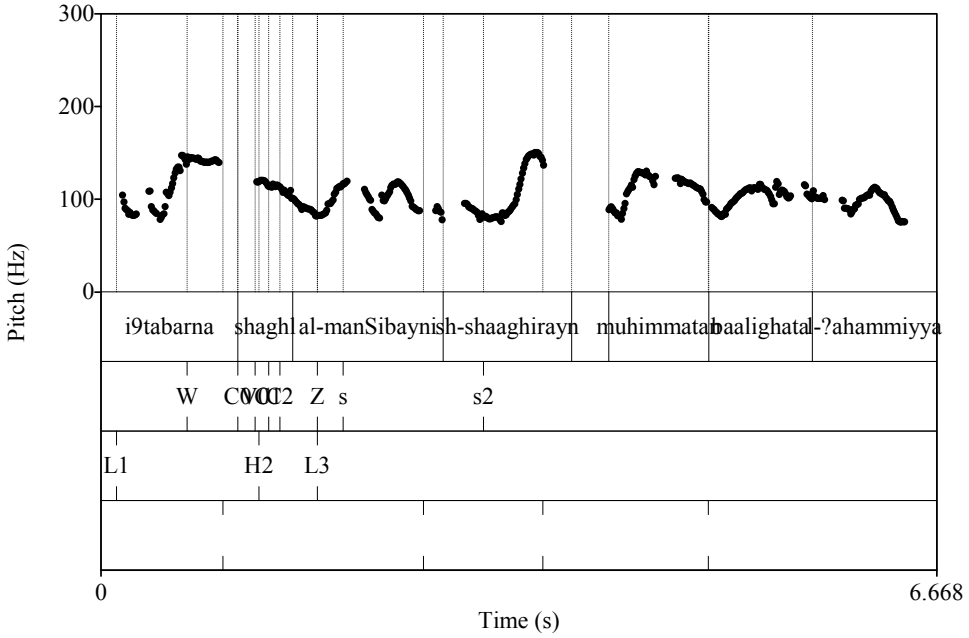
The incidence of both phrase breaks and secondary accents is of course intimately related to speech rate, which in turn can itself be viewed as a measure of formality and informality. The findings here are therefore consistent with the claim that greater use of phrasing boundaries and secondary accents is part of a wider strategy used by speakers to enhance rhythmicity in EFA (El Zarka & Hellmuth 2007).

Figure 7: EFA token, showing a secondary accent on the function word 'wala' "nor" (speaker CM).



laa γaaliba wa-la maγluub fii liqaaʔi-l qimmati-l ʔaxiir
NEG winner and-NEG defeated in meeting-DEF summit-DEF last
"There was no winner and no loser in the last summit."

Figure 8: EFA token, showing a secondary accent on a non-primary stressed syllable (speaker CM).
The word [ˌmansˁiˈbayni] bears an accent on both the 1st and 3rd syllables.



iʕtabarna	shaγla-l	manSibayni-ʃ	ʃaaγirayn	muhimmatan	baaliγata-l	ʔahammiyya
considered -we	occupation-DEF [of]	position-DU-DEF	vacant-DU	task	utmost-DEF	importance

