Journal of Phonetics (1998) **26,** 3–25

Stability of tonal alignment: the case of Greek prenuclear accents

Amalia Arvaniti*

University of Cyprus, Department of Foreign Languages and Literatures, P.O. Box 537, Nicosia 1678, Cyprus

D. Robert Ladd and Ineke Mennen

University of Edinburgh, Department of Linguistics, AFB, George Square, Edinburgh, EH8 9LL, U.K.

Received 5th July 1996, and in revised form 24th June 1997

Greek prenuclear accents show a sharp rise that starts near the onset of the accented syllable and peaks on the following unaccented syllable (if there is one). We have presented elsewhere evidence for analyzing these accents as consisting of a L(ow) and a H(igh) target. In a first experiment exploring the factors that affect the location of the H target, we discovered that in words with antepenultimate lexical stress the H is consistently aligned just after the onset of the first postaccentual vowel. In a second experiment we replicated this finding, showing that the alignment of the H is not affected by variations in the duration of the accented syllable. A third experiment showed that for some speakers the alignment of the H *may* be affected by "tonal crowding", when the accented syllable is close to the end of the word and/or close to the next accent. Overall, however, the results show that the L and H targets are independently aligned relative to the segmental string: the accentual rise is neither of fixed slope nor of fixed duration. This result, which replicates and extends earlier findings of Prieto, van Santen & Hirschberg (1995) for Mexican Spanish, is difficult to accommodate in a theory that views pitch movements as the primes of intonational structure.

1. General introduction

1.1. *Theoretical background*

Since the groundbreaking work of Bruce (1977) on Swedish, it has been known that at least some fundamental frequency (f_0) patterns in the European languages are characterized by the relative phonetic stability of certain *tonal targets*. Tonal targets are specific points in the f_0 contour that can be seen as the realization of a *tone* or some similar phonological object. Phonetically, such tonal targets can vary along two dimensions, the "horizontal" time dimension and the "vertical" f_0 dimension. The phonetic/phonological

* Correspondence to A. Arvaniti.

0095–4470/98/010003 + 23 $25.00/0/jp970063

properties corresponding to these two dimensions have come to be known as *alignment* and *scaling*, respectively.

Bruce's original discovery had to do with the phonetic basis of the Swedish word accent distinction.[1] He found that accented syllables are in every case accompanied by a fall from a local f_0 peak to a valley, and that the distinction between the two accent types is conveyed by the fact that the fall occurs just before the accented vowel for accent I and during the accented vowel for accent II. More specifically, he found that the local f_0 peak, whether it occurs before the accented syllable (accent I) or right at the beginning of the accented vowel (accent II), is aligned very precisely in time with the segmental material, and that its f_0 value is extremely constant across multiple repetitions of the same utterance by the same speaker. Bruce interpreted this to mean that, for the Swedish accentual distinction, "reaching a certain pitch level at a particular point in time is the important thing, not the movement (rise or fall) itself" (1977: 132).

On the basis of these findings, Bruce proposed that f_0 contours are most appropriately analyzed as a series of local maxima and minima — "turning points" — which he notated H(igh) and L(ow). In this view, rises and falls of pitch are essentially only transitions between one turning point and another. With this analysis, Bruce thus independently arrived at conclusions similar to those being developed within the autosegmental approach to intonation that was emerging in the United States at about the same time. Works such as those by Liberman (1975), Goldsmith (1976), and Leben (1976), for theoretical reasons largely unrelated to Bruce's concerns, were exploring the idea of analyzing the f_0 contour in English as a string of *level tones* like those used to describe languages with lexical level tones, such as Yoruba. Although phonological level tone cannot be equated directly with the quite specific notion of phonetic turning point, Bruce's approach and that of the autosegmentalists share an important assumption, namely that local f_0 movements are not primitives of the linguistic analysis, but are defined in terms of their beginning and ending points. This is the assumption that underlies much recent research on intonational phonology and phonetics (Pierrehumbert, 1980; Silverman *et al.*, 1992; Ladd, 1983, 1996).

Among the works based on this assumption are a number of studies of the scaling and alignment of putative tonal targets. Many of these studies have demonstrated the same kind of stability for f_0 targets in other languages as that found by Bruce for Swedish. This is particularly true of target scaling, which has been the topic of a number of experimental studies and quantitative models (e.g., Liberman & Pierrehumbert, 1984; Ladd, 1988; Shriberg, Ladd, Terken & Stoleke, 1996). Fewer studies have dealt with alignment than scaling, but those few that have done so provide evidence for the same kind of stability observed by Bruce. Two studies are worth summarizing, namely Silverman & Pierrehumbert (1990) and Prieto, van Santen & Hirschberg (1995).

Silverman & Pierrehumbert (1990) investigated the alignment of the prenuclear high (H*) accent in English under a variety of prosodic conditions (speech rate, the proximity of the following word boundary, and the number of unaccented syllables intervening between the test syllable and the following accent). They found that the accentual f_0 peak is preferentially aligned "past the end of the rhyme", i.e., into the following unaccented syllable (1990: 87), but they also found that this alignment is subject to

[1] Swedish, like several European languages, exhibits a lexically conditioned distinction between two types of "pitch obtrusion" (Bruce, 1977: 11) or accent, which associates with one of the word's stressed syllables. The two types of accent in Swedish have traditionally been described as accent I (or acute) and accent II (or grave) (for details see Bruce, 1977, and references therein).

systematic shifts as a function of the prosodic conditions mentioned above. For example, they found that the peak is aligned earlier if there is another accent on the immediately following syllable. They also suggest that this variability in peak alignment can best be described quantitatively if alignment is expressed as a proportion of the duration of the stressed vowel, rather than as an absolute measure.

Prieto *et al.* (1995) studied the alignment of H* accents in Mexican Spanish. Unlike Silverman & Pierrehumbert, who looked only at f_0 peaks, Prieto *et al.* measured the alignment of both the preceding valley and the peak of the accentual rises in question. They found that the valley is very consistently aligned just before the beginning of the onset consonant of the accented syllable. The alignment of the peak is more variable, being affected by a number of factors of the sort investigated by Silverman and Pierrehumbert (e.g., word boundary and stress clash), although Prieto *et al.* did not find any evidence for Silverman & Pierrehumbert's hypothesis that peaks are best expressed as a proportion of the accented syllable's rhyme length. Like Silverman & Pierrehumbert, they also found no support for what they call the "invariance hypothesis", i.e., the hypothesis that the slope and duration of the f_0 rise is constant. Instead, they conclude that "the absolute rise time (of an accent) is different in almost all the conditions (they) investigated. ... (W)hen peak location is retracted, absolute rise time shortens, and it is the velocity of the accent rise that accommodates to the time constraints imposed" (Prieto *et al.*, 1995: 449f).

The study reported here is intended to contribute to this line of research and to provide, among other things, results from a different language, Modern Greek. Our materials are very similar to Prieto *et al.*'s, though our study was designed prior to the publication of Prieto *et al.* (1995). In particular, we were originally motivated by rather specific questions of phonological representation arising within the autosegmental analysis of intonation, and by the assumption—based on Silverman & Pierrehumbert (1990)—that peak alignment would be variable and affected by various prosodic factors, such as the proximity of the word's boundary and of the following accent. We outline the phonological background briefly in Section 1.2, in order to set our original experimental approach in context.

1.2. *Modern Greek prenuclear accents*

Modern Greek prenuclear accents are characterized by a sharp rise, from an f_0 minimum at the onset of the accented syllable to a peak that is normally reached on the *following* unaccented syllable. We first examined these accents in Arvaniti & Ladd (1995) (henceforth AL95), where we looked at the beginning and ending points of the accentual rise in sentences containing two consecutive accents separated by one to five unaccented syllables. We found that the initial f_0 minimum is very stable in terms of both scaling and alignment: it occurs approximately 5 ms before the onset of the accented syllable, and its scaling is not affected by the number of unaccented syllables intervening between accents. The position of the peak, on the other hand, is variable: the larger the number of unaccented syllables following the accented one, the later the peak occurs.

In AL95 we interpreted these results on the scaling and alignment of the valley and the peak as evidence for two tonal targets, a L and a H. In particular, we suggested that the prenuclear accent is a bitonal accent L*+H, i.e., an accent with a stable low "starred tone" and a high "trailing tone". However, the results of AL95 also suggest an alternative autosegmental analysis of the observed rise, namely as a sequence of a L* accent followed by a H "phrase tone" aligned with the edge of the accented word. In particular, in AL95 we showed that the alignment of the peak was affected mainly when the accent was separated from the following accent by one, two, or three unaccented syllables. Since in those materials this interval corresponded to words with final, penultimate, and antepenultimate lexical stress, respectively (lexically stressed syllables being those with which pitch accents are associated), an alternative interpretation could be that there are indeed two tonal targets, a L and a H, which however do not form a bitonal accent: the accent is a L*, and the H is some kind of edge tone that marks the right boundary of the accented word. This analysis would be analogous to Pierrehumbert's analysis (1980) of the falling nuclear accent in English as a sequence of a H* accent and a L phrase tone said to be aligned with the edge of the accented word.

