DURATION AND PITCH ANCHORING AS CUES TO WORD BOUNDARIES IN GREEK

Argyro Katsika

Yale University
argyro.katsika@yale.edu

ABSTRACT

This investigation is part of a larger study of the role of fine phonetic details in word segmentation in Greek connected speech. The present paper investigates whether and how Greek speakers use durational and pitch alignment acoustic cues to mark word boundaries in identical segmental strings differing only in the word boundary affiliation. Duration modification mechanisms are evident in cuing words, while different F0 alignment is not detected.

Keywords: word segmentation, duration modification mechanisms, pitch anchoring.

ACKNOWLEDGMENTS

My deep gratitude goes to Rachel Smith (University of Glasgow) and Francis Nolan (University of Cambridge) for their support and suggestions. I would also like to thank all the participants for their contribution to this research.

1. INTRODUCTION

Words in connected speech undergo variations caused by the sound forms of adjacent words and a range of other factors. It has traditionally been assumed that, as a consequence of this variation, the signal is phonetically ambiguous. But, this phonetic ambiguity might be resolved by the lexical and contextual information [4] revealed by the signal's fine phonetic details. This study seeks to establish whether native Greek speakers use particular acoustic features (duration and pitch alignment) to mark word boundaries in their speech, creating potential perceptual underpinnings for listeners.

Although durational acoustic cues to *phrase* boundaries are well established, attempts to discover whether such systematic effects apply to *word* boundaries have given conflicting results (cf. [8], [10], [11], [14]). As for pitch alignment, researches have shown that F0 landmarks are aligned with identifiable and predictable points of the segmental string ([1], [2]), with such anchoring functioning as a cue for word segmentation (cf. [6], [9]). Does Greek provide evidence for use of acoustic durational cues at word level and is there F0 anchoring on Greek word edges? This study is of intrinsic interest for Greek phonetics, with reference to cross-linguistic comparisons and implications to psycholinguistic and computational models of word processing.

2. ANALYSIS PROCEDURES

Five male native Standard Greek speakers aged between 23 and 28 participated in a speech production experiment. Utterances containing identical segmental strings differing on the word boundary affiliation were elicited: the nasal consonant /n/ (in order to avoid local perturbation in F0) constituted the coda of the monosyllable singular masculine article in the one member of the minimal junctural pair [/tɔn#ˈɐnɔ/ (*the upper*)] and the stem's onset of the singular neutral article's accompanied noun or adjective [/tɔ#ˈnɐnɔ/ (*the dwarf*)] in the other member of the pair. The noun modified by this article was placed in a context designed to elicit the pitch accent L+H* (intonational hat-pattern). All junctural pairs were embedded in the same elliptical meaningful sentence. 10 filler sentences were used. There was a dialogue session, followed by two reading tasks, one in casual and one in careful speech. 5 repetitions of each sentence were analyzed. Sentences' order was random.

Recordings were carried out in a soundproof booth. After digitization at a sampling rate of 16 kHz and conversion to wave files, the data were analyzed using the phonetic software package Praat. Duration was measured at the edges of each segment of the minimal junctural pairs except the initial stop of /tɔ(n)/. Pitch was recorded at the edges of the test segments, at the midpoints of the ambiguous consonant /n/ and the vowels preceding and following it, and at the lowest and highest points of the pitch contour (judged by eye from the pitch contour superimposed upon the spectrogram).

It was hypothesized that word-final [3] and word-initial lengthening ([10], [11], [14]) would modify the durational length of the pre-word boundary and post-word boundary segment/syllable respectively. Additionally, the F0 minimum was expected to be placed before or early within the ambiguous consonant when the latter is attached to the post-boundary word, and after or late within the consonant when it belongs to the pre-boundary word [6].

3. DURATIONAL MODIFICATIONS

This section examines the durational and pitch properties of /ɔ/ of the definite article, the ambiguous /n/, the vowel following /n/ (VL2), the segment preceding the final vowel (BFVL3), the final vowel (FVL3) and the durational sum of the two latter (FSYL), in both junctural cases. Table 1 illustrates the segmental material that falls into each category for each junctural pair.

Table 1: Segmental content of the categories /t/, /ɔ/, /n/, VL2, BFVL3, FVL3 and FSYL per junctural pair.

