TONAL ALIGNMENT AND THE REPRESENTATION OF ACCENTUAL TARGETS

Amalia Arvaniti & D. Robert Ladd

Department of Linguistics, University of Edinburgh, U.K.

ABSTRACT

This paper examines the tonal composition and alignment of prenuclear accents in Greek. Our experimental results suggest that these accents, which show an initial dip and late peak alignment, are best described as L*+H since (a) their initial L tone is invariant in scaling and alignment and not affected by declination, and (b) the H tone is more variable in alignment and affected by the position of the accented syllable within the word.

1. INTRODUCTION

Prenuclear pitch accents in Greek show a slow rise that begins on the accented syllable and reaches its peak towards the end of this syllable, or on the following one (see Figure 1 for an example).

In the standard autosegmental-metrical framework of intonational analysis, as exemplified by [4], such accents in English are described as L*+H and said to be different from the two other types of rising accent, H* and L+H*, in both scaling and alignment: H* does not show the bitonal accents' initial dip, while the difference between L*+H and L+H* relates to the variable alignment of the "unstarred" tone (the trailing H and the leading L respectively). The evidence, however, for the three accent types has been disputed, e.g. by Ladd [3], who has argued that all rising accents are instances of H* with variable peak alignment, not distinct categories.

This issue is still unresolved in English. However, if it could be shown that another language, in the present case Greek, uses at least one of the bitonal accents, then the necessity of differentiating between single and bitonal rising accents in the universal inventory of accent types would have been demonstrated.

2. METHOD

In order to examine the tonal composition of Greek prenuclear accents, and in particular whether the L tone needs to be specified in their phonological representation, we devised two sets of sentences in which two accents (A1 and A2) within the same intonational phrase were separated by progressively more unaccented syllables. Table 1 gives the details of one of the sets of sentences. The second set was constructed along similar lines, but gave less useful results because speakers tended to divide some of the sentences into two prosodic phrases. This is discussed further below.

The hypothesis was as follows: if the F_0 dip observed at the beginning of A2 is due to declination between two H* accents, then it would become deeper as the number of unaccented syllables between A1 and A2 increased; if the F_0 dip is due to the specified L tone of a bitonal accent, the alignment and scaling of this tone would remain relatively stable regardless of the number of unaccented syllables between A1 and A2.

The test sentences were recorded in a sound treated booth in the Phonetics Laboratory of the University of Oxford. Three native speakers of standard Greek, naive as to the purposes of the experiment, recorded seven repetitions of the sentences of both sets, in random order. Durational and F_0 measurements of the four most natural (in the first author's judgement) repetitions were obtained from waveforms and F_0 traces respectively, using Waves+. The F_0 measurements were transformed into ERB scale (see [1], [2]).

Table 1: One of the two sets of test sentences. The syllables bearing A1 and A2 are underlined.

1. [tilefo'no sto 'mano ja to 'parti]
   "I'm calling Mano about the party."
2. [tilefo'nusa sto 'mano ja to 'parti]
   "I was calling...."
3. [tilefo'nusame sto 'mano ja to 'parti]
   "We were calling...."
4. [tilefo'nusame me to 'mano sti 'meri ja to 'parti]
   "We and Mano were calling Mary...."
5. [tilefo'nusame apo to 'mano sti 'meri ja to 'parti]
   "We were calling from Mano's to Mary...."

The data were analysed statistically using analyses of variance in which the independent variables were speaker and number of unaccented syllables between accents. Where necessary, the ANOVAs were followed by Scheffé tests; p-levels for these are presented below.

3. RESULTS AND DISCUSSION

The results on scaling show that for speakers CN and NP the value of the L tone of A2 (L2) is not affected by the number of unaccented syllables between accents (see Figure 2). ET's data, however, show a weak effect of this factor: the value of L2 is higher for sentence 1, with one unaccented syllable between accents, than for sentence 4, with four unaccented syllables ($p = 0.03$).