Yet there are difficulties with both analyses, as with any other possibilities suggested by autosegmental theory. The most important problem for any analysis is the fact that neither the L nor the H is actually aligned with the accented syllable: the L comes just before it, and the H comes after it. This makes it difficult to designate either tone as starred, since starred tones are supposed to align with the accented syllable (Pierrehumbert, 1980). Our first experiment was an attempt to decide between analyzing the Greek prenuclear accent as a L*+H bitonal pitch accent and analyzing it as a monotonal L* with a following H edge tone. In particular, we looked for factors that cause the variability in the alignment of the H in order to evaluate the different possible autosegmental interpretations of our previous results (AL95). As it turned out, we did not find the kind of variability we expected (for a discussion of the ramifications of this for the phonological question originally addressed in our research see Arvaniti, Ladd & Mennen, forthcoming). However, the kind of invariance that we did find has a number of implications for the issues investigated by Silverman & Pierrehumbert (1990) and by Prieto *et al.* (1995), and this is the focus of the present paper.

2. Experiment 1

2.1. *Introduction*

As we just noted, in Experiment 1 we were looking for factors that cause the variability of prenuclear peaks in order to decide between different interpretations of our AL95 results. Specifically, in Experiment 1 we wanted to test three hypotheses regarding the factors regulating the alignment of the H target.

Hypothesis A: The distance between the L and H targets is a relatively stable pitch excursion (once factors such as tonal pressure are removed). In autosegmental terms, this would be interpreted as further evidence in favor of our original analysis of the accent as L*+H (AL95), as it would be consistent with the H being the "trailing" tone of a bitonal accent—i.e., a tone which is said to be aligned at a fixed distance from its starred counterpart (e.g., Grice, 1995*a*). In this case, we would expect the alignment of the H not to be affected by the segmental composition of the postaccentual syllables (with which, as noted, the H normally coincides).

Hypothesis B: The H is a tone that demarcates the right edge of the word. In this case, we would expect the alignment of the H to be affected by the duration of the postaccentual syllables. That is, if the canonical alignment of the H is a fixed distance from the end of the word, then we would expect the H to coincide with different segments (possibly even syllables) depending on the segmental makeup of the postaccentual syllables.

Hypothesis C: The alignment of the H is regulated, at least in part, by the oncoming accent. As mentioned, in AL95 we show that the alignment of the H is affected by the number of postaccentual syllables. This may not be due to the word boundary but could be the result of pressure to realize the L target of the oncoming accent. If this hypothesis is correct, we would expect the H to be aligned earlier when the oncoming prenuclear accent is closer, for instance, when it is separated from the previous accent by shorter (rather than longer) syllables.

2.2. *Method*

2.2.1. *Speech materials*

In order to examine these hypotheses, a corpus of 22 sentences was designed, 6 relating to Hypothesis A and 16 relating to Hypotheses B and C. To limit the range of variability and thereby make for more stable and (we hoped) more interpretable results, the test words, i.e., the words bearing the prenuclear accent, were exclusively *proparoxytones* (words with lexical stress on the antepenultimate syllable). Thus, the effects we were interested in were those due to segmental duration differences rather than to differences in the number of unaccented syllables following the accent.

For Hypothesis A the test words were followed by a word that had one unaccented initial syllable. This means that in all cases there was a sequence of three unaccented syllables following the accented syllable of the test word. To test for segmental effects, this three-syllable sequence included only stops, fricatives or nasals; there were two test sentences for each of these conditions.

Hypotheses B and C were tested together but in such a way that the possible effects of one could be distinguished from those of the other. For Hypothesis B, in particular, we created materials in which the two postaccentual syllables of the test words were combinations of long and short syllables. "Long syllables" contained low or mid vowels ([a, e, o]) and, in some cases, consonant clusters as onsets, while "short syllables" contained short consonants, such as [ɾ], and high vowels ([i, u]). We hypothesized that if the H was aligned relative to the end of the word, then its distance from the preceding L target would be shorter if the postaccentual syllables were short.

In order to combine these materials with those appropriate for Hypothesis C, each test word used for Hypothesis B was incorporated into two sentences; in one of them it was followed by a word with a long initial unaccented syllable, in the other by a word with a short initial unaccented syllable. That is, the materials included altogether eight combinations of postaccentual syllable durations, shown in Table I. There were two sentences for each combination, yielding 16 test sentences in all (see Table II for sample sentences).

2.2.2. *Procedure*

The materials were recorded by five native speakers of Greek, three female (DA, SP and KA), and two male (IA, AH). (The results of a third male speaker had to be discarded, because he devoiced unaccented vowels to such an extent that it was not possible to obtain a sufficient number of acceptable repetitions for all the test sentences.)

TABLE I. Experiment 1: the combinations of short and long post-accentual syllables that were used to test Hypotheses B and C. # notes the location of the word boundary

	a: long syllable follows	b: short syllable follows
1.	short–short # long	short–short # short
2.	long–long # long	long–long # short
3.	short–long # long	short–long # short
4.	long–short # long	long–short # short

TABLE II. Experiment 1: sample sentences from the corpus, with the test words underlined. Sentences 1, 2, and 3 relate to Hypothesis A, and sentences 4 and 5 to Hypotheses B and C. Sentences 4 and 5 demonstrate how the combinations of long and short syllables were achieved by using the same test word ([a'pomakɾa] in 4a and 4b, and ['ðiɣlosi] in 5a and 5b) followed once by an unaccented long syllable ([vɾa] and [ksa] in 4a and 5a, respectively) and once by an unaccented short syllable ([vu] and [ði] in 4b and 5b, respectively)

1.	['teseɾa pa'ɾanoma no'mizmata 'vɾeθikan stin kato'çi tu] "Four illegal coins were found in his possession."
2.	[to ɾo'ðitiko to'pio 'ine apo ta oɾe'oteɾa tis e'laðas] "The landscape of Rhodes is among the most beautiful of Greece."
3.	[ti θi'mosofi ɟa'ɟa tus 'itan a'polafsi na tin a'kus] "Listening to their wise grandmother was a pleasure."
4.a	[ta a'pomakɾa vɾa'xaca 'isa pu ðia'kɾinodan 'mesa stin o'mixli] "The distant rocks could just be seen in the fog."
4.b	[sta a'pomakɾa vu'na me'notan mɲa tɾome'ɾi kate'ɟiða] "In the distant mountains raged a terrible storm."
5.a	[i 'ðiɣlosi ksa'ðeɾfi tu tu mi'luse si'neçia agli'ka ɟa na ton pi'ɾaksi] "His bilingual cousin continually spoke English to him to tease him."
5.b	[i'ðiɣlosi ði'motes a'petisan na ði'ðaskonde ce i 'ðio 'ɣloses sta sxo'lia tis 'polis] "The bilingual citizens demanded that both languages be taught in the city's schools."

The speakers were all in their twenties, and at the time of the recording had been in Edinburgh, as students at the University, for periods ranging from a few months to four years. All speakers were brought up in Athens and spoke Greek with a standard Athenian accent. None of the speakers had any known speech or hearing problems and they were all naïve as to the purpose of the experiment.

The recordings were made on digital audio tape (DAT) in the recording studio of the Department of Linguistics, University of Edinburgh. The speakers read each test sentence a total of eight times from a randomized list typed in Greek. Prior to the recording the speakers were instructed to read the sentences as naturally as possible and were given some time to look at the material. The recordings were monitored, and the speakers were asked to repeat any misread sentence.

Six out of the eight repetitions of each sentence were used for measurement. These six tokens were selected by the third author who chose those repetitions that sounded most fluent and natural. In particular, utterances were discarded if they contained a disfluency, or if the speaker had produced a phrase boundary immediately after the test word, or an otherwise different intonation contour from the one we had anticipated. Different contours were produced to a considerable extent, especially in the materials testing for Hypotheses B and C, so that for Hypothesis B it was not possible to find six acceptable repetitions of every sentence for all the speakers. For this reason, the data used for Hypothesis B include three, rather than six, repetitions of every sentence. However, we were able to include six repetitions from all speakers in the subset of sentences used to test Hypothesis C (for details see Section 2.3).