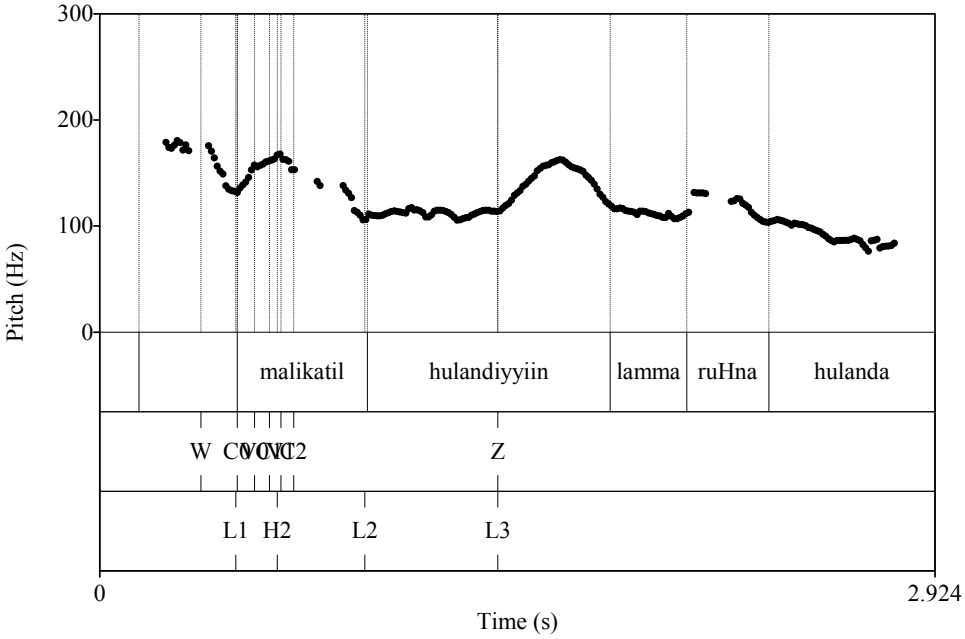
"We considered the occupation of the two vacant positions an issue of utmost importance."

The incidence of post-accentual valleys ('L2') in EFA and ECA

As in the subset of data analysed in Hellmuth & El Zarka (2007), close observation of the positioning of the L2 turning point in the full dataset reveals a large number of cases in which there is a low plateau between successive H peaks, rather than a single valley low turning point. This plateau can be interpreted phonologically as a second L pitch target or tone, following the H pitch target, which then forms a plateau running up to the L leading tone of the subsequent accent.

The additional L tone seems to be aligned either at the beginning of the following lexical item or at the end of the target word itself. When two content words are strictly adjacent these two positions of course coincide. However in cases where there is additional grammatical material occurring between two successive lexical words the two positions do not coincide. In the stimulus materials there were many target words which appeared as the first item in a 'construct state' annexation construction ('idˁaafa'), which means that they were followed by a second lexical item marked as definite (using the definite marker 'il'). See, for example, the target word 'malika' ("queen") in 'malikit il-hulandiyyiin' ("queen of the Dutch") in Figure 9 below. In such cases, the most common pitch contour observed was a steady fall all the way through to the beginning of the following word 'hulandiyyiin', stretching over the determiner, so that the L2 target was located at the beginning of the second lexical item. This could also be interpreted as alignment with the end of some phonological unit such as the 'prosodic word'. Similar behaviour of the L2 target has been observed in cases involving function words besides the determiner 'il', such as prepositions or the relative pronoun 'illi' (El Zarka 2008).

Figure 9: ECA token showing a post-accentual valley 'L2' (speaker AM).



ʃufna	malikit-il-hulandiyyiin	lamma	ruħna	hulanda
we-saw	queen-[of]-the-Dutch	when	we-went	[to] Holland

"We saw the queen of the Dutch when we went to Holland."

Looking again at the overall qualitative results in Table 4, it can be seen that the 'plateau' strategy is used by all speakers, and is observed in both EFA and ECA. Indeed, the proportion of L2 occurrences in both varieties is surprisingly similar, at a little over 50% (56.7% for ECA and 56.9% for EFA). Although there are some differences among speakers, the general picture is fairly uniform.

A subset of the data (that of the male speakers only) was examined in greater detail in order to determine what factors might influence the (non-)occurrence of L2. Specifically, the incidence of potentially conditioning factors was counted in all of the tokens in which no L2 was observed.

The three potentially conditioning factors examined were:

i) cases in which there are three or fewer unstressed syllables intervening between the target word and the following lexical item;
ii) cases with extensive downstep of the following accent (that is, of the accent on the lexical item following the target word, usually observed when the following word is final in the phrase); and,
iii) cases with suspended high pitch or a rise all through the target word (with or without a phrase break following the target word).

Factor ii) was included because it is generally not possible to identify any valley between successive accents when the second of the two accents is final in the utterance. Note that more than one factor may hold of any particular token. Investigation of cases without an L2 (rather than of cases with an L2) was carried out because we suspected that the primary factor conditioning the occurrence of L2 would be the number of interaccentual unstressed syllables, and we wanted to exclude the possibility that there was some other explanation for the non-occurrence of L2 than a lack of accentless syllables.

