VARIATION IN PHONETIC REALISATION OR IN PHONOLOGICAL CATEGORIES? INTONATIONAL PITCH ACCENTS IN EGYPTIAN COLLOQUIAL ARABIC AND EGYPTIAN FORMAL ARABIC

Sam Hellmuth & Dina El Zarka

Universität Potsdam, Universität Graz
sam.hellmuth@uni-potsdam.de, dina.elzarka@uni-graz.at

ABSTRACT

This paper uses qualitative and quantitative methods to compare the intonation of formal and colloquial varieties of Egyptian Arabic in a corpus of elicited read speech, to explore the widely held assumption that spoken formal Arabic will have the intonational characteristics of the speaker's colloquial variety. Speakers are found to use broadly parallel phonological systems in each register, reflected in parallel distribution and type of pitch accents. A quantitative analysis of the pitch target alignment to the segmental string reveals only minor differences in the phonetic realisation of pitch accents across registers.

Keywords: Arabic, intonation, sociolinguistic variation, pitch accent alignment.

1. INTRODUCTION

This paper explores the hypothesis that the intonation of spoken formal Arabic ('fusha') will display the same intonational properties as those of the mother tongue dialect of the speaker ('ʕaammiyya'). This hypothesis arises from the commonly held assumption that, in general, the phonological (and phonetic) properties of spoken formal Arabic are transferred from the speaker's colloquial dialect. For example, although all spoken Arabic dialects display quantity sensitive stress assignment, the exact rules for assigning primary stress vary from dialect to dialect [12]; in spoken formal Arabic speakers are expected to apply the particular stress assignment rules of their own mother tongue dialect [4]; studies have shown this to be the case for Egyptian Arabic [e.g. 19].

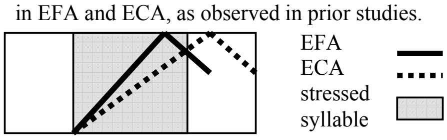
The distinction between standard and spoken varieties of Arabic is complex, and has been the subject of much attention in the literature [3, 7]. We adopt here Mitchell's three way classification [20] whereby the primary distinction is between *formal* and *informal* Arabic, with a sub-division of informal Arabic into 'careful' and 'casual' registers. Our study addresses possible intonational differences between Egyptian Formal Arabic (EFA), defined as the rendition of formal Arabic ('fusha') by Egyptians, and Egyptian Colloquial Arabic (ECA, 'ʕaammiyya').

Cross-linguistic intonational variation is itself a relatively new area of research, although a number of possible parameters of intonational variation have been suggested [8-11, 15, 16]. Ladd [16:119] proposes a taxonomy of intonational variation in four categories: semantic (different meaning/use of phonologically identical contours), systemic (different inventory of phonologically distinct contours), realisational (differences of detail in the phonetic realisation of what is phonologically the same contour) and phonotactic (differences in contour-text association).

Variation between EFA and ECA could occur in any of these four areas, however semantic and/or systemic differences require investigation in a large corpus of (preferably) spontaneous speech, which is beyond the scope of the present study. We thus focus here on the search for possible variation in just two categories: realisational differences, which we suggest would constitute evidence of *phonetic* variation between the two varieties, and phonotactic differences of contour-text association, to yield evidence of potential *phonological* variation. To this end our study comprises both qualitative and quantitative analysis.

In particular we explore potential differences in the alignment of tonal events (L and H pitch targets) to the segmental string, since earlier studies suggest H peaks may align earlier in EFA (within the stressed syllable [24]) than in ECA (early in the second mora of the stress foot [13]) as illustrated in Figure 1.

Figure 1: Schematised peak alignment in CV syllables in EFA and ECA, as observed in prior studies.



An extensive body of 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 [22], but also extensive cross-linguistic variation in the fine detail of alignment patterns [1, 17]. For example, in a study of alignment in varieties of German, Atterer & Ladd [2] found consistent differences of alignment between Northern and Southern German, which they analyse as cross-dialectal variation in the phonetic realisation of a single phonological object (a pre-nuclear rising pitch accent), both of which are however quantitatively different from alignment patterns in, say, English. Indeed, in a corpus of speech in English produced by the same two sets of German speakers, German-like alignment patterns were found, and speakers did not produce English-like alignment patterns.

