THE FOOT AS THE DOMAIN OF TONAL ALIGNMENT OF INTONATIONAL PITCH ACCENTS

Sam Hellmuth

Universität Potsdam
sam.hellmuth@uni-potsdam.de

ABSTRACT

This paper presents evidence from an experimental investigation of the alignment properties of pre-nuclear rising pitch accents in Egyptian Arabic, across different syllable types (CV, CVV and CVC). In both CVC and CVV syllables, the H peak is aligned within the second mora of the syllable, but in CV syllables the H peak appears in the following (light) syllable. We argue that these generalizations support adoption of the (metrical) foot as the domain of tonal alignment of intonational pitch accents in Egyptian Arabic, and discuss this finding in the context of current debate regarding tonal alignment in intonation languages.

Keywords: intonation, tonal alignment, tonal domain, foot, Egyptian Arabic.

1. INTRODUCTION

A large body of research has built up in recent years regarding the best way to understand the tonal alignment properties of intonational pitch accents (see surveys in [9, 14, 15]). Cross-linguistic evidence suggests that, other things being equal (that is, in the absence of influence from the surrounding prosodic environment, such as a prosodic boundary), individual L and H pitch targets in a bitonal rising pitch accent may each display independent 'segmental anchoring' to landmarks in the segmental string, as first shown for Greek [1]. However, there is also increasing evidence that strict interpretation of this finding in autosegmental terms cannot fully capture the fine phonetic detail of tonal alignment, such as individual behaviour of pitch targets whereby the L tone is typically closely anchored to the start of the stressed syllable but alignment of the H target is more variable [15].

Ladd [9] argues that the facts of segmental anchoring suggest intonational pitch accents be analysed as gestures (in the sense of [4]), which display strong association to some tonal domain, such as the stressed syllable, foot or word. This paper presents evidence of alignment across syllable types in colloquial Egyptian Arabic (henceforth EA, defined here as the spoken dialect of Cairo) to support adoption of the foot as the appropriate domain of tonal alignment in EA.

The most common pre-nuclear pitch accent in colloquial EA has been observed to be rising, with the peak situated at or near the end of the stressed syllable [13, 17]. EA stress assignment is based on bimoraic trochaic feet [7], which can be either monosyllabic (CVV,CVC) or bisyllabic (CV.CV). EA also has a phonemic vowel length distinction. It is therefore possible to test whether pitch targets in EA pre-nuclear pitch accents are best understood in terms of association with the syllable or the foot as the domain of tonal alignment by comparing pitch target alignment in all three types of possible non-word-final stressed syllable types (CV, CVC and CVV): if the syllable is the domain of tonal alignment in EA, then patterns of alignment observed in all three syllable types should be parallel; if the foot is the relevant tonal domain then alignment will be later in CV syllables (initial in a CV.CV bisyllabic foot) than in CVC and CVV syllables (monosyllabic feet).

2. METHODS

2.1. Materials

In order to test the hypothesis that the domain of tonal alignment in EA is the foot, 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, we sought target words 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, and each was placed early in a sentence frame to elicit a pre-nuclear pitch accent. The 18 target sentences were pseudo-randomised and interspersed with distractors and presented to speakers typed in Arabic script, using EA spelling conventions and lexical items, to elicit colloquial spoken EA.

The 18 sentences were read 3 times each by 15 speakers of EA (9 male, 6 female), yielding a potential 810 tokens. All participants were mother tongue speakers of EA, born and raised in Cairo, aged 21-34 years, with no reported auditory or speech production difficulties; subjects were paid a small fee for their participation. Digital recordings were made in Cairo at 44.1KHz 16bit using ProTools 6.0 on MBox in a draped classroom using a head microphone. Each sound file (re-sampled at 22.05KHz 16bit) was segmented and labelled by the author with reference to F0 contour and spectrogram extracted using Praat 4.2 at default settings [2]. Tokens containing a disfluency (N=59) or phrase boundary (N=42) on or near the target word were discarded, leaving 709 tokens for analysis.

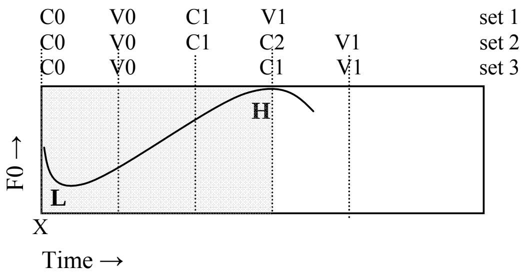
2.2. Analysis

Segmental landmarks and pitch events in target syllables were labelled as in Table 1 and Figure 1.