The number of tokens without an L2 in each of these conditions is shown in Tables 5 & 6. The residual number of cases, in which there was no obvious conditioning factor, is also shown.

Table 5: Incidence of relevant factors in cases without an L2, among male speakers, by register.
Conditioning factors: plateau or rising contour on the target word (P/R), downstep on the following word (D), three or fewer intervening syllables (≤3), or no obvious conditioning factor (none).

Colloquial						**Formal**					
speaker	**P/R**	**D**	**≤ 3**	**none**	**TOTAL**	**speaker**	**P/R**	**D**	**≤ 3**	**none**	**TOTAL**
AM	7	8	11	5	30	**AM**	2	0	**36**	7	45
BM	1	18	**28**	14	53	**BM**	9	0	**37**	25	63
CM	1	**23**	11	6	40	**CM**	**29**	6	19	18	63
TOTAL	9	49	50	25	123	**TOTAL**	40	6	92	50	171
%	7.3%	39.8%	40.7%	20.3%		%	23.4%	3.5%	53.8%	29.2%	

Table 6: Summary of relevant factors in cases without an L2, among male speakers, across both registers.

	Plateau/Rise	**Downstep**	**≤ 3 syllables**	**no obvious factor**	**no L2**
TOTAL	**49**	**55**	**142**	**75**	**294**
	16.6%	18.7%	48.3%	25.5%	

The summary in Table 6 shows that the most important factor conditioning the non-occurrence of L2 is that of having a small number of intervening unstressed syllables between accents. In almost half of the cases in which there was no L2 there were 3 or fewer intervening unstressed syllables between the target word and the following word. Other factors, such as downstep of the following accent and occurrence of a different tonal configuration on the target word itself, appear to play less of a role. In addition, the effect of the conditioning factors is not uniform across speakers, and also not even across registers within the speech of the same speaker. Speaker CM, for example, in ECA frequently uses renditions in a single phrase with a following downstepped accent, while he produces many EFA targets with a flat or rising intonation (on the target word), mostly in combination with a manifest phrase break afterwards.
These findings suggest that under certain conditions, most notably when there is a high number of intervening unstressed syllables between words (up to 6 in some of our stimulus sentences), a second low target is very likely to occur. As already mentioned, in the design of the stimulus materials we aimed to avoid a small number of intervening syllables between the accented syllable of a target word and the accented syllable of a following word. Hence, for example, the choice of the phrase ‘malikit il-hulandiyyiin’ (“queen of the Dutch”), which has five intervening unstressed syllables, rather than, say, ‘malikit holanda’ (“queen of Holland”) which has three intervening unstressed syllables. This means that our materials probably contained an unusually large number of word-word sequences in which there are more than three unstressed syllables between words, and thus created a large number of contexts in which the L2 post-accentual valley would appear.

Nevertheless, there is some freedom of realisation, as even in some of the examples with six intervening unstressed syllables a gradual fall to the low target of the following stressed syllable could be observed; and, as the final column in Table 6 shows, there are also a non-trivial number of cases (25%) where no obvious reason for the non-occurrence of L2 could be identified.

As regards any differences between EFA and ECA in the incidence of L2, the percentage of cases without a second low target (ECA 43.3 % vs. EFA 43.1%) and the number of cases without a potential conditioning factor for the non-occurrence of L2 (ECA 20.3% vs. EFA 29.2%) was surprisingly similar. There is a difference in the number of downstepped accents observed after the target word, but this is due to the fact that in the often shorter ECA stimulus sentences the target word more often appeared as the penultimate lexical item than in the EFA sentences. The L2 analysis does however reveal one other factor that appears to differ between the two varieties, namely the contour type used on the target word. Speaker BM and especially speaker CM used many more rising or high plateau contours on the target word in EFA (34.4%) than in ECA (2.2%). Although this observation does not hold for speaker AM, the high or rising contours of BM and CM were frequently accompanied by phrase breaks, matching our general findings regarding the incidence of phrase boundaries in EFA and ECA.

