

PHONETIC AND PHONOLOGICAL PROPERTIES OF TONAL TARGETS IN MODERN GREEK INTONATION

**ESRC Award Reference No. R000 23 5614
Attachment to END OF AWARD REPORT**

2. FULL REPORT OF RESEARCH ACTIVITIES AND RESULTS

1. Background

The research reported here is part of the developing approach to the study of sound in language known as "Laboratory Phonology" (Beckman and Kingston 1990). In this approach, data from the phonetics laboratory are used to investigate not the physical aspects of speech production and perception, but the nature of the phonological structures that underlie speech; in Kingston's succinct formulation, "phonological representations embody hypotheses about phonetic behaviour". Specifically, our project aimed to shed light on a number of general theoretical problems within the autosegmental approach to the phonology of intonation (Pierrehumbert 1980, Ladd 1996), by means of several carefully controlled studies of intonation in Modern Greek, building on our own earlier work on Greek (Arvaniti 1991, Arvaniti 1994, Arvaniti and Ladd 1995).

The issues we were especially interested in studying revolve around the phonetic properties of tonal targets. Unlike many earlier approaches to intonation, which describe contours in terms of global shapes and overall slopes, the autosegmental theory assumes that contours are best analysed as strings of tones - sequences of Highs and Lows that occur at well-defined points in the utterance. In this view, the tones are manifested as specific pitch targets whose "scaling" (pitch level) and "alignment" (temporal coordination with the segmental string) can be measured instrumentally, and the pitch in between these tonal targets is determined by simple interpolation. There is good empirical evidence for several specific aspects of this general theory, but at the same time there remain a number of fairly fundamental questions relating to the notion of "tone" and "tonal target", and to the possible structural configurations that tones can enter into. The overall goal of the project was to provide a wealth of solid experimental phonetic data on the factors that affect targets in different contexts, and to bring this evidence to bear on the theoretical issues.

2. Objectives

The specific objectives are laid out below as stated in the original proposal, followed by a brief discussion of how each one was met. More detail is provided in the Results section below.

(1) To provide a thorough phonological description of the intonation of Modern Greek, on the basis of a substantial new body of acoustic and perceptual data. As outlined under Results, we have produced the substantial body of new acoustic data as planned, including extensive datasets relevant to intonation in yes-no questions, WH-

questions (question-word questions), emphatic statements, and prenuclear accents in declarative utterances; we have perceptual data on the distinction between emphatic statements and yes-no questions; and we have (as yet unanalysed) data on the intonation of intransitive statements with varying word order. With the exception of the unanalysed data on intransitives, we believe that the correct phonological analyses of the intonational types listed emerge quite clearly from our findings.

We are uncertain at this point whether a professional journal in linguistics or phonetics would publish a straightforward descriptive overview of the Greek intonational system. For one thing, such a paper would substantially overlap the publications that the project has already produced or is about to produce. For another thing, the datasets we have produced leave certain lacunae, most notably to do with nuclear and post-nuclear accents in non-emphatic statements, the intonational treatment of clitics, and the range of possible intonational resolutions of stress clash. However, we do intend to produce a technical report containing more quantitative detail than is possible in a typical journal article, covering at least the main intonational types for which we have new data. Moreover, Arvaniti has been asked to provide a concise description of the pronunciation, rhythm, and intonation of Greek for the grammar of Greek in the Council of Europe's "Threshold" series, and this will be heavily based on our findings. See further under Outputs.

(2) To use this body of data as a foundation for more satisfactory solutions to certain persistent problems of the generally successful autosegmental theory of intonation, including (a) the "pitch accent"/"boundary tone" distinction and (b) the identification of intonational targets.