The fine detail of alignment patterns can thus in principle serve as an indicator of realisational differences within a single phonological category (as the German North/South distinction) or of the transfer of phonotactic contour-text association patterns from one variety to another (as in the survival of dialectal alignment differences in speakers' L2 productions). We therefore take the fine detail of tonal alignment in EFA and ECA to be a good potential source of evidence of variation in phonetic and/or phonotactic realisation across the two varieties, and the body of research into alignment variation allows us to investigate the question using established methodology. In the remainder of the paper we set out the details of the materials and analysis employed (§2), a survey of the qualitative and quantitative results (§3), a discussion (§4) and a brief conclusion (§5).

2. METHODOLOGY

To test for variation in phonetic realisation between EFA and ECA (with alignment of pitch targets to the segmental string as our dependent variable), pairs of target words were sought containing parallel segmental content in the stressed syllable in each variety. 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. To facilitate location of F0 events in the pitch track, we sought target words with sonorant consonants around the stressed vowel, but due to the limited number of lexical items meeting these criteria we were able to create only 6 such lexical sets, and thus also included in the design a further 6 lexical sets in which there is greater variety of segmental content in the stressed syllable. The 12 lexical sets per syllable type, in each of two varieties (EFA/ECA), yield 12x3x2=72 target words in total.

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

	CV	CVV	CVC
ECA	<u>ma</u>lik 'king'	<u>maa</u>lik 'owner'	<u>mal</u>Ha 'salty'
EFA	<u>ma</u>lik 'king'	<u>maa</u>lik 'owner'	ta<u>mal</u>mul 'nervousness'

Each target word was placed in its own frame sentence designed to be as natural as possible and also to facilitate elicitation of the correct register of Arabic (see Figures 2 & 3 for sample sentences). The prosodic context was controlled to avoid tonal clash from a following adjacent accented syllable, and a target word was sometimes placed in more than one carrier sentence, to ensure at least one useable rendition. A total of 43 ECA sentences and 42 EFA sentences resulted. These were pseudo-randomised and interspersed with distractor sentences in the relevant register and presented to speakers on a computer screen one at a time typed in Arabic script. For ECA we used Egyptian lexical items and spelling (i.e. as used in cartoons) to elicit the correct register.

The full set of materials were read three times each in EFA and then, after a break, three times each in ECA, by two speakers ('A' and 'B'), yielding a parallel corpus of 126 tokens in EFA and 129 tokens in ECA for each speaker (255x2=510 tokens in total). Both speakers were male, aged under 30, born and raised in Cairo, and mother tongue speakers of ECA. Both have university level formal training in the grammar of standard Arabic and are fluent speakers of EFA. Digital recordings were made at 44.1KHz 16bit directly to digital format in a quiet room (re-sampled to 22.05KHz 16bit). Due to technical recording difficulties 26 tokens (2 + 24 from each speaker respectively) were discarded, leaving 484 tokens for analysis.

The corpus was submitted to qualitative and quantitative analysis with reference to F0 contour and wideband spectrogram using Praat 4.5.2 [5]. Each sound file was annotated by the first author and each annotation subsequently checked by the second author; discrepancies were re-checked by both authors in order to reach consensus. All 484 target sentences were included in this qualitative analysis phase, one of whose functions was to identify tokens which were suitable for inclusion in the quantitative analysis. The following cases were tagged for exclusion: tokens containing a disfluency or assignment of word stress in an unexpected position in the target syllable, tokens in which an additional accented syllable results in less than three intervening syllables between the target syllable and the following accented syllable (a 'clash'), and tokens in which the speaker inserts a phrase boundary after the target word ('boundary') or places focus on the target word itself ('focus'). The prosodic effects of clash, boundary and focus on the alignment of pitch events relative to the segmental string are well-documented [21, 23, 27]; hence these tokens were excluded from quantitative analysis.

For the quantitative analysis the segmental and pitch events listed in Table 2 were labelled by hand by the first author and checked by the second author. We calculated the following derived variables following [23]: L-C0 'L to onset of stressed syllable' and 'peak delay' (H-C0, 'H to onset of stressed syllable'), as well as the duration of the stressed syllable, to calculate a proportional measure of peak delay relative to the duration of the stressed syllable: 'relative peak delay' (rpd).