In CN and NP's data the number of unaccented syllables did not affect the F_0 difference between L2 and the H tone of A1 (H1). In the data from ET, however, this difference increases as unaccented syllables are added: it is smaller for sentence 1 than for sentences 3, 4 and 5 ($p = 0.018$, $p = 0.001$ and $p = 0.0001$ respectively); it is also smaller for sentence 2 than for sentence 5 ($p = 0.02$).

In terms of alignment, L2 is consistently aligned with the beginning of the stressed syllable bearing A2, in the data from all speakers (see Figure 3).

In contrast, the alignment of both H1 and H2 (the H tone of A2) exhibits greater inter- and intra-speaker variability as Figure 3 shows. In the case of H1 in particular, the peak is reached further from the beginning of the accented syllable as the number of unaccented syllables increases (for speakers ET and CN): H1 is closer to the beginning of the accented syllable in sentence 1 than in sentences 3, 4 and 5 (for ET, $p = 0.009$, $p = 0.007$ and $p = 0.001$ respectively; for CN, $p = 0.0001$ in all cases).

As these differences in alignment level off once the number of three unaccented syllables is reached, i.e. once the accent is placed on the antepenult (see Figure 4), the results suggest that the alignment of this H tone may depend on the position of the accented syllable relative to the right boundary of the accented word.

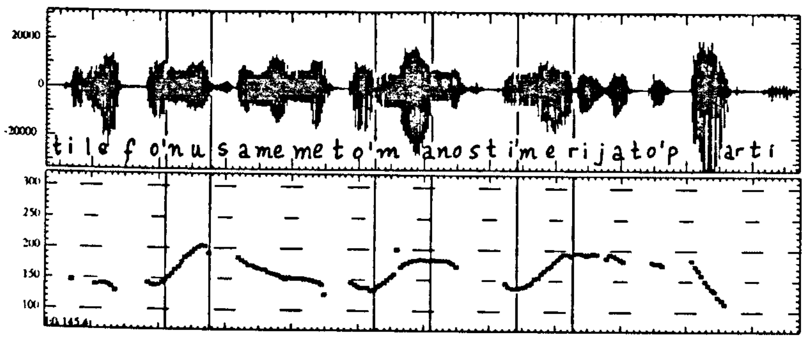

Figure 1: Waveform and F_0 contour of the test sentence [tilefo'nusame me to 'mano sti 'meri ja to 'parti]. The parts of the contour corresponding to prenuclear accents are between vertical lines.

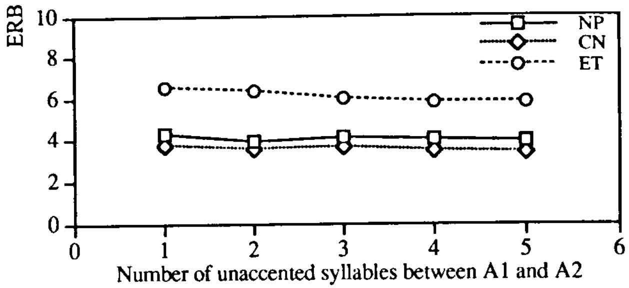

***Figure 2:** Scaling of L2 as a function of the number of unaccented syllables between accents, for each speaker separately. Mean values are in ERB scale (standard deviations are too low to be shown).*

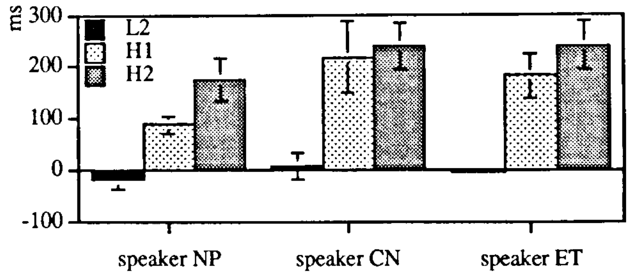

***Figure 3:** Alignment of L2, H1 and H2, i.e. distance of the tone from the beginning of the accented syllable (in ms), for each speaker separately; means and standard deviations are shown.*