Conclusion

This paper used qualitative analysis to examine the occurrence of three variables, which were thought to potentially vary between EFA and ECA, namely the incidence of phrase breaks, secondary accents and 'L2' post-accentual valley turning points.

Two of the parameters investigated do show a difference between EFA and ECA as expected: EFA sentences showed almost twice as many utterance-internal phrase breaks than ECA sentences, and there was a higher incidence of secondary accents in EFA than in ECA. As already observed informally in Hellmuth & El Zarka (2007) however, we find that the occurrence of an L2 post-accentual target does not seem to differ between EFA and ECA, being observed roughly equally in both.

This paper thus provides additional evidence from read speech data to support the general hypothesis that the phonology of spoken formal Arabic reflects that of the speaker's mother tongue dialect. We again find the same broad phonological categories used in both EFA and ECA, with the clearest differences being at what might be termed a stylistic and/or rhythmic level, with greater use of phrasing boundaries in EFA and an increased number of pitch accents in EFA, brought about by the insertion of secondary accents.

There is however much more work to be done to fully understand the range of factors influencing intonational patterns in both EFA and ECA, particularly in naturally occurring speech, and we hope that the present paper will inspire others to engage in this work.

REFERENCES

Arvaniti, A., Ladd, D.R., & Mennen, I. 1998. Stability of tonal alignment: the case of Greek prenuclear accents. *Journal of Phonetics*, 26, 3-25.

Badawi, A.-S.M. 1973. Mustawayaat al-ʕarabiiya al-muʕaasira fii misˁr: baʕθ ʕilaaqat al-luɣa bi al- ħadˁaara. Cairo, Daar al-Ma'aarif.

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.

Boersma, P. & Weenink, D. 2006. Praat: doing phonetics by computer (Version 4.5.2).

El Zarka, D. & Hellmuth, S. 2007 The intonation of colloquial and formal Egyptian Arabic. (Paper presented at DGfS 2007, Siegen.).

El Zarka, D. 2008 Leading, linking and closing tonal contours in Egyptian Arabic. (Paper presented at the 22nd ALS Conference, Univ. of Maryland).

Ferguson, C.A. 1957. Two problems in Arabic phonology. *WORD: Journal of the International Linguistic Association*, 13, 461-479.

Hayes, B. 1995. *Metrical stress theory: principles and case studies.* Chicago, University of Chicago Press.

Hellmuth, S. 2006. Intonational pitch accen distribution in Egyptian Arabic. Unpublished PhD thesis, SOAS.

Hellmuth, S. & D. El Zarka (2007). Variation in phonetic realisation or in phonological categories? Intonational pitch accents in Egyptian Colloquial Arabic and Egyptian Formal Arabic. *Proceedings of the 16th ICPhS, Saarbruecken, Germany*.

Ladd, D.R. 1996. *Intonational phonology*. Cambridge, CUP.

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.

Lickley, R., Schepman, A., & Ladd, D.R. 2005. Alignment of 'phrase accent' lows in Dutch fallingrising questions: theoretical and methodological implications. *Language and Speech*, 48, 157-183.

McCarthy, J. 1979. On stress and syllabification. *Linguistic Inquiry*, 10, 443-466.

Mitchell, T.F. 1984. Soziolinguistische und stilistische Aspekte des gesprochenen Arabisch der Gebildeten in Ägypten und der Levante. *Sitzungsberichte der Sächsischen Akad.d.Wissenschaften in Leipzig, Philolog.-hist.Klasse*, 123, 1-21.

Prieto, P. & Shih, C. 1995. Effects of tonal clash on downstepped H* accents in Spanish. *Proceedings of EUROSPEECH'95*, 2, 1307-1310.

Prieto, P. & Torreira, F. 2006, in press. The segmental anchoring hypothesis revisited. Syllable structure and speech rate effects in Spanish. *Journal of Phonetics*.

Prieto, P., van Santen, J., & Hirschberg, J. 1995. Tonal alignment patterns in Spanish. *Journal of Phonetics*, 23, 429-451.