The materials were digitized at 16 kHz, with appropriate low-pass prefiltering, and were analyzed using ESPS Waves$^+$ on a SUN workstation. f_0 was measured from tracks obtained using the pitch-tracking facility of Waves$^+$, which took measurements of f_0 (in Hz) every 10 ms over a 49 ms cosine^4 window. The H, which was generally easy to locate, was measured at the highest f_0 point around the end of the accented syllable. No attempt was made to normalize for possible microprosodic effects; for instance, if the rise showed two successive peaks, because of a shallow trough caused by a nasal or a voiced fricative, we took the higher of the two peaks to be the location of the H target. The measurement of the L target and of the interval between the L and the H proved problematic, as in some test words the accented syllable began with a voiceless obstruent; in such cases it was not easy to decide whether the low f_0 observed just before the beginning of the accented syllable was the L target, a microprosodic effect, or a value higher than the intended target which happened to be aligned with the voiceless consonant and hence impossible to locate. Because of this problem, we decided that in all cases we would consider the onset of the accented syllable as equivalent to the L target and measure the distance of the H target from that point. This interval is referred to as "C0toH". This decision was justified by our previous results (AL95), which showed that the L target of prenuclear accents is aligned on average 5 ms before the onset of the accented syllable.

For reasons that will become apparent shortly (see Section 2.3), further durational measurements, made from waveforms in combination with spectrograms, were also obtained at a later stage. Standard criteria of segmentation were followed for these measurements (Peterson & Lehiste, 1960). The durational measurements included (i) the interval from the onset of the accented syllable to the onset of the first postaccentual vowel (C0toV1), (ii) the duration of the first postaccentual vowel (V1toC2), and (iii) the distance of the H target from the onset of the first postaccentual vowel (V1toH). Fig. 1 illustrates how the various measurements were taken.

To examine the effect of segmental make-up on the alignment of the H (Hypothesis A), we ran a two-way ANOVA in which the independent variables were SEGMENT TYPE and SPEAKER. The variable SPEAKER had five levels and the variable SEGMENT TYPE three (stops, fricatives, and nasals).

To examine the effect of postaccentual syllable length on the alignment of the H target (Hypothesis B), we ran an ANOVA, in which the independent variables were SPEAKER and PREBOUNDARY INTERVAL (i.e., the combination, in terms of length, of the two postaccentual syllables that belonged to the same word as the accent); this variable had four levels: long–long, short–short, long–short, and short–long.

Finally, in order to see whether the distance of the following accent would affect the alignment of H (Hypothesis C), we ran an ANOVA in which the independent variables were SPEAKER and POSTACCENTUAL INTERVAL (i.e., the combination, in terms of length, of the three syllables following the accented one); this variable had four levels; long 3 (all three syllables were long), short 3 (all three syllables were short), long 2 (two of the

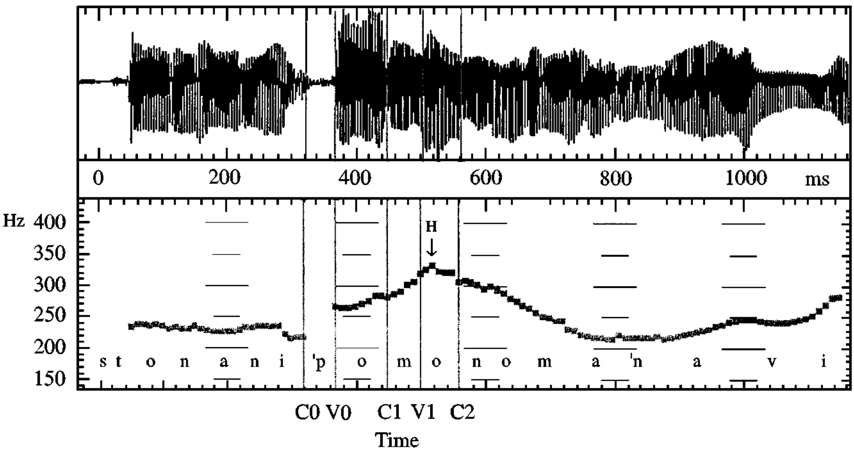

Figure 1. Waveform and f_0 trace for one of the sentences used in Experiment 1, showing the durational measurements obtained and the point at which the H target was measured.

syllables were long, the other short), and short 2 (two of the syllables were short, the other long).

In addition to the ANOVAs, if an analysis showed interaction between factors we ran *post hoc* Scheffé tests in order to determine the source of the interaction.

2.3. *Results*

Hypothesis A: The results show that both SEGMENT TYPE [F (2, 160) = 50.92, $p < 0.0001$] and SPEAKER [F (4, 160) = 9.45, $p < 0.0001$] affected the alignment of the H. However, the analysis also showed that there was interaction between the two factors [F(8, 160) = 3.66; $p < 0.006$]. Scheffé tests showed that the interaction was due to the fact that SEGMENT TYPE affected alignment somewhat differently for each speaker (see also Fig. 2). Specifically, speaker SP showed no effect of SEGMENT TYPE; for the other speakers the earliest alignment of the H target occurred when the postaccentual syllables contained nasals, and the latest when they contained fricatives [for IA, $p < 0.0001$; for KA, $p < 0.0002$; for AH, $p < 0.006$; for DA, $p < 0.0001$]. For these four speakers, stops were somewhere in between nasals and fricatives in terms of their effect on peak alignment, though for KA and AH neither of the comparisons (nasals *vs.* stops, and stops *vs.* fricatives) reached the 0.05 significance level; for IA peak alignment was later with stops than with nasals [$p < 0.0001$] but no different from fricatives, while for DA it was earlier with stops than with fricatives [$p < 0.0001$] but no different from nasals.

Overall the results were not as we expected since, according to Hypothesis A, there should be no effect of segmental composition on the alignment of the H and in particular on its distance from the L tone. The results, however, suggest that postaccentual segmental composition does affect the alignment of the H, with nasals resulting in earlier alignment of the H, fricatives resulting in later alignment, and stops grouping with either one or the other category, depending on the speaker.

Hypothesis B: The length combinations of the first two postaccentual syllables showed a significant effect of both SPEAKER [F (4, 214) = 8.927, $p < 0.0001$] and

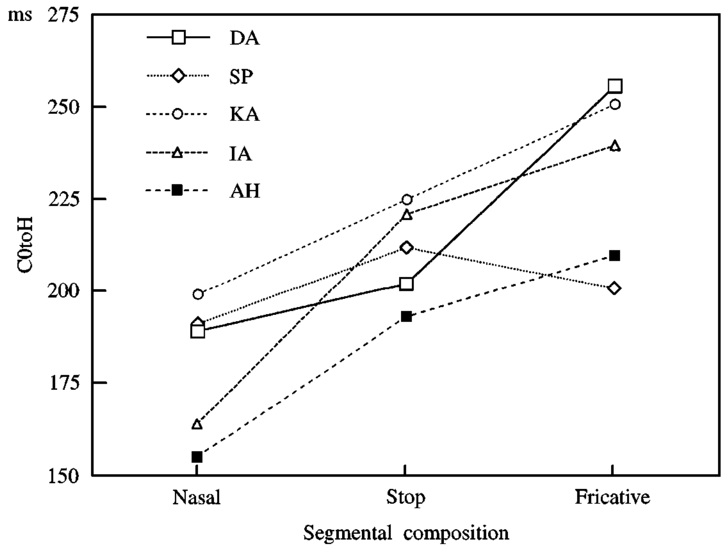

Figure 2. Experiment 1: duration means of the interval C0toH when the postaccentual syllables include nasals, stops or fricatives; data for each speaker separately.

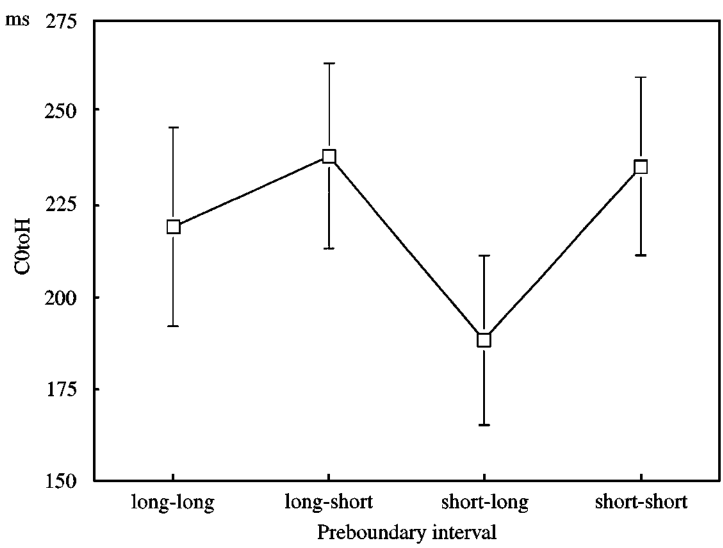

Figure 3. Experiment 1: duration means of the interval C0toH for each of the length combinations of the two postaccentual syllables that belong to the same word as the accent (preboundary interval); data from all speakers together.