Table 2: 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

3. RESULTS

3.1. Results of the qualitative analysis

The qualitative analysis of 484 tokens confirms that EFA and ECA share two key intonational properties. Firstly, both varieties display highly populated pitch accent distribution, with an accent realised on almost every content word (cf. for ECA [13] and for EFA [24, 25]). Secondly, in the overwhelming majority of cases, in non-phrase-final ('pre-nuclear') position, the accent shape placed on each content word is the same: an accent whose most salient property is an H peak realised towards the end of the stressed syllable. These properties are illustrated in Figures 2 and 3.

Figure 2: ECA token (speaker A).

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

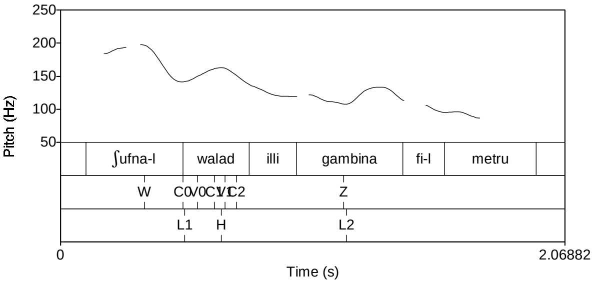
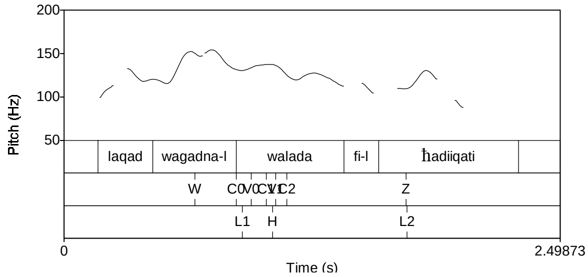


Figure 3: EFA phrase boundary token (speaker A).

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



The properties of EA pre-nuclear pitch accents have received different phonological analyses, variously in EFA as H*+L [24] and as H* (with flanking fall and rise) [26], and in ECA as L+H* [13]. Our present study suggests that this apparent variation may well 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. Close observation of the positioning of the L2 turning point reveals a sizeable number of cases in which there is a low plateau between successive H peaks, rather than a 'single' valley low turning point, but this 'plateau' strategy is used by both speakers in both registers, thus we do not explore it further here. (Due to this variation in L2 behaviour, we confine the quantitative analysis below to the patterning of the L1 turning point immediately before the H peak.)

However, the qualitative analysis does reveal differences between the two registers in what may be termed the more global properties of our target sentences. In particular we observe a greater proportion of sentences containing an internal phrase boundary in EFA than in ECA. This holds for both speakers but most clearly so for speaker B. The number of tokens containing boundaries, as a proportion of total fluent tokens is as follows: speaker A, ECA 24.5%, EFA 54.2%; speaker B, ECA 9.0%, EFA 37.1%. A sample EFA token containing a boundary is provided in Figure 3 above. We suggest that greater use of phrasing boundaries is part of a wider strategy used by speakers to enhance rhythmicity in EFA [6].

3.2. Results of the quantitative analysis

Following recent proposals [18] we use two types of dependent variable to investigate the alignment of pitch events to landmarks in the segmental string: when reporting absolute distances from a pitch event to a segmental landmark we report relative to the closest suitable landmark, in order to minimise variance; when comparing over lexical sets containing different segmental material we report a proportional variable 'relative peak delay' (rpd) which expresses 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 A, 91 tokens: ECA 58 + EFA 33; for speaker B, 136 tokens: ECA 71 + EFA 65).

3.2.1. Alignment of L1

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, and 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 one-way ANOVA by register, with L1-C0 as the dependent variable (performed for each speaker independently), indicates that the position of L1 does not vary significantly between registers for speaker A ($F(1,89) = 0.300$; $p = .585$; $\alpha = 0.05$); for speaker B the variation in the distribution of L1-C0 values is too great to permit a statistical test to be performed. We thus take 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.