Junctural Pairs	/t/	/ɔ/	/n/	VL2	BFVL3	FVL3	FSYL
tɔ(n)ˈɐnɔ	t	ɔ	n	ɐ	n	ɔ	nɔ
tɔ(n)ˈɔmɔ	t	ɔ	n	ɔ	m	ɔ	mɔ
tɔ(n)ˈiki	t	ɔ	n	i	k	i	ki
tɔ(n)ˈipiɔ	t	ɔ	n	i	i	ɔ	iɔ
tɔ(n)ˈɛvrɔ	t	ɔ	n	ɛ ×	r	ɔ ×	rɔ

For the analysis of the data both by-speaker and by-junctural-pair analysis, each for both careful (CR) and casual (CS) speech, were applied.

3.1. Durational Modification

Figures 1 and 2 illustrate the durational differences of the test segmental material in careful and casual speech respectively. The categorization of A versus N represents the condition where the ambiguous nasal consonant /n/ attaches to the preceding ***article (A)*** (e.g. /tɔn#ˈɐnɔ/ or to the following ***noun (N)*** (ex. /tɔ#ˈnɐnɔ/). The F-ratios resulting from the by-speaker and the by-junctural pair analyses are symbolized as F_1 and F_2 respectively. For each segmental portion, the data were submitted to two sets of 2x2 repeated-measures ANOVAs, with Word Boundary and Speech Style as factors. In one set, the repeated factor was Speaker and in the other, Junctural Pair. Only the statistically interesting points are presented.

For segment /ɔ/, the interaction between Word Boundary and Speech Style was found to be strongly significant by junctural pairs, but not by speakers [F_1 (1, 4) = 3.055, p=0.155 and F_2 (1, 4) = 30.565, p=0.005]. Paired samples t-tests showed that /ɔ/ was significantly shorter (9 ms) in A than in N in CR (t=-3.052, df =4, p=0.038) and significantly longer (4 ms) in A than N in CS (t=3.292, df=4, p=0.03). Additionally, /ɔ/ was not significantly longer (2 ms) in CR than in CS in A (t=1.017, df= 4, p=0.366), but it was significantly longer by an average of 11 ms in CR than CS in N (t=10.690, df=4, p<0.0001). Figure 3 illustrates this interaction. Contrary to prediction, the tokens of segment /n/ showed no durational differences between A and N conditions [F_1 (1, 4) = 0.390, p=0.566 and F_2 (1, 4) = 0.331, p=0.596].

The tokens of VL2 were significantly longer (8ms) in A than in N by the analysis by-junctural pair [F_1 (1, 4) = 2.213, p=0.211 and F_2 (1, 4) = 17.355, p=0.014]. The difference between A and N was larger by an average of 8.6 ms in CR than CS, with 22.4 ms and 13.8 ms being the mean value of durational difference between A and N in each condition of speech rate respectively. The interaction between test factors was close to significance in the by-junctural pairs analysis [F_1 (1, 4) = 1.289, p=0.320 and F_2 (1, 4) = 4.782, p=0.094].

BFVL3 was significantly longer (7ms on average) in A than in N through the by-speaker analysis [F_1 (1, 4) = 13.865, p=0.02 and F_2 (1, 4) = 3.458, p=0.136]. Although A was longer than N in CR by 4.8 ms, and by 9.3 ms in CS, the interaction between Word Boundary and Speech Style was marginal only through the by-junctural pair analysis [F_1 (1, 4) = 1.346, p=0.311 and F_2 (1, 4) = 5.307, p=0.083].

The fact that FVL3 was longer by an average of 7.8 ms in A than N was almost significant by junctural pair [F_1 (1, 4) = 3.891, p=0.12 and F_2 (1, 4) = 7.496, p=0.052]. However, FSYL was significantly longer by an average of 15 ms in A than in N [F_1 (1, 4) = 11.571, p=0.027 and F_2 (1, 4) = 12.328, p=0.025].

3.2. Pitch Anchoring

Figures 4 and 5 show that there are not differences in pitch excursion between A and N (cf. [7]) and that the F0 minimum is probably placed somewhere within the segment /n/. For the examination of Ladd's and Schepman's [7] hypothesis, the durational distances of F0min from the beginning (F0minBeg) and end (F0minEnd) of the ambiguous consonant /n/ were calculated. Data-points were discarded when speakers showed glottal stops. However, comparison of means through paired samples t-tests using both analyses by speakers and by junctural pairs showed that F0min did not present different alignment in relation with the location of word boundary, opposing to predictions.