We feel that this has been one of the two clear theoretical accomplishments of the project, in particular with regard to (a). As outlined in the Phonology paper included with this report, our data on "phrase accents" in yes-no questions show clearly, and our data on WH-questions strongly suggest, that such accents associate with lexically stressed syllables wherever possible. Since "phrase accents" are supposed to be comparable to boundary tones in orthodox autosegmental theory, our clear empirical documentation that they behave like pitch accents is important, and we have outlined what we see as the theoretical implications in the paper. As for subgoal (b), our most important general contribution is providing a great deal of new evidence that putative F0 targets are quite invariant phonetically, and thus bolstering the view that F0 contours are underlyingly sequences of targets. One of our more specific contributions in this regard is to document the phonetic invariance of the F0 valley that precedes prenuclear accents. We feel that this invariance should be taken as evidence that this valley is an intonational target corresponding to a Low tone, but we are aware that not everyone in the field shares this view. Some discussion of this issue is included in Ladd's chapter in the Bruce festschrift (see Outputs), which unfortunately cannot be included with this report. (3) By providing a detailed body of data on a language clearly distinct from the Germanic languages on which much intonation theorising is based, to gain new insight into the cross-linguistic typology of intonation systems, particularly with respect to the notions of alignment and secondary accent.

This is what we see as the other important theoretical contribution of this project, in particular our new discovery about the segmental anchoring of alignment, which is sketched under Results below and reported in detail in the Journal of Phonetics paper

included with the report. As noted under Results, this unexpected discovery made it necessary for us to devote considerable time to an unplanned follow-up experiment (Expt. 2a) and to rethink the theoretical implications, which somewhat curtailed the attention we could give to other issues. As for "secondary accent", some of the questions we planned to investigate under this heading form part of our findings on "phrase accent" mentioned just above. We had also planned to address this in greater detail in the analysis of the intransitive statement dataset, which as we noted in the discussion of the first objective was recorded but remains unanalysed.

(4) Possibly to contribute to the methodology of the investigation of intonational phenomena, in the interplay between acoustic and perceptual studies.

We do not feel we have made any notable contributions in this area, primarily because, as already mentioned, the perceptual side of the project was less extensive than originally hoped.

(5) Possibly to provide a basis for more successful approaches to Greek intonation in Greek (or multi-language) speech synthesis systems.

We think our data, especially our data on question intonation, should be directly applicable in the improvement of synthetic intonation. We discuss this briefly in the papers we are about to submit to the Journal of the Acoustical Society of America (see Outputs), which unfortunately cannot be included with this report.

3. Methods

As planned, we performed two main types of experiments: (i) production experiments, in which carefully controlled speech materials were read aloud under studio recording conditions; and (ii) perception experiments, in which the intonation of naturally-spoken utterances was systematically modified by means of digital resynthesis and the resulting stimuli were presented to listeners for judgement (e.g. question vs. statement). The procedures and techniques involved in these experiments were fairly standard and we had experience with them from previous work. These include constructing controlled speech materials, digital resynthesis of intonation, and segmenting utterances from spectrographic displays. The production experiments all had factorial designs which were analysed by Analysis of Variance (ANOVA).

4. Results

a. Segmental anchoring of tonal targets

The first, most unexpected, and possibly most important finding of the project was the discovery that certain tonal targets are aligned relative to landmarks in the segmental string. Specifically, the F0 valley and the F0 peak associated with prenuclear rising accents are aligned with the beginning of the onset consonant of the accented syllable and the beginning of the following unaccented vowel, respectively. This finding is described in detail in the Journal of Phonetics paper included with this report. As we argue in the paper, this finding has implications for the long-standing debate over whether to treat targets or movements as the primitive elements of intonation. In our data, the targets exhibit constant alignment and have roughly constant scaling for any

given speaker, whereas the properties of the rises that one might expect to be constant if pitch movements were primes (e.g. slope and duration) are quite variable, being entirely determined by the alignment and scaling of the targets.

The general theoretical issue of targets and pitch movements supersedes the more parochial issue of whether the H tone is an accent tone or an edge tone, which is what we were attempting to resolve by our experiments. The clear implication of our findings is that the alignment of the H is determined by the location of the accented syllable, not the word boundary, and is therefore part of the accent. (Among other things, this meant that it was less important for us to do a follow-up perceptual experiment, which we had tentatively called Expt. 2p, manipulating the alignment of a F0 peak with the aim of affecting the percept in the direction of either an accent tone or an edge tone; we still hope that it will be possible to do a related experiment in the relatively near future, but this is in no sense part of the project.) On the other hand, the new findings do raise questions for what it means for a tone to be "associated" with an accented syllable, since in the case of the Greek prenuclear accent neither the L nor the H is aligned in time with the accented syllable itself, though the alignment of both is determined relative to that syllable. We spent a considerable amount of time discussing the implications of this fact for autosegmental theory, in particular for the notion of the "starred" tone, which is supposedly the one tone in a bitonal accent that co-occurs with the accented syllable. This discussion is essentially what we reported in the paper presented to the 1996 Laboratory Phonology conference (see [Activities](#)).