PREBOUNDARY INTERVAL [F (3, 214) = 54.039, $p < 0.0001$], and no interaction between the two factors. Yet the results did not support our second hypothesis that the alignment of the H would occur later if the postaccentual syllables were longer. As shown in Fig. 3, by far the earliest alignment occurred with the combination short–long (the C0toH interval in short–long was significantly shorter than in all the other combinations; $p < 0.0001$ in all cases, according to Scheffé tests). Relatively early alignment of the H occurred also with the long–long combination (which we had expected to show the latest alignment of all); H alignment occurred earlier in the long–long combination than in both long–short [$p < 0.0004$] and short–short [$p < 0.003$]. The latest alignment occurred with the combination long–short (which we did not expect to be different from short–long) and, most surprisingly, with the combination short–short (there was no difference between these two). Again, our hypothesis was not supported, but the results clearly suggest that (a) the H is not aligned relative to the word boundary, and (b) segmental composition plays an important part in the alignment of the H tone.

Hypothesis C: The conclusions just stated were further supported by the analysis of the effect of the postaccentual interval length on the alignment of the H. For this particular analysis of the data we included only a subset of the 16 sentences of the original materials. The subset was selected so that all cells in the ANOVA were of equal size; i.e., we used all the sentences that had only long or only short postaccentual syllables, but not all the sentences that contained combinations of short and long syllables. In order to achieve equal cells in the analysis, we included only sentences for which we had a full set of six repetitions. The analysis is based on two test words for each of the four POSTACCENTUAL INTERVAL levels (long 3, long 2, short 2, short 3).

The results showed that both SPEAKER [$F(4, 219) = 17.22$, $p < 0.0001$] and POSTACCENTUAL INTERVAL [$F(3, 219) = 9.81$, $p < 0.0001$] affected alignment, and that there was interaction between the two [$F(12, 219) = 2.06$, $p < 0.02$]. Scheffé tests once more showed that different speakers adopted different strategies (see also Fig. 4). SP and AH showed no effect of POSTACCENTUAL INTERVAL. For DA and KA, alignment was earlier when the interval included three long syllables [for DA: long 3 *vs.* long 2, $p < 0.02$; long 3 *vs.* short 3, $p < 0.007$; long 3 *vs.* short 2, $p < 0.0006$; for KA: long 3 *vs.* long 2, $p < 0.01$; long 3 *vs.*

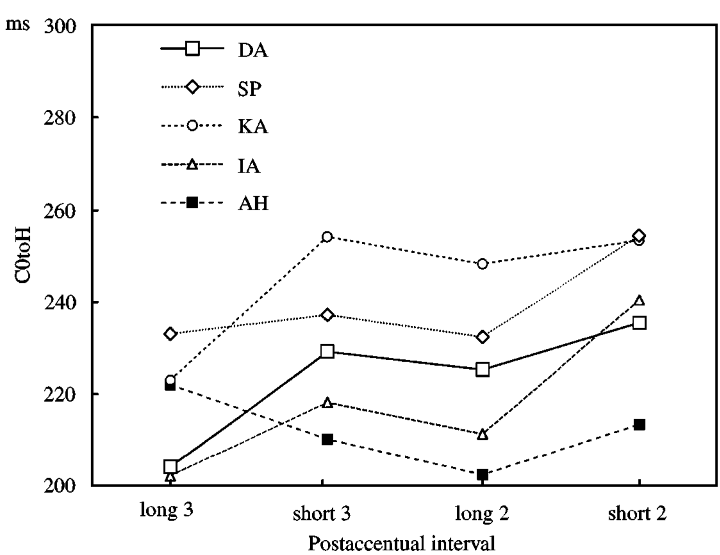

Figure 4. Experiment 1: duration means of the interval C0toH for each of the length combinations of all three postaccentual syllables (postaccentual interval); data for each speaker separately.

short 3, $p < 0.003$; long 3 *vs.* short 2, $p < 0.002$], but there were no differences among the other categories. Finally, for IA the earliest alignment occurred when there were two or three long syllables; in his data, both these categories showed earlier alignment of the H than short 2 [for long 3 *vs.* short 2, $p < 0.0005$; for long 2 *vs.* short 2, $p < 0.01$].

These results do not support the hypothesis that the H aligns earlier because of pressure from the oncoming accent, i.e., that it tries not to be too close to the L target of that accent, in that they do not show H to be aligned later if the interval between one accent and the next is longer. As with Hypothesis B, we found that the alignment either remained unaffected (for a minority of speakers), or was affected in ways that ran completely against our predictions.

Overall then, our results failed to support the more general hypothesis that the alignment of the H is somehow regulated by length differences in the postaccentual syllables. Instead what we found were large, and at first glance largely unexplainable, effects of segmental composition on the alignment of the H. However, by marking the segment with which the H co-occurred, we observed that in the majority of the data this segment was the first postaccentual vowel. In fact, further measurements showed that the H occurs on average 10.6 ms (SD = 14.1) after the beginning of this vowel, irrespective of the length of the accented syllable, or that of the following unaccented syllables. This was rather different from what we had expected: first, the H did not seem to be a fixed distance from the L tone, as one would expect if the H were the trailing tone of a $L^* + H$ accent, or if the accent consisted of a fixed pitch excursion; second, the H did not seem to be aligned relative to the actual accented syllable but was affected by the length of this syllable and the length of the postaccentual consonant.

In order to support the new hypothesis that emerged from the data, namely, that the H is aligned a fixed distance from the beginning of the first postaccentual vowel, we ran two correlations in which we included all the valid data relating to Hypotheses B and C. We reasoned that if our new hypothesis is correct then (a) the interval C0toH would correlate highly with C0toV1, and (b) the interval V1toH would show no correlation with the duration of the first postaccentual vowel (V1toC2).

Our new hypothesis was indeed confirmed by the data. As can be seen in Fig. 5 the location of the H is highly correlated with the interval stretching from the onset of the accented syllable to the onset of the first postaccentual vowel [$R^2 = 0.806$], but is not at all related to the duration of that vowel [$R^2 = 0.003$].

Not only do our correlation data confirm our new hypothesis, they also help us understand the puzzling results we obtained from the original analyses. First, it is not surprising that postaccentual nasals resulted in an earlier placement of the H, while postaccentual fricatives resulted in a later placement: if the alignment of the H is related to the interval from the onset of the accented syllable to the onset of the postaccentual vowel, it is natural that the longer fricatives will delay the occurrence of the H tone, while the shorter nasals will bring it forward (on durational differences between nasals and fricatives, see Carlson & Granström, 1986; Crystal & House, 1988; Fletcher & McVeigh, 1993).[2] This effect must have been strengthened by the fact that in one of the two test

[2] Strictly speaking, Carlson & Granström, Crystal & House, and Fletcher & McVeigh all report longer durations for voiceless fricatives than for nasals, but similar durations for nasals and voiced fricatives. Our corpus included both voiced and voiceless fricatives in equal numbers, but inspection of our data reveals that Greek voiced fricatives, for which there are no quantitative data available, are longer and articulatorily more "robust" than their English and Swedish equivalents.

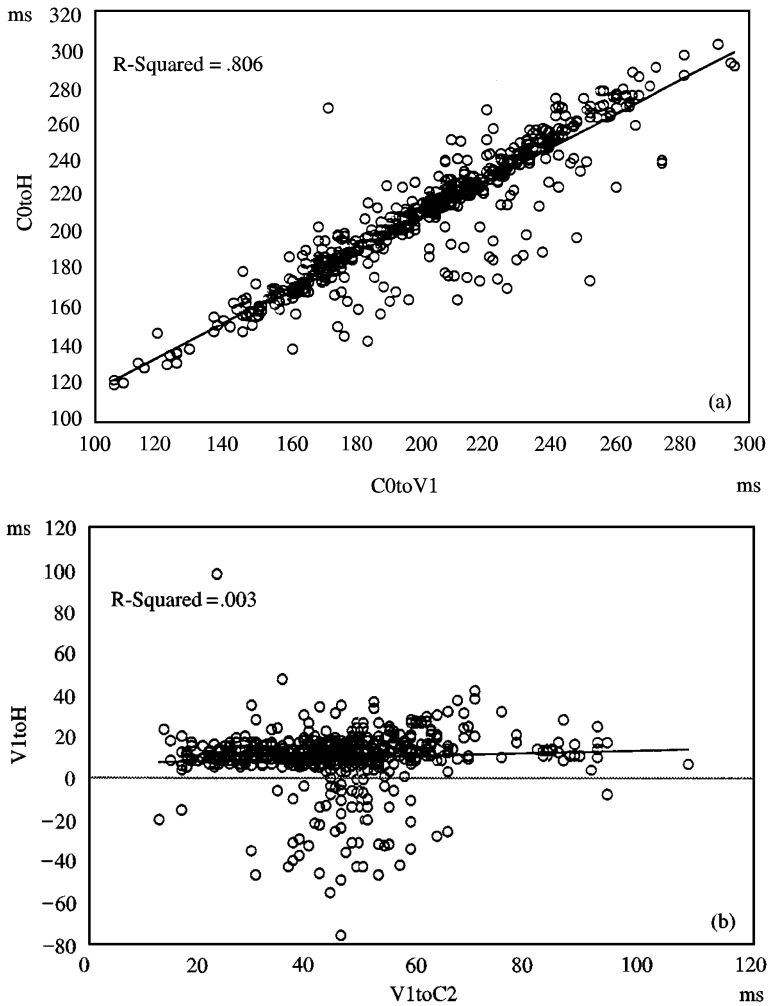

Figure 5. Experiment 1: (a) the interval C0toH as a function of the combined duration of the accented syllable and the postaccentual consonant; (b) the interval V1toH as a function of the duration of the postaccentual vowel; data from all speakers together.

words with postaccentual nasals the accented syllable is [ɾa], i.e., involves a flap, the shortest consonant of Greek, with an average duration of approximately 46 ms in accented syllables (Arvaniti, 1987).