Figure 1: Durational differences values in milliseconds of /ɔ/, /n/, VL2, BFVL3, FVL3 and FSYL in A and N respectively in careful speech (CR).

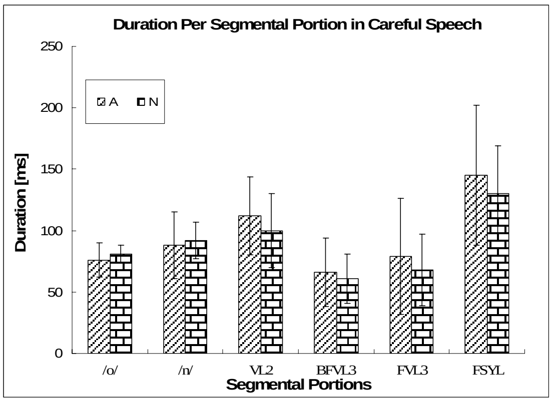


Figure 2: Durational differences values in milliseconds of /ɔ/, /n/, VL2, BFVL3, FVL3 and FSYL in A and N respectively in casual speech (CS).

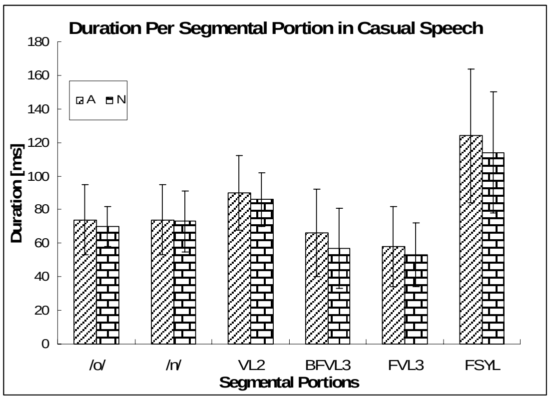


Figure 3: The interaction of the factors Word Boundary (A/N) and Speech Style (CR/CS) for /ɔ/.

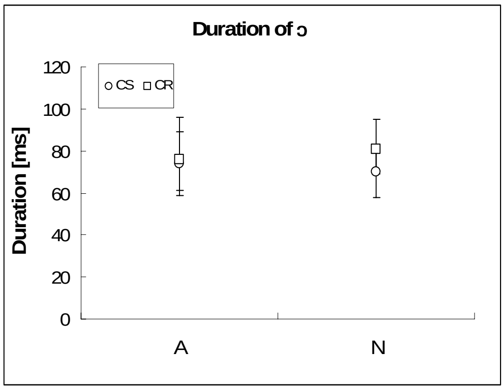


Figure 4: Pitch contour for careful speech.

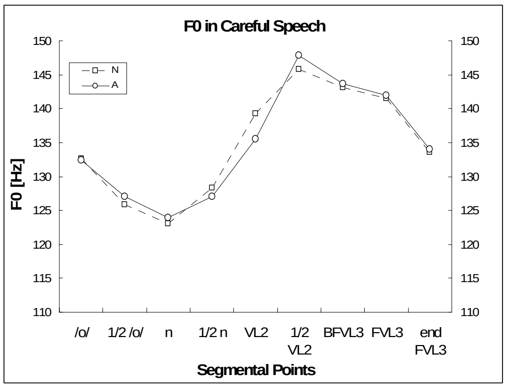
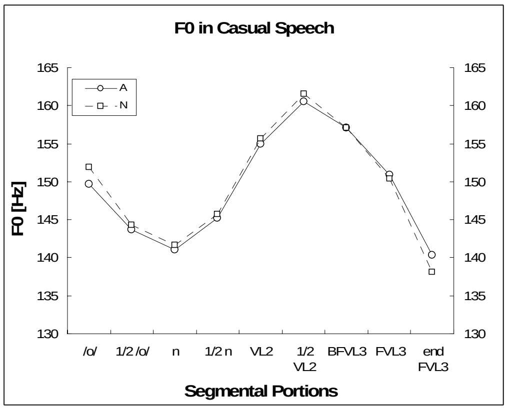


Figure 5: Pitch contour for casual speech.



4. DISCUSSION AND CONCLUSIONS

The locus of word boundary affects the durational properties of segments. Although word-final lengthening does not occur (no significant durational differences of the ambiguous consonant were observed across the two boundary conditions), word-initial lengthening is in play (the vowel following the pivotal nasal was longer when the latter preceded the boundary) (cf. Dutch and English [10], [11], [14]). The fact that word-final lengthening does not apply might be accounted for by Greek's phonological preference for open syllables, which causes resyllabification in connected speech. Thus, in both boundary conditions the nasal belongs to the post-boundary syllable. Resyllabification might also explain the finding that F0 minima were located at the beginning of the ambiguous consonant in both word boundary conditions, since segmental pitch anchoring functions as syllable marker [7].