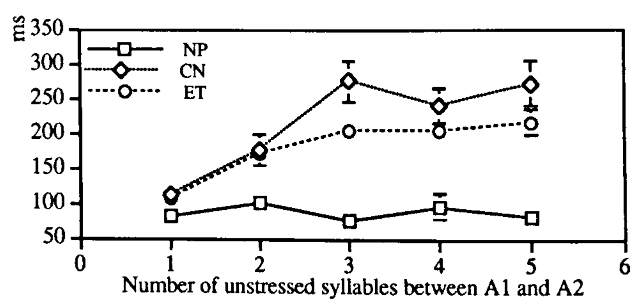

***Figure 4:** Alignment of H1, i.e. distance of the tone from the beginning of the accented syllable (in ms), as a function of the number of unaccented syllables between accents, for each speaker separately; means and standard deviations are shown.*

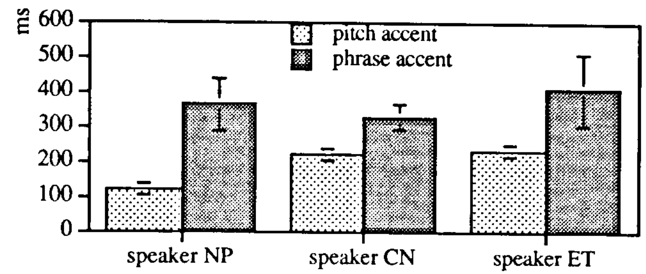

***Figure 5:** Mean distances between L2 and H2 (and standard deviations) in the second set of sentences, for each speaker separately; light grey bars show the distance when the L and H tones form a bitonal accent, and dark grey bars when they form the sequence L* H⁻ (see penultimate paragraph of text).*

Thus the data suggest that prenuclear pitch accents in Greek are best represented as L*+H, since the L tone (a) is clearly specified and not a result of declination and (b) shows more stable scaling and alignment than the H tone.

It might be argued that the results of the H tone alignment – viz. the fact that the H tone seems to align with the edge of the accented word – could justify considering it a type of phrase accent that demarcates the end of the word. However, the data from the *second* set of test sentences suggest that this is not an appropriate interpretation. The data from this set, as noted above, could not all be used in the main analysis, because in many cases speakers divided the sentences into two prosodic phrases. (In the cases where speakers did *not* divide the sentence into two phrases, the data from the second set agreed with the first set.) In autosegmental terms, when speakers divided the sentence into two prosodic phrases, they inserted a H^- phrase accent after the word which was intended to have A2, and replaced this accent with L*. In these cases, the alignment of the L and H is markedly different from the same tones in the L*+H accent ($F(1, 2) = 19.09$, $p = 0.04$, for the distance between L2 and H2 in the two configurations). As can be seen in Figure 5, when the two tones form a bitonal L*+H accent the distance between them is shorter and varies less; when the H tone is in fact a H^- then it is placed further away from the preceding L* accent, and the distance between the two tones is highly variable. In other words, the rise associated with prenuclear accents in Greek must be the trailing tone of a bitonal accent.

In conclusion, although further research on the alignment of the trailing tone of these accents and on the function and alignment of single H* accents is still necessary, the present results suggest that the prenuclear accents of Greek are best represented as bitonal L*+H accents.

REFERENCES

[1] Glasberg, B. R. & B. C. J. Moore (1990) Derivation of auditory filter shapes from notched-noise data. *Hearing Research* 47: 103-138.

[2] Hermes, D. & J. van Gestel (1991) The frequency scale of speech intonation. *Journal of the Acoustical Society of America* 90: 97-102.

[3] Ladd, D. R. (1983) Phonological features of intonational peaks. *Language* 59: 721-759.

[4] Pierrehumbert, J. (1980) *The phonology and phonetics of English intonation.* Ph.D. dissertation, MIT.