Our new hypothesis also explains why the combination of a short and a long postaccentual syllable yielded much earlier alignment of the peak than the other combinations, while the combination of two long postaccentual syllables did not result in later alignment of the H target as expected. Examination of the materials reveals that the combination short–long included the words [ef'xaɾistos] "pleasant" and [po'litimo] "valuable". In [ef'xaɾistos] the first postaccentual consonant is again a flap. Similarly, the accented syllable of [po'litimo] contains the shortest Greek vowel, [i], with an average duration between 65 and 80 ms in stressed syllables (Fourakis, 1986*a*, *b*; Botinis, Fourakis & Katsaiti, 1995). In other words, in both cases the C0toV1 interval must have been short compared with the same interval in the long–short combination (to which we had expected it to be similar): there the words were [ˈðiɣlosi] "bilingual" and [akaˈtastati] "untidy". In both cases, the postaccentual syllable started with a cluster which, for most speakers, would considerably delay the occurrence of the H. Finally, the combination short–short showed a later alignment of the H than the combination long–long. Again, the reason for this surprising difference becomes evident once the postaccentual segmental composition of the test words in each category is considered. The combination short–short included the words [fiˈlomusi] "music lover" and [vaθiˈkocini] "dark red", both of which have [o], the second longest Greek vowel (Fourakis, 1986*a*, *b*; Botinis *et al*., 1995) as their accented vowel; in contrast, the combination long–long included the words [aˈpomakɾa] "distant" and [kseˈkuɾðisto] "unwound", with [o] and [u] as the accented vowels. While durational differences in the postaccentual consonants are probably evened out in these four words, [u] is the second shortest vowel of Greek (Fourakis, 1986*a*, *b*; Botinis *et al*., 1995), and thus brings forward the alignment of the H in the long–long combination. The results for Hypothesis C are very similar to the above and can be explained in the same way.

2.4. *Discussion*

In Experiment 1 we tested three hypotheses regarding the alignment of the H target of prenuclear accents in Greek. In a previous experiment (AL95) this target had exhibited high variability in terms of alignment and provided evidence that alignment may be related either to the position of the right boundary of the word carrying the accent, or to the position of the following accent. However, in the earlier study the variability in the alignment correlated with gross differences of duration based on different numbers of syllables between the accent of the test word and the following one. In the present study, we held the number of syllables constant and varied the duration by manipulating segmental composition. Under these conditions, none of the hypotheses we tested was supported by our data. Instead what we found was that, when the position of the accent in the word and the number of following unaccented syllables are held constant, the alignment of the H is unaffected by small differences in duration: the H is aligned just after the onset of the first postaccentual vowel. In other words, the data showed that in the case of Greek prenuclear peaks we are not dealing with a H trailing tone, in the "orthodox" sense of a tone aligned a fixed distance from the starred tone of a bitonal accent (Pierrehumbert, 1980; Grice, 1995*a*), i.e., the data do not support our initial analysis of the accent as L*+H; nor do they support our alternative analysis of the H as an edge tone aligned a fixed distance before the following word edge or before the following accent. Finally, the data show that we are not dealing with a pitch excursion of fixed slope or duration.

3. Experiment 2

3.1. *Method*

Experiment 2 was designed to replicate the unexpected result of Experiment 1, namely, that the H target of Greek prenuclear accents is aligned a fixed distance from the beginning of the first postaccentual vowel. In addition, in Experiment 2 we measured the f_0 interval between the L and H targets. If our hypothesis is correct and we are indeed dealing with two independently aligned targets, then the f_0 difference should not correlate with the distance in time between the two targets. By contrast, if accentual pitch excursions are assumed to be specified in terms of a certain more or less fixed slope, we should expect a strong correlation between the duration of the rise and the range of frequencies it spans.

3.1.1. *Speech materials*

A corpus of 25 different sentences was designed, in which the duration of the accented syllable of the test words was highly variable, ranging from rather long (e.g., [ðja] with a low vowel and an onset consonant cluster, or [lev] with a mid-vowel and a coda consonant), to rather short (e.g., [ɾi] with a high vowel and a short flap as the onset consonant).[3] The target accent was always on the antepenult of the test word and was followed by two to five unaccented syllables. Care was taken to choose test words that included only sonorants in the relevant syllables, in order to avoid the problems with the target alignment measurement we had encountered in Experiment 1. Examples of the sentences included in this corpus are given in Table III.

3.1.2. *Procedure*

The speakers who participated in Experiment 1 took part in Experiment 2 as well. Each test sentence was read twice from a randomized list typed in Greek. Practice sentences were added at the beginning (seven fillers) and the end (five fillers) of the list, to avoid discourse-initial and final f_0 effects. The third author, following the same criteria as in Experiment 1, selected for measurement the more natural of the two repetitions of each sentence.

Digitization and measurements were done as for Experiment 1. The only difference was that in Experiment 2 we were able to locate the L target with ease (since the materials included only sonorants and voiced fricatives), so we did not have to substitute the onset of the accented syllable for the actual L target as we had done in Experiment 1. For the measurement of the L, we followed the same criteria we used in Experiment 1 for the location of the H: i.e., we took the target to be the absolute f_0 minimum located in the vicinity of the accented syllable's onset; no attempt was made to compensate for any segmental effects on f_0. The following duration measurements were obtained: (i) the distance between the L and H targets (LtoH)—which corresponds to C0toH in Experiment 1; (ii) the distance between the onset of the accented syllable and the onset of the first postaccentual vowel (C0toV1); (iii) the distance of the H target from the onset of the first postaccentual vowel (V1toH); and (iv) the duration of the first postaccentual vowel (V1toC2).

[3] We excluded tautosyllabic clusters in the postaccentual syllable, as Experiment 1 had alerted us to the fact that speakers adopt two different strategies regarding the alignment of the H in cases such as [a'naɣlifa] "relief": some speakers treat the second consonant as part of the vocalic element of the syllable, and consequently align the H with it, while others treat it as part of the consonantal element and align the H with the vowel. (To a large extent, this is also the reason why Fig. 5(b) shows some negative values for the interval V1toH.) This problem did not occur in words in which the two consonants were heterosyllabic, such as [pa'remvasi] "intervention". For this reason, we did include heterosyllabic clusters in the postaccentual syllable. We also used words with tautosyllabic clusters in the onset of the accented syllable, since these did not seem to cause similar problems with the alignment of the L tone.

TABLE III. Experiment 2: sample test sentences; the test words are underlined

1.	['ama 'vjenume 'ekso to 'vɾaði ta pe'ðja ta kɾa'tai i mi'teɾa mu] "When we go out in the evening, my mother looks after the kids."
2.	[ta ɣliko'lemona pu a'ɣoɾases ðen 'itan 'fɾeska] "The limes you bought were not fresh."
3.	[i pa'ɾemvasi tu ipuɾ'ɣu ðen 'efeɾe to poθi'to apo'telezma] "The minister's intervention did not have the desired result."

3.2. *Results and discussion*

The results clearly replicated those of Experiment 1, although they showed greater variability due to the nature of the materials (different sentences, as opposed to repetitions of the same sentence): the H target was aligned on average 17 ms (SD = 31.9) after the onset of the first postaccentual vowel.

As can be seen in Fig. 6, there was a significant correlation between the distance from the L to the H target and the interval C0toV1 [$R^2 = 0.453$]. In contrast, no correlation was found between the distance of the H from the onset of the unaccented vowel and the duration of that vowel [$R^2 = 0.007$].

The results of the scaling difference between the two tones support, to a large extent, our hypothesis: for four of the speakers they show that indeed there was no correlation between the f_0 interval between the H and the L and the duration of that interval (for DA, $R^2 = 0.161$; for SP, $R^2 = 0.099$; for KA, $R^2 = 0.048$; for AH, $R^2 = 0.048$); for the fifth speaker, IA, however, there is a weak correlation between these intervals ($R^2 = 0.205$).

The results confirm the hypothesis we developed as a result of Experiment 1, and strongly support the idea that the distance between the two turning points of the accent is not fixed. Although the distance between the L and the H is variable, both are aligned with particular points in the segmental string. Moreover, there appears to be no systematic relationship between the duration of the f_0 excursion and the amount of pitch change; it would appear that the L and the H are independently scaled.