A striking result is the consistent durational properties of the final syllable: Although speech style has an impact on the duration of FSYL (20 ms longer in CR than CS), the difference between A and N remains constant (15 ms) across both speech styles. The syllable following the ambiguous consonant is constantly accented in both junctural cases, implementing accentual lengthening (cf. [12], [13]). This type of lengthening can extend to at least one following syllable (cf. [1], [5]). Hence, the syllable preceding the final one might undergo double lengthening in the case where the dubious nasal is attached to the first syllable (accentual and word-initial lengthening), whereas its counterpart would undergo only accentual lengthening. This double lengthening of the post-boundary syllable might be manifested by transferring not only the accentual lengthening, but the word-initial lengthening too to the following syllable, since it would be difficult for speakers to pronounce an even longer initial syllable in a natural way.

Additionally, the double lengthening effect could be intensified by the (apparently optional) lengthening of the segment that precedes the final vowel. This indicates that some durational cues are speaker-specific, whereas others are intrinsic to words' properties (ex. lengthening the first vowel following the juncture -VL2- when the ambiguous consonant precedes the juncture).

The data analyzed in this study indicate that the speech signal provides invariant fine-grained phonetic markers that are systematically present in relation to contextual factors, supporting the view that different linguistic structures are distinguished at the word level [4]. Important questions for future research are to what extent these different marking mechanisms can be teased apart and whether they ultimately derive from production or perceptual factors. One line of useful evidence is likely to come from whether Greek speakers are able to use the details identified here to track speech in word-size chunks, and how they interact with the different kinds and degrees of additional information available to the perceiver of speech.

5. REFERENCES

[1] Arvaniti A, Ladd D. R. & Mennen I. (1998). 'Stability of Tonal Alignment: the case of Greek Prenuclear Accents'. *Journal of Phonetics*, January 1998.

[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] Cho T. & Johnson E.K. (2004). 'Acoustic correlates of phrase-internal lexical boundaries in Dutch' *Proceedings of the 8th International Conference on Spoken Language Processing (ICSLP 2004)* , vol.2, 1297-1300. Jeju, Korea. 2004.

[4] Hawkins, S. (2003). Roles and representations of systematic fine phonetic detail in speech understanding. *Journal of Phonetics* **31**, 373–405.

[5] Kastrinaki A. (2003). The temporal correlates of lexical and phrasal stress in Greek, exploring rhythmic stress: Durational patterns for the case of Greek words. MSc thesis, University of Edinburgh, 2003.

[6] Ladd. D.R. and Schepman, A. (1999). Segmental anchoring of tones as a word-boundary correlate in English. *Proceedings of the International Congress of Phonetic Sciences*, San Francisco, August 1999.

[7] Ladd D.R. (2004b). Segmental anchoring of pitch movements: autosegmental phonology or speech production? .In H. Quené and V. van Heuven (eds.), *On Speech and Language: Essays for Sieb B. Nooteboom*, Utrecht: LOT, pp. 123-131.

[8] Lehiste, I. (1960). An acoustic-phonetic study of internal open juncture. *Phonetica* **5**(supplement), 5-54.

[9] Prieto, P., (2005): "Word-edge tones in Catalan", *Italian Journal of Linguistics* **17** (special issue on Autosegmental-metrical approaches to intonation in Europe: tonal targets and anchors, ed. by Mariapaola D'Imperio).

[10] Quené H. (1992). Durational cues for word segmentation in Dutch. *Journal of Phonetics* **20**, 331-350.

[11] Quené H. (1993). Segment durations and accent as cues to word segmentation in Dutch. *Journal of Acoustical Society of America* **94**, 2027-2035.

[12] Turk A. E. & Sawusch J. R. (1997). The domain of accentual lengthening in American English. *Journal of Phonetics* **25**, 25-41.

[13] Turk A. E. & White L. (1999), Structural influences on accentual lengthening in English. *Journal of Phonetics* **27**, 171-206.

[14] Turk A.E. & Shattuck-Hufnagel S. (2000). Word-boundary-related duration patterns in English. *Journal of Phonetics* **28**. 397-440.