Table 1: 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
V1	start of following vowel
C2	end of coda consonant (in closed syllables)
X	left edge of the word
L	start of the F0 rise
H	end of the F0 rise

Figure 1: Schematised labelling diagram of segmental landmarks and pitch events.



[malik]	m	a	l	i	k		set 1 CV.CV
[manga]	m	a	n	g	a		set 2 CVC
[ma:liH]	m	a:		l	I	H	set 3 CVV

Standard alignment variables were calculated, such as L-C0 'L to onset of stressed syllable' and H-C1 'H to end of stressed vowel', as well as Peak Delay (H-C0, 'H to onset of stressed syllable' [16]). In addition the duration of a number of potential tonal domains was calculated as shown in Table 2.

Table 2: Measures of stressed syllable/foot duration.

label	variable	CV.CV	CVC	CVV
Canonical syllable	*sylldur1*	C1-C0	C2-C0	C1-C0
Strong syllable	*sylldur2*	V1-C0	C2-C0	C1-C0
Foot Duration	*footdur*	Y-C0	C2-C0	C1-C0
Word Duration	*worddur*	Y-X	Y-X	Y-X

Two definitions of 'syllable' are included, both the canonical syllable (CV, CVC and CVV), and the strong syllable, an ambisyllabic conception of the syllable which involves incorporating into the stressed syllable the foot internal intervocalic consonant in sequences of light syllables ['CVC.V']. Two main types of analysis were carried out, following [11], using both 'close' alignment variables such as L-C0 and H-C1, which indicate the distance of a pitch target from its closest segmental landmark, as well as proportional measures of peak delay (H-C0) relative to the various potential tonal domains.

Since 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, it was not possible to find suitable words with only one stressed vowel quality. Similarly, in order to ensure elicitation of colloquial rather than formal EA, target 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. A series of linear regressions were carried out to assess 'Vowel Quality' and 'Clash Context' (encoded as categorical variables, vowel type/ number of intervening syllables) as potential predictors of Peak Delay, alongside Speaker and Syllable Duration (the 'strong syllable' was used). 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, followed by Speaker (3.09%). The effects of following Clash, preceding Clash and Vowel Quality were negligible.

3. RESULTS

Mean values of the closest measures of L and H alignment provide a first picture of the alignment facts of EA, as listed in columns 1-3 of Table 3.

Table 3: Mean and standard deviation of alignment and other variables in milliseconds by syllable type.

Set #		L1-C0	H-C1	H-V1	H-C0	sylldur1
Set 1	Mean	7.2	43.5	-5.3	177.9	134.5
CV	N	215	215	215	215	215
	SD	26.3	37.4	36.0	48.4	29.6
Set 2	Mean	16.2	25.9	-114.5	155.5	200.5
CVC	N	249	249	249	249	249
	SD	29.3	29.1	46.7	36.6	43.5
Set 3	Mean	10.4	3.6	-48.2	177.4	173.8
CVV	N	245	245	245	245	245
	SD	33.6	46.5	46.7	57.4	44.8
Total	Mean	11.5	23.5	-58.5	169.9	171.3
	N	709	709	709	709	709
	SD	30.2	41.5	62.5	49.2	48.3

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 within the 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 Figure 2 below.

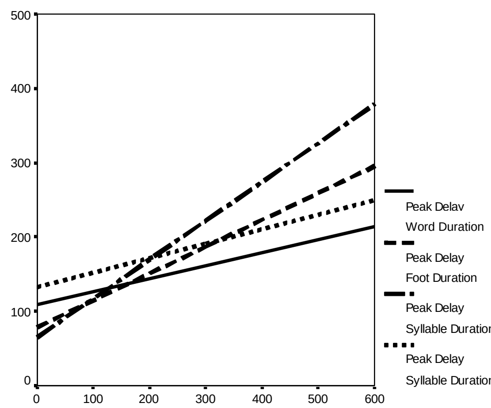
Figure 2: Schematised alignment of pitch events.

	L		H			
a.	m	a	l	i	k	
	L		H			
b.	m	a	n	g	a	
	L			H		
c.	m	a:		l	i	ħ

The surface alignment facts in EA seem to pattern closely with those observed in [10] for 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 (in the coda in CVC and at the end of the stressed vowel in CVV).