Rastegar-El Zarka, D. 1997. *Prosodische Phonologie des Arabischen.* Unpublished PhD thesis, Karl-Franzens-Universität Graz.

Rifaat, K. 1991. *The intonation of Arabic: an experimental study.* Unpublished PhD thesis, University of Alexandria.

Rifaat, K. 2003. The structure of Arabic intonation: a preliminary investigation. In Perspectives on Arabic Linguistics XVII-XVIII: Papers from the seventeenth and eighteenth annual symposia on Arabic linguistics. M.T.Alhawary and E.Benmamoun, eds. pp. 49-67. Amsterdam/Philadelphia, John Benjamins.

Yeou, M. 2004 Effects of focus, position and syllable structure on F0 alignment patterns in Arabic. JEP-TALN 2004.

APPENDIX

EFA dataset

qaamat bi-dˤarbi ħatta nazafati d-dimaaʔu min ʔanfi
bitnaa fii nazlati r-riyaadˤiyyiina fi l-qaryati l-ʔuulimbiyya
kaana l-ħukmu naazii qad tasabbaba fii xasaaʔira faadiħatin fi l-ʔarwaaħ
haaða ʃayʔun naazilun mina s-samaaʔ
laqad raʔaynaa nadʒafa fii maktabati l-dʒaamiʕa
laqad raʔaynaa nadʒwa fi l-mataadʒiri l-kabiira
laqad kaana naadʒi fii masraħiyyati l-dʒaamiiʕa
laqad nadamati l-mudarrisatu ʕalaa maa faʕalat
sawfa yaʃtarikuuna fi n-nadwati baʕda l-iltiħaaqi bi-l-dʒamʕiyya
ʔinna l-baaʔiʕa munaadin ʕala l-fawaakih
fi l-ħaqiiqati laqad namati n-nabataatu bi-surʕatin kabiira
wadʒadat salwaa namlatan fii ħaliibihaa
laqad naamati l-mudarrisatu baʕda t-tadriis
ʔinna l-ʔiskandariyyata qad ɣalabat iz-zamaalika min qabl
tamma taɣalɣulu n-nuuraaniyyiina fii maħaafili l-maasuuniyya
laa ɣaaliba wa-laa maɣluuba fii liqaaʔi l-qimmati l-ʔaxiir
laqad raʔaynaa malika s-swiidiyyiina ʕindama zurna s-swiid
laqad raʔaynaa malikata l-ʔurduniyyiina bi-l-ʔams
laqad raʔaynaa malikata l-huulandiyyiina bi-l-ʔams
duhiʃa t-talamiizu bi-malmalati l-mudarrisati l-gadiida
ʔinna tamalmula l-mudarrisiina l-gududi muʃkilatun kabiira
inxafadˤa tamalmulu l-mudarrisiina l-gududi baʕda t-tadriibi l-ħukuumi
maa tamaalaka l-ʔasaatiðatu ʕan tasˤħiiħi ʔaxtˤaaʔi tˤ-tˤullaab
ʔinnana qad zurnaa maalika l-mataadʒiri laylata l-xamiis
al-ʔaklu l-maaliħu laysa mufiidan li-sˤ-sˤiħħa
yuudʒadu tˤaaʔirun bismi maaliki l-ħaziin fii waadi n-niil
yuudʒadu ħaywaanun bismi manati fi l-wilaayaati l-muttaħida
ʔa-lam tasmaʕ ʕan qadˤiyyati manʕi l-mudawwanaati fii misˤr
ʔinnahu laa yuudʒadu maaniʕun fii binaaʔi kanaaʔisa dʒadiidatin fii haaðihi l-mantˤiqa
ʔuxtii munaa rasamati l-ʕimaarata l-waaqiʕata ʕala dˤaffati n-niil
haaðihi hiyya rasmatu l-ʕimaaraati l-waaqiʕati ʕala dˤaffati n-niil
laqad kaana raasim fii masraħiyyati l-dʒaamiʕa
muna qad inʃaɣalat bi-d-diraasa
iʕtabarnaa ʃaɣla l-mansˤibayni ʃ-ʃaaɣirayni muhimmatan baaliɣata l-ʔahammiyya
kaan haaða ʃ-ʃuɣla ʃ-ʃaaɣila fi-s-sanatayni l-maadˤiyatayni bi-n-nisbati lahu
laqad ʔakaltu tiinatan wa balaħatan baʕda l-ɣadaaʔ
wa-baʃʃarahuu bi-l-dʒannati ʕalaa balwaa tusˤiibuh
haaða l-mawdˤuuʕu yaʃɣalu baalii munðu ʔayyaamin kaθiira
laqad wadʒadna l-walada fi l-ħadiiqa
inxafadˤa muʕaddalu l-wafayaati fi-l-waldati fii ʔaryaafi misˤra xilaala l-qarni s-saabiq
samiʕnaa tawalwula n-nisaaʔi baʕda wafaati ʃ-ʃayx
laqad taħaddaθna maʕ waalidi tˤ-tˤaalibatayni tˤawiilan