It is worth emphasising that we discovered the segmental anchoring of tonal alignment largely by accident, while we were preparing to investigate the issue of whether the H tone is an accent tone or an edge tone. Specifically, Expt. 1 was intended as a control experiment for our planned Expt. 2, which would have explored the influence of the number of unaccented syllables between accents on the alignment of the H; before doing Expt. 2, we wanted to have some idea of how much variability in alignment is conditioned by differences in the segmental composition of the experimental materials. As it turned out, there appeared to be very little variability in alignment, so long as alignment is characterised in the appropriate way relative to segmental landmarks. Expt. 2a was a straightforward test of this claim, and as such was not planned for in the original proposal. Expt. 2b was more like the original experiment we had planned, in the sense that it explored the effects of varying the number of unaccented syllables and the location of word boundaries relative to the accented syllable; however, in the new context its focus was on determining to what extent "tonal crowding" would give rise to deviations from the basic segmentally-anchored alignment pattern discovered in Expts. 1 and 2a. Note that Expts. 1, 2a, and 2b correspond to what we call Expts. 1, 2, 3 in the [Journal of Phonetics](#) paper.

b. Stress-seeking alignment of "phrase tones"

The second major finding of the project was the experimental confirmation of the "impressionistic" linguistic description on which the proposal was based, specifically with regard to the behaviour of phrase tones in yes/no questions in Expt. 3. We established that the high peak that occurs late in the typical yes/no question contour is aligned with the last syllable of the sentence if the nuclear accent is on the final word, but with the last stressed syllable of the sentence if the nuclear accent is on an earlier word. As a very general implication, this work exposes the shortcomings of traditional

instrumental descriptions of intonation, which emphasise quantitative data reduction (e.g. characterisation of the overall "declination" of a contour) at the expense of local detail. In one of the papers we are about to submit to JASA, we emphasise this aspect of our work, and provide enough quantitative details of the alignment that our work could be used as the basis for improving Greek speech synthesis. It is worth noting here that the alignment data exhibit the same degree of constancy that we found in Expts. 1, 2a, and 2b: as a rule, the peak is aligned just after the onset of the vowel when it occurs on the last syllable of the sentence, and approximately in the middle of the vowel when it occurs on the last stressed syllable. However, in both cases, "tonal pressure" is present when the last word of the sentence has lexical stress on the last syllable, and this pressure substantially affects the alignment.

From the point of view of the autosegmental theory of intonational phonology, the important implication of this finding is, as we noted in the proposal, that the final F0 peak of the question contour is a "phrase tone" which nevertheless seeks a stressed syllable. As we discuss at length in the Phonology paper included with this report, this poses problems for the Beckman/Pierrehumbert conception of the phrase tone as the edge tone for a specific type of prosodic domain, viz. the "intermediate phrase". Rather, we think that this vindicates Bruce's (and Pierrehumbert's) original idea that the phrase tone is an additional accent. This in turn has two important implications, one of which we have been able to confirm. The first is that intonational tunes are not merely strings of accents and boundary tones, but have an internal structure that specifies metrical relations of relative strength or prominence (in particular, phrase tones are accent tones that form a constituent with the nuclear tone and are subordinate to it). The second is that in nuclear falling accents, which according to Pierrehumbert consist of a H accent tone and a L phrase tone, the alignment of the L phrase tone should be affected by the location of lexical stress, in exactly the same way that the H phrase tone of the yes/no contour is affected. It is this second implication that we have successfully tested, in Expt. 4.

In Expt. 4 we manipulated the length and stress pattern of the low-pitched stretch of speech following a nuclear falling accent. Specifically, we constructed WH-questions (which in Greek normally have the nuclear accent on