Whilst absolute values of peak delay 'H-C0' appear similar in CV and CVV syllables (column 4 Table 3), CVV syllables are on average longer than (canonical) CV syllables (column 5 Table 3). We therefore used correlations to measure peak delay as a proportion of the four potential domains of tonal alignment: canonical syllable, strong syllable (whose status as a prosodic constituent is discussed below), foot and word. The durations of all four domains prove to be significantly correlated to peak delay 'H-C0' (Spearman's Rank: rho; $p<0.001$), ruling out a fixed delay analysis of H alignment. However, as illustrated in Figure 3, the highest correlation is between peak delay and sylldur#2, suggesting the strong syllable as most appropriate tonal domain.

Figure 3: Peak delay x 4 tonal domain durations.



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, alongside each of the four potentially explanatory domains: canonical syllable (*sylldur1*), strong syllable (*sylldur2*), foot (*footdur*) and word (*wddur*). The domain which best predicts the variation in the model is again the strong syllable (36.4%), followed closely by the foot (30.7%)

4. DISCUSSION

The structural domain which is most closely correlated to peak delay in the EA data appears to be the 'strong syllable', matching results obtained for Dutch [10]. Prieto & Torreira [15] find no evidence to support anchoring of H to the syllable in Spanish, basing their calculations on a canonical definition of the syllable. However, their finding that H aligns later in closed than in open syllables would be consistent with treating the domain of tonal alignment in Spanish as the strong syllable.

There is, however, no such prosodic constituent as the 'strong syllable'. Although the notion of ambisyllabicity has been used to analyse a range of phonological phenomena ([3, 5, 12], the context of these can also be shown to generalise to a more traditional prosodically defined environment, in that they occur foot-internally [6, 8]. I suggest therefore that alignment of H peaks in EA inside the intervocalic consonant in CV.CV feet is an indication that the tonal domain of the rising pitch accent gesture in EA is the foot.

Although the head of the foot is leftmost in EA (a bimoraic trochee), it is arguably the second mora that appears to attract the perceptually salient H tone, though with considerably great variation in the positioning of H than of the initial L target (cf. [15]. Alternatively, one could argue that the H tone is phonologically *associated* with the leftmost head mora of the foot but that production restrictions mean we expect the peak to in fact align just after the strong mora - that is, in the second mora of the foot, as observed: cf. [18] who show that speakers need at least 124 ms to achieve a 4 semitone pitch rise (the mean F0 excursion of rises in the present dataset is 3.2 semitones).

Under either scenario however, the claim of this paper is that across syllable types the position of the H peak of rising pitch accent gestures in EA is best understood relevant to the foot as the tonal domain to which the whole gesture is associated.

5. CONCLUSION

This paper investigated the surface phonetic alignment of pitch targets in EA rising pre-nuclear pitch accents, on the basis of 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 simplest generalisation describing the alignment patterns of the H tone is that it falls somewhere within the second mora of the foot. We argue that the simplicity of this generalisation speaks for adoption of the foot as the domain of tonal alignment in EA.

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] Boersma, P. & Weenink, D. 2004 *Praat: doing phonetics by computer (Version 4.2)*.

[3] Borowsky, T., Itô, J., & Mester, R.-A. 1984. The formal representation of ambisyllabicity: evidence from Danish. *NELS*, 14, 34-48.

[4] Browman, C. & Goldstein, L.M. 1986. Towards an articulatory phonology. *Phonology Yearbook*, 3, 219-252.

[5] Hall, T.A. 1989. Lexical Phonology and the distribution of German [ç] and [x]. *Phonology*, 6, 1-17.

[6] 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.

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

[8] Jensen, J.T. 2000. Against ambisyllabicity. *Phonology*, 17, 187-235.

[9] Ladd, D.R. 2006. Segmental anchoring of pitch movements: autosegmental phonology or gestural coordination? *Rivista di Linguistica*, 18, 19-38.

[10] 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.

[11] 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.

[12] McCarthy, J. 1993. A case of surface constraint violation. *Canadian Journal of Linguistics/Revue Canadienne de Linguistique*, 38, 169-195.

[13] Norlin, K. 1989. A preliminary description of Cairo Arabic intonation of statements and questions. *Speech Transmission Quarterly Progress and Status Report*, 1, 47-49.

[14] Prieto, P., D'Imperio, M., & Gili Fivela, B. 2006. Pitch accent alignment in Romance: primary and secondary associations with metrical structure. *In Intonation in language varieties.* P.Warren, ed.

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

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

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

[18] Xu, Y. & Sun, X. 2002. Maximum speed of pitch change and how it may relate to speech. *J.Acoust.Soc.Am.*, 111, 1399-1413.