ECA dataset

ʃufna nagafa fi l-maktaba illi gambina
ʃufna nagwa fi d-dakakiin illi gambiina
ʃufna naagi fi masraħiyyit il-gamʕa imbaariħ
muna nadahit ʕala saami
ʃufna nadya fi d-dakakiin illi gambina n-nahardʿa
ʃufna naadi fi d-dakakiin illi gambiina imbaariħ
il-bint is-swidiyya ʕandaha namaʃ ʕala wiʃʃaha
salwa laʔit namla fi l-makaroona
fatʿma naamit baʕd il-madrasa
miʃ hastaslim mahma nazafit guruuħi
mʿaamʿa kat nazla min itʿ-tʿayyaara
bʿaabʿa kaan naazil min itʿ-tʿayyaara
ʃufna malik is-swidiyyiin lamma ruħna s-swiid
ʃufna malikit il-hulandiyyiin lamma ruħna hulanda
ʃufna malikit il-ʔurduniyyiin imbaariħ
istaɣrabu t-talamza bimalmalt il-mudarrisa l-gidiida
ruħna malʕab il-ʔalmaan lamma ruħna ʔalmanya
zurna malgaʔ il-masakiin iʃ-ʃahr illi faat
iʃ-ʃurba l-malħa hiyya miʃ kuwayyisa li-sʿ-sʿaħħa
il-ʔakl il-maaliħ huwwa miʃ kuwayyis li-sʿ-sʿaħħa
fii tʿeer ismu maalik il- ħaziin mawguud fi waadi n-niil
zurna maalik id-dukkaan illi gambina imbaariħ
ʃufna maalik id-dakakiin il-gidiida illi fi madiinit nasʿr lamma ruħna l-mool
ʃufna ħaywaan ismu manati lamma ruħna flurida
min ʃaʔʔitna l-gidiida il manzʿar ʕala l-gineena mumtaaz
mafiiʃ ʕandi maaniʕ inni k-kanaayis titbini tʿaalama l-ʕadad miʃ kifaaya
ʔuxti muna rasamit il-ʕimaara illi ʔuddaam il-baħr
ʔaadi rasmit il-ʕimaraat illi ʔuddaam il-baħr
muna lissa rasma l-ʕimarteen illi ʔuddaam il-baħr
saami lissa raasim il-ʕimaraat illi ʔuddaam il-baħr
iskindiriyya ɣalabit iz-zamaalik ʔabl kida
iskindiriyya lissa ɣalba l-ismaʕiliyya
iz-zamaalik lissa ɣaalib il-baladiyya
muna inʃaɣalit bi-d-dirʿaasa
il-mawadîiʕ kaanit ʃaɣla kull ʔafkaaru
il-mawdʿuuʕ kaan ʃaaɣil kull ʔafkaaru
il-mawdʿuuʕ ʃaaɣil baalak min ʔasabiiʕ
ʔakalt tiina wi-balaħa baʕd il-ɣada
id-dawa tʿallaʕ il-balɣam illi fi sʿidru
xalli baalak min it-talamza
ʃufna l-walad illi gambina fi l-mitru
ʔuxti lissa walda min ʃahreen
ʃufna waalid isʿ-sʿuɣayyariin fi s-sinima