4. Experiment 3

4.1. *Introduction*

The results of Experiments 1 and 2 together showed very clearly that for words with antepenultimate stress the alignment of the H is invariant. Experiment 3 was designed in order to explore the extent to which regularities comparable to those we found for proparoxytones could be observed with paroxytones and oxytones (words with penultimate and final lexical stress). We believed that a possible effect of the word boundary on the alignment of the H could not be rejected outright: our results from Experiments 1 and 2 had clearly shown that the H is not sensitive to various types of durational differences in the postaccentual syllables, hence it could not be an edge tone; however, the fact that it was aligned relative to a particular syllable also meant that its alignment is probably phonologically, rather than phonetically, conditioned. If so, then peak placement in

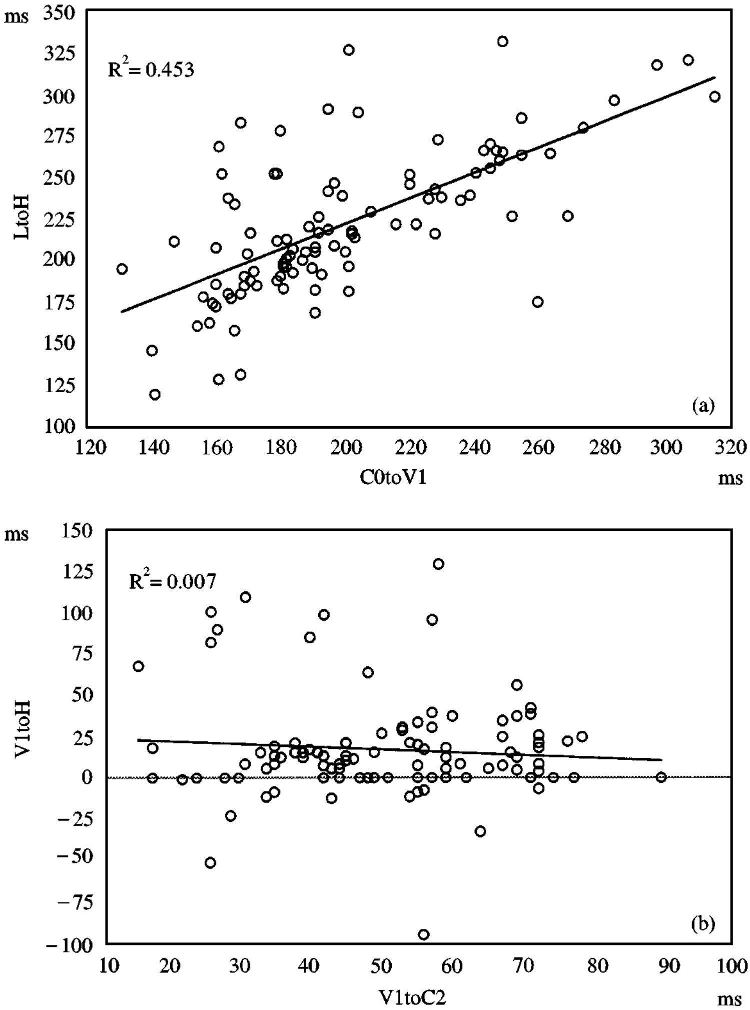

Figure 6. Experiment 2: (a) the interval LtoH (in ms) as a function of the combined duration of the accented syllable and the postaccentual consonant; (b) the interval V1toH as a function of the duration of the postaccentual vowel; data from all speakers together.

oxytones might be a problem: if the H is meant to be canonically aligned with the first postaccentual syllable, then speakers would have to choose between aligning the H earlier and keeping it within the boundaries of the accented word, and aligning it canonically but outside the accented word.

In addition, we wanted to test the effect that pressure from an oncoming accent would have on the alignment of the H, and in particular to see whether *tonal crowding* would result in regular temporal adjustments as reported in Silverman & Pierrehumbert (1990), Grice (1995*b*), and Prieto *et al.* (1995). Our prediction was that in such cases, both the number of postaccentual syllables and the position of the accent itself relative to the word boundary would affect the position of the H tone. In particular, we expected that the presence of accents on adjacent syllables would create tonal crowding which, in turn, could result in earlier alignment of the H tone (earlier peak placement can be observed in some of the clash data involving adjacent accented—as opposed to stressed but unaccented—syllables, reported by Arvaniti, 1994).

We also wanted to see whether the speakers would attempt to cope with the pressure caused by tonal crowding using strategies such as undershooting one of the targets, which are not among those suggested by Silverman & Pierrehumbert or by Grice. In particular, we expected that the scaling of the H could be affected as the number of unaccented syllables decreased, and that there could be a trade-off between scaling and alignment: under pressure the target would either not be reached (i.e., the observed f_0 maximum would be lower than under pressure-free conditions) or would be brought forward (i.e., closer to the L tone). Some evidence that undershooting is a possibility comes from AL95: here we found that for one of our speakers, ET, the L was scaled higher when there was only one unaccented syllable between accents, i.e., the L target was undershot, but not altogether missing from the contour. This also suggests that tonal crowding may result not only when the two accents are clashing, but even when there is one unaccented syllable between them. The possibility that one intervening syllable is not sufficient to alleviate the pressure caused by tonal crowding is consistent with Mirambel's (1959) observation that the rhythm of Greek is not based on alternation between stressed and unstressed syllables, but rather on a pattern involving one stressed syllable followed by two unstressed ones. This idea is supported by quantitative data reported by Arvaniti (1994).

4.2. *Method*

4.2.1. *Speech materials*

Given the consistent results of Experiments 1 and 2, we decided to use proparoxytones as controls and to compare these to paroxytones, and oxytones, while at the same time manipulating the number of syllables intervening between the test accent and the following accent.

We designed four sets of sentences, with six sentences each. The test word in each of these sentences was either an oxytone, paroxytone, or proparoxytone, and the number of following unaccented syllables ranged from none to two. Obviously, the position of these unaccented syllables relative to the word boundaries varied: for instance, in proparoxytones both syllables belonged to the test word, while in oxytones they could only belong to the following word. The set of conditions we tested is shown in Table IV. In order to avoid alignment differences due to segmental make-up, we designed the materials in such a way that the accented syllable of the test word and the following unaccented syllable remained constant within a set. In all cases, these two syllables were simple CV syllables, and the consonants were sonorants or voiced fricatives (see Table V).

4.2.2. *Procedure*

The materials were recorded by three speakers who had taken part in the previous experiments (KA, AH, and DA), following the same procedure as before. The sentences were read seven times, and our aim was to use for the analysis each speaker's "best" six repetitions ("best" being defined following the same criteria as in the previous

TABLE IV. Experiment 3: the combinations of oxytones, paroxytones and proparoxytones with the number of postaccentual syllables used in the experiment. # notes the location of the word boundary; σ stands for syllable. The boxes enclose the conditions for which two different analyses of variance were run (for details see text)

	A. oxytone	B. paroxytone	C. proparoxytone
(0) No postaccentual syllable	σ́#σ́		
(1) One postaccentual syllable	σ́#σσ́	σ́σ#σ́	
(2) Two postaccentual syllables	σ́#σσσ́	σ́σ#σσ́	σ́σσ#σ́

TABLE V. Experiment 3: one of the four sets of the test sentences. The words containing the sequence of accents under investigation are underlined; the capitals in front of each sentence refer to accent placement in the first word (A = oxytone, B = paroxytone, C = proparoxytone), and the digits to the number of following unaccented syllables (see also Table IV)

A0.	[tin <u>o'ðo'ɾijenas</u> tis pa'ʎas lefko'sias tin 'kseɾo a'pekso canakato'ta] "I know <u>Rigenas Street</u> in Old Nicosia inside out."
A1.	[me sinxo'ɾite a'la tin <u>o'ðo ɾi'jilis</u> ðen tin 'kseɾo] "I'm sorry, but I don't know [where] <u>Rigilis Street</u> [is]."
A2.	[pa'ʎa 'emenan stin <u>o'ðo ɾina'nias</u> sto 'cendɾo ala 'toɾa pça 'menum sta 'voɾia pɾo'astia] "In the old days they lived in <u>Rinanias Street</u> in the centre, but now they live in the north suburbs."
B1.	[apo 'tote pu i <u>maɾka'ðoɾi 'leɾosan</u> ton kana'pe i ma'ma mu apa'ɣoɾefse na zoɣɾa'fizo sto sa'loni] "Ever since <u>the markers stained</u> the sofa, mum has forbidden me to paint in the sitting room."
B2.	[epi'ði a'fti i <u>maɾka'ðoɾi le'ɾonum</u> se paɾaka'lo na 'ise pɾosekti'kos 'otan tus xɾisimopi'is] "As these <u>markers stain</u>, please be careful when you use them."
C2.	['exondas ta psi'la ɟa to <u>filo'ðoɾima 'etima</u> 'fonakse to ɡaɾ'soni ɟa na pli'ɾosi] "Having the change ready for the tip he called the waiter to pay."

experiments). However, this was not possible in most cases, mainly because the speakers often did not produce the type of contour we had anticipated, i.e., one in which both the test word and the following word were accented with the type of prenuclear accent under investigation. In many cases the speakers inserted an intonational phrase boundary after the test word, in others they deaccented it, and in still others they used a "hat pattern" (i.e., f_0 did not fall for a L target after the accent of the test word). It is possible that the larger variability in intonational patterns observed in this experiment is itself an effect of the tonal crowding we were investigating, i.e., the contours used were in effect strategies for coping with having too many accents too close together. Because of the observed variability, we had to discard a large part of the corpus and use six repetitions of each of the two sets of sentences (out of the four sets originally recorded) from each speaker.