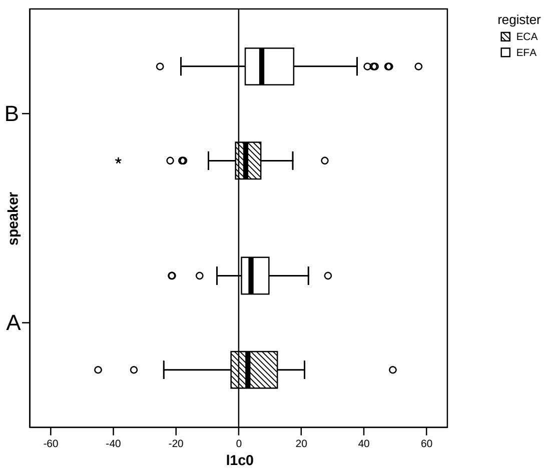
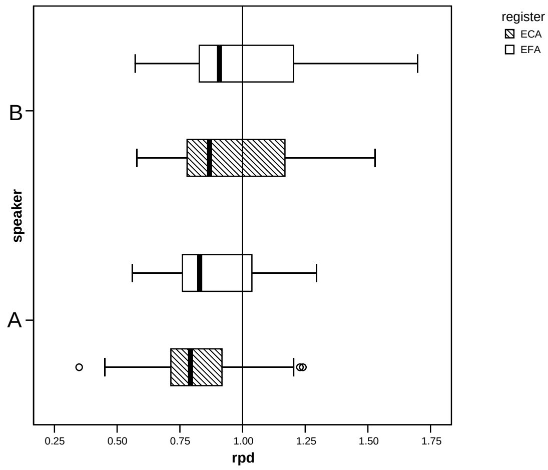


Figure 5: Median and interquartile values of relative peak delay (H-C0/syllable duration) in ms. The vertical line indicates the end of the accented syllable; EFA plain box, ECA striped box.



3.2.2. Alignment of H

To investigate the position of the H peak we use a proportional measure. Figure 5 displays the median and interquartile ranges of 'relative peak delay' (rpd) by speaker and by register. 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 A. 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 A only: speaker A (F (1,89) = 4.525; p = 0.036; α = 0.05); speaker B (F (1,134) = 1.676; p = 0.198; α = 0.05). Thus, H aligns differently in ECA and EFA for speaker A only, while speaker B aligns H in parallel fashion in ECA and EFA.

Figure 6: Median and interquartile values of relative peak delay (H-C0/syllable duration) in ms. for speaker A by syllable type (CV plain box, CVV striped box, CVC hatched box; EFA at top, ECA at bottom).

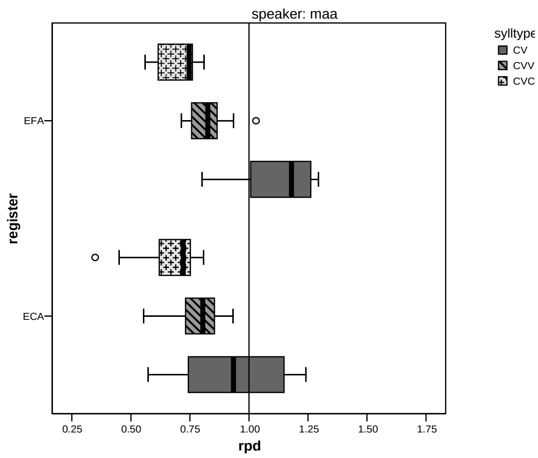
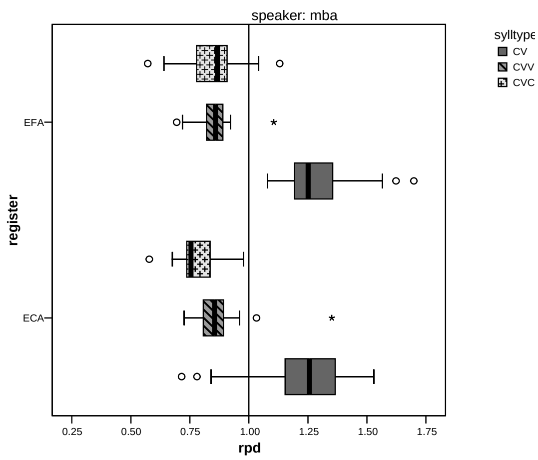


Figure 7: Median and interquartile values of relative peak delay (H-C0/syllable duration) in ms. for speaker B by syllable type (CV plain box, CVV striped box, CVC hatched box; EFA at top, ECA at bottom).