Measurements and analyses were done as in the previous experiments. In particular, we again measured the distance between the L and H tones, and the f_0 difference of the L and H targets in Hz, using the same criteria as before.

For the statistical analysis, we ran separate ANOVAs for each speaker, since the three speakers did not produce the same full sets of test sentences. We performed two types of analysis. First, we ran two-way ANOVAs on the data that included oxytones and paroxytones with one and two following unaccented syllables.[4] For these ANOVAs the independent variables were STRESS TYPE, which had two levels (oxytone and paroxytone), and NUMBER OF UNACCENTED SYLLABLES, which also had two levels (one and two). In addition, we ran ANOVAs on the materials that included oxytones, paroxytones, and proparoxytones with two unaccented syllables following. For these analyses the independent variable was STRESS TYPE (in this case, a variable with three levels, oxytone, paroxytone, and proparoxytone). The dependent variables were the LtoH interval and the value of the H target in Hz.

4.3. *Results and discussion*

For speakers AH and KA neither NUMBER OF UNACCENTED SYLLABLES nor STRESS TYPE affected the placement of the H target. For speaker DA, however, NUMBER OF UNACCENTED SYLLABLES did affect H alignment: specifically, in her case, the H target was aligned earlier when there was one unaccented syllable between accents than when there were two [$F(1, 43) = 8.812$, $p < 0.0049$]. This result was the same for both oxytones and paroxytones (i.e., STRESS TYPE was not significant).

Similarly, the alignment of the H was not affected by STRESS TYPE when only words in which the postaccentual interval included two syllables were compared. For speakers KA and DA, alignment was stable in all three stress types (oxytones, paroxytones and proparoxytones). The data of AH, on the other hand, show an effect of the position of the accented syllable in the test word: in his case, the LtoH interval was much longer in proparoxytones than in oxytones and paroxytones (for proparoxytones *vs.* paroxytones, $p < 0.008$; for proparoxytones *vs.* oxytones, $p < 0.009$, according to Scheffé tests).

In short, the data suggest that the alignment of the H is largely unaffected by the position of the accent in the word, and quite possibly by the number of unaccented syllables, if there is at least one unaccented syllable. However, the data are not consistent for all the speakers. For AH the position of the accent seems to be important; it appears that for him the canonical shape of the accent requires not only two unaccented syllables, but two syllables that belong to the accented word. If these conditions are not met, the H is brought forward, irrespective of the number of postaccentual syllables. For DA, on the other hand, the position of the accent within the word is not important, but having only one postaccentual syllable constitutes enough pressure for her to bring the alignment of the H target forward. This result is reminiscent of the data reported for ET, the speaker in AL95 mentioned above. The results of DA and ET are not the same (ET seemed to "undershoot" the scaling of the following L rather than the alignment of the preceding H), but considered together they suggest that for some speakers having only one syllable intervening between accents constitutes tonal crowding which they try to alleviate using various strategies. Thus, there seem to be two types of speakers: those for whom the number of following unaccented syllables is the most important factor regulating the alignment of the H, and those for whom the most important factor is the

[4] In the analyses presented here we have not included the materials that had no unaccented syllables between accents (stress clashes). These results exhibited so much variation both within and across speakers, that we are unable to draw any useful generalizations at present.

accented word's right edge and its distance from the accented syllable. However, in all cases it must be kept in mind that we are dealing with "phonological distance" expressed in number of syllables, and not "phonetic distance" expressed in ms.

The data on the LtoH difference in Hz largely confirm those of alignment. Specifically, for the two speakers AH and KA for whom neither STRESS TYPE nor NUMBER OF UNACCENTED SYLLABLES affected the alignment of the H, the scaling of the H was also unaffected by these factors. In contrast, in the data of DA, who showed earlier alignment of the H when there was only one postaccentual syllable, the difference between the L and H targets was also smaller under this condition [$F(1,43) = 6.2$, $p < 0.01$], providing further evidence that for this speaker one unaccented syllable between accents constitutes tonal crowding.

The results of the f_0 difference between L and H from all three stress types are not so readily interpretable. The data were extremely variable and appeared to be affected by STRESS TYPE for all the speakers, but not in exactly the same way [for AH: $F(2,33) = 6.35$, $p < 0.004$; for DA: $F(2,33) = 3.91$, $p < 0.029$; for KA: $F(2,33) = 4.96$, $p < 0.01$]. For AH and KA, proparoxytones showed a larger f_0 difference between targets than paroxytones [for AH: $p < 0.004$; for KA: $p < 0.01$], but oxytones did not (i.e., paroxytones and oxytones behaved in the same way); for DA, paroxytones showed a smaller difference than oxytones [$p < 0.045$], but not proparoxytones.

These results cannot be easily explained but, as already mentioned, they may be the result of different utterances having somewhat different contours; the differences may have been slight, or the contours may have been variants, so to speak, of the same basic structure, but the differences between them may have been sufficiently large to create a statistically significant result. Unfortunately, this is a distinct possibility, as our materials were not uniform, in that the test words were not always incorporated in the same syntactic structure in exactly the same part of the sentence. This was not possible given that we were trying to create natural-sounding sentences in which the accented syllable and the postaccentual syllable had the same segmental structure, while the position of the accent and its distance from the following one were manipulated.

The results of Experiment 3 must be considered with caution, because of the above-mentioned limitations and also because they are based on few speakers and few utterances, and show unexplainable variation among speakers. What we can conclude, however, from the results of all three experiments considered together is that the canonical conditions which allow us to observe the full form of the Greek prenuclear accents involve at least two unaccented syllables following the accented one, preferably within the same word as the accent. If these conditions are met, then we observe a very stable pattern of alignment of the H tone. If they are not met, then it appears that there are two main factors that may influence the H tone: the position of the accent relative to the word's right boundary and the number of postaccentual syllables. The exact role of each of these two factors cannot be discerned with any certainty from the present data. It is clear that further research, with better controlled materials, more speakers, and more repetitions is necessary to address this issue.

5. Discussion and conclusion

Our research on Greek prenuclear accents was originally motivated by issues arising within the autosegmental theory of intonational phonology. Our previous results (AL95) had suggested that there are two tonal targets involved in Greek prenuclear accentual rises, a L and a H. In the first experiment reported in the present paper, we were interested in providing evidence for the correct phonological analysis of these putative targets. As expected, this experiment replicated the results of AL95 regarding the stable alignment of the L target, but instead of shedding light on the phonological question, it produced the quite unexpected finding that the H target is very precisely aligned just after the beginning of the first postaccentual vowel. In a second experiment we replicated this new finding, showing that the alignment of the H is stable relative to the onset of the first postaccentual vowel, even in the presence of large differences in the combined duration of the accented syllable and the immediately following consonant. We also showed that there is no systematic relationship between the duration of the rise and the f_0 difference between the L and the H targets. The third experiment explored what happens when the H target is close to a word boundary and/or to a following accented syllable, and revealed a variety of phonetic effects that were slightly inconsistent from speaker to speaker. As a result of these findings, our conclusion about the appropriate phonological representation for Greek prenuclear accents is somewhat tentative: we think that in light of all the evidence these accents are best represented as bitonal accents of the form L+H*, but we are aware that the issue is not yet settled (for a thorough discussion on this and alternative representations see Arvaniti *et al.*, forthcoming).

Despite these limitations with respect to the phonological issues involved, we believe that our study is relevant to a much older and more basic issue in intonation research, namely, the so-called "levels *vs.* configurations" debate. As we noted in the Introduction, Bruce and the early autosegmental theorists converged on the general notion that the primitives of intonation systems are f_0 targets or level tones, and that f_0 movements are only transitions or interpolations between tonal targets. However, there remains a variety of researchers who reject the idea that tonal targets are primary and that rises and falls are merely transitional. These sceptics espouse the traditional idea (e.g., Pike, 1948; Jakobson, Fant & Halle, 1952; Crystal, 1969; Lehiste, 1970) that linguistic pitch can only operate in terms of relative f_0 within a specific utterance; e.g., high pitch can only be perceived relative to lower pitch somewhere in the immediate phonetic context. This in turn implies that pitch features of language ultimately involve pitch *change* or *movement*. Among the adherents of this view are the followers of the "British" tradition of intonation description (e.g., O'Connor & Arnold, 1973; Brown, Currie & Kenworthy, 1980), and the intonation researchers of the IPO school. In particular, the IPO model of Dutch intonation (now extended to numerous other languages) is based on the idea that pitch accents are rises and falls of various types, accompanying lexically stressed syllables and rendering them prominent in the utterance (see 't Hart, Collier & Cohen, 1990, for a full summary and reference to other works; see also Ladd, 1996, Section 1.2, for a review). For the researchers who espouse these principles, the autosegmental view is at best a trivial notational variant of a description expressed in terms of pitch movements: in Bolinger's words, "to the extent that intonation is dynamic, it makes no difference, in describing a movement, whether one says 'first you are going to be up and then you are going to be down' or 'you are going to go down'" (1986: 225f).

In our view, however, it does make a difference. In particular, the two approaches make different predictions with regard to the *rate* at which f_0 changes and the *duration* of f_0 movements. In a theory in which movements are primary, it is reasonable to expect that a given type of movement will have relatively constant rate or slope. As a result, its duration should be a function of how far it has to go: if the command is simply "you are going to go down", then the system might be expected to respond by starting down at some standard rate and continue going down until it reaches some appropriate level. This expectation is made explicit by the adherents of the traditional British view (among others, Ashby, 1978). The IPO researchers suggest that not only the rate is invariant, but the duration of the movement (what they term "size") is constant as well. The "size" of a pitch movement is one of the distinctive characteristics of intonational units ('t Hart *et al.*, 1990: 72). In either version of this "invariance hypothesis", the duration of movements is not supposed to be influenced (at least not to any considerable extent) by segmental effects.

By contrast, in a theory in which targets are primary, it is obvious that the slope and duration of a pitch movement will depend entirely on where the targets are. If, in the Modern Greek case, the command is "first (at the onset of the accented syllable) you are going to be down, and then (at the onset of the following unaccented vowel) you are going to be up", then the system obviously has to calculate its trajectory so that it gets to the right place at the right time. The whole notion of "the right place at the right time" is the essence of the target-based theory, and is, at the very least, unexpected in a theory that treats movements as primes. Very clearly, our results, like Prieto *et al.*'s, support this implicit prediction of the pitch target view: the timing and scaling of the beginning and ending of the prenuclear accentual rise in Greek is not determined by properties of the rise *qua* pitch movement, but rather the exact opposite is true. The L and the H of the accent are anchored to segmentally defined positions, and the duration and slope of the pitch movement are completely determined by the segmental composition of the accented word.

The debate between the pitch movement view and the tonal target view has long been dogged by the fact that much of the empirical evidence is consistent with either theory. We believe that the findings reported here now make it possible to argue more strongly for one view over the other. Unlike much previous experimental evidence for target levels and for the consistency of target alignment, our results are extremely difficult to reinterpret in a way that makes sense in terms of the pitch movement view. We believe that it is now incumbent upon those who espouse that view to make clear how the notion that movements are primitives can be reconciled with the evidence presented here.

The research reported here was supported by the U.K. Economic and Social Research Council (ESRC) through grant no. R000 23 5614 to Edinburgh University. This support is gratefully acknowledged.

References

Arvaniti, A. (1987) *Temporal coordination of initial consonant clusters in Greek*. MPhil dissertation, University of Cambridge

Arvaniti, A. (1994) Acoustic features of Greek rhythmic structure, *Journal of Phonetics*, **22,** 239–268

Arvaniti, A. & Ladd, D. R. (1995) Tonal alignment and the representation of accentual targets. In *Proceedings of the 13th international congress of phonetic sciences*, *Stockholm* (K. Elenius & R. Branderad, editors) Vol. 4, pp. 220–223

Arvaniti, A., Ladd, D. R. & Mennen, I. (forthcoming) What is a starred tone? Evidence from Modern Greek. Paper presented at the 5th Conference on Laboratory Phonology, Northwestern University, 6–8 July 1996

Ashby, M. (1978) A study of two English nuclear tones, *Language and Speech*, **21,** 326–336

Bolinger, D. (1986) *Intonation and its parts*. Stanford: Stanford University Press

Botinis, A., Fourakis, M. & Katsaiti, M. (1995) Acoustic characteristics of Greek vowels under different prosodic conditions. In *Proceedings of the 13th international congress of phonetic sciences*, *Stockholm* (K. Elenius & R. Branderad, editors) Vol. 4, pp. 404–407

Brown, G., Currie, K. L. & Kenworthy, J. (1980) *Questions of intonation*. London: Croom Helm
Bruce, G. (1977) *Swedish word accents in sentence perspective*. Lund: Gleerup
Carlson, R. & Granström, B. (1986) A search for durational rules in a real-speech data base, *Phonetica*, **43,** 140–154
Crystal, D. (1969) Prosodic systems and intonation in English. Cambridge: Cambridge University Press
Crystal, T. H. & House, A. S. (1988) Segmental durations in connected-speech signals: current results, *Journal of the Acoustical Society of America*, **83,** 1553–1573
Fletcher, J. & McVeigh, A. (1993) Segment and syllable duration in Australian English, *Speech Communication*, **13,** 355–365
Fourakis, M. (1986*a*) An acoustic study of the effects of tempo and stress on segmental intervals in Modern Greek, *Phonetica*, **43,** 172–188
Fourakis, M. (1986*b*) A timing model for word-initial CV syllables in Modern Greek, *Journal of the Acoustical Society of America*, **79,** 1982–1986
Goldsmith, J. (1976) *Autosegmental phonology*. PhD dissertation, MIT, published by Garland Press, New York, 1979
Grice, M. (1995*a*) Leading tones and downstep in English. *Phonology*, **12,** 183–233
Grice, M. (1995*b*) *The intonation of interrogation in Palermo Italian: implications for intonation theory*. Tubingen: Niemeyer
't Hart, J., Collier, R. & Cohen, A. (1990) *A perceptual study of intonation: an experimental-phonetic approach*. Cambridge: Cambridge University Press
Jakobson, R., Fant, G. & Halle, M. (1952) *Preliminaries to speech analysis*. Cambridge, MA: MIT Press
Ladd, D. R. (1983) Phonological features of intonational peaks. *Language*, **59,** 721–759
Ladd, D. R. (1988) Declination 'reset' and the hierarchical organization of utterances. *Journal of the Acoustical Society of America*, **84,** 530–544
Ladd, D. R. (1996) *Intonational phonology*. Cambridge: Cambridge University Press
Leben, W. (1976) The tones in English intonation. *Linguistic Analysis*, **2,** 69–107
Lehiste, I. (1970) *Suprasegmentals*. Cambridge, MA: MIT Press
Liberman, M. (1975) *The intonational system of English*. PhD dissertation, MIT
Liberman, M. & Pierrehumbert, J. (1984) Intonational invariance under changes in pitch range and length. In *Language sound structure* (M. Aronoff & R. Oerhle, editors) pp. 157–233. Cambridge, MA: MIT Press
Mirambel, A. (1959) *La langue grecque moderne: description et analyse*. Paris: Librairie Klincksieck
O'Connor, J. D. & Arnold, G. F. (1973) *Intonation of colloquial English*. London: Longman
Peterson, G. E. & Lehiste, I. (1960) Duration of syllable nuclei in English, *Journal of the Acoustical Society of America*, **32,** 693–703
Pierrehumbert, J. (1980) *The phonology and phonetics of English intonation*. PhD Dissertation, MIT
Pike, K. L. (1948) *Tone languages: a technique for determining the number and type of pitch contrasts in a language, with studies in tonemic substitution and fusion*. Ann Arbor: University of Michigan Press
Prieto, P., van Santen, J. & Hirschberg, J. (1995) Tonal alignment patterns in Spanish, *Journal of Phonetics*, **23,** 429–451
Shriberg, E. E., Ladd, D. R., Terken, J. & Stolcke, A. (1996) Modeling pitch range variation within and across speakers: predicting f_0 targets when "speaking up". In *Proceedings of the 4th international conference on spoken language processing*, Supplement, Philadelphia, PA, pp. 1–4
Silverman, K. & Pierrehumbert, J. (1990) The timing of prenuclear high accents in English. In *Papers in laboratory phonology I* (J. Kingston & M. Beckman, editors) pp. 72–106. Cambridge: Cambridge University Press
Silverman, K., Beckman, M., Pitrelli, J., Ostendorf, M., Wightman, C., Price, P., Pierrehumbert, J. & Hirschberg, J. (1992) TOBI: A standard for labeling English prosody. In *Proceedings of the 2nd international conference on spoken language processing*, Banff, Canada, **2,** 867–870