The nature of the difference between ECA and EFA for speaker A is revealed when we look at alignment in different syllable types individually. Figures 6 & 7 display the median and interquartile ranges for values of rpd for each speaker individually, by register and by syllable type. The vertical line indicates the syllable end (CV.CV/CVV./CVC.). For both speakers, in CVV and CVC syllables H is aligned within the accented syllable, in both registers. In CV syllables however a different pattern emerges. Speaker B (Figure 7) aligns H mostly outside the syllable (within the intervocalic consonant) in both EFA and ECA. In EFA speaker A (Figure 6) also aligns H within the intervocalic consonant, but in ECA H falls most often within the syllable. The difference in H alignment between EFA and ECA for speaker A is thus due mostly to a difference in the alignment behaviour of H in CV syllables. For speaker A, a oneway ANOVA (within each syllable type independently) indicates that the difference in alignment between registers is significant in CV syllables only (CV: F (1,33) = 5.852, p = 0.021, α = 0.05; CVV: F (1,31) = 0.988, p = 0.328, α = 0.05; CVC: F (1,21) = 0.307, p = 0.036, α = 0.05). For speaker B a parallel analysis indicates no significant differences in H alignment between registers.

In sum then we find a small quantitative difference in H alignment between EFA and ECA, but in only one syllable type (CV) and for only one speaker (speaker A).

4. DISCUSSION

The pattern observed in speaker A's ECA CV syllables is similar to that observed in the corpus of broadcast EFA in [24]. We suggest that this earlier alignment of H may in fact be a hallmark of 'professional speech' in both ECA and EFA, rather than a hallmark of EFA itself, since speaker A's productions in both registers were in general more deliberate (or planned) than those of speaker B. Indeed it is possible that we have here in fact elicited three levels of Arabic, by Mitchell's [20] classification: formal (speakers A and B in EFA), informal careful (speaker A in ECA), and informal casual (speaker B in ECA).

The alignment difference observed in this 'professional speech' register may plausibly be an example of phonological (phonotactic) variation. The patterns of alignment observed in speaker B's productions are consistent with the analysis that the H pitch target seeks to align with the second mora of the stress foot [13, 14]. The stress foot in

EA is a moraic trochee consisting of either one heavy or two light syllables [12], thus whilst in CVV and CVC syllables H aligns within the syllable (co-extensive with the foot) in CV syllables the H aligns within the following syllable, containing the second mora of the foot. The pattern in speaker A's ECA productions, in which H aligns within the syllable in all syllable types, suggests that the H target may be seeking instead to align with the end of the syllable.

5. CONCLUSION

This paper provides evidence from intonational patterns to support the general hypothesis that the phonology of spoken Standard Arabic reflects that of the speaker's mother tongue dialect. We find the same broad phonological categories used in both EFA and ECA, with the clearest differences being at the global level, in greater use of phrasing boundaries in EFA. There is some indication of a possible alignment difference, which we suspect reflects patterns of speech used by 'professional speakers' in both ECA and EFA, but further investigation is required before this can be definitively classified, as a difference either of phonetic realisation or of intonational phonotactics in the two registers of spoken Arabic.

6. REFERENCES

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

[2] 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.

[3] Badawi, A.-S.M. 1973. *Mustawayaat al-ʕarabiiya al-muʕaasira fii misr: bahθ ʕilaaqat al-luγa bi al-hadaara.* Cairo, Daar al-Ma'aarif.

[4] 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.

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

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

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

[8] Fitzpatrick, J. 2000. On intonational typology. *In Methodological issues in language typology*. P.Siemund, ed. pp. 88-96.

[9] Fox, A. 1999. *Prosodic structure and features.* Oxford, OUP.

[10] 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.

[11] Gussenhoven, C. 2004. *The phonology of tone and intonation.* Cambridge, CUP.

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

[13] Hellmuth, S. 2006. *Intonational pitch accent distribution in Egyptian Arabic.* Unpublished PhD thesis, SOAS.

[14] Hellmuth, S. 2007. The foot as the domain of tonal alignment of intonational pitch accents. *Proceedings of the 16th ICPhS, Saarbruecken, Germany*.

[15] Jun, S.-A. 2005. Prosodic typology. *In Prosodic transcription and typology: a unified approach.* S.-A.Jun, ed. pp. 430-458. Oxford, OUP.

[16] Ladd, D.R. 1996. *Intonational phonology*. Cambridge, CUP.

[17] 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.

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

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

[20] 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.

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

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

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

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

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

[26] Rifaat, K. 2005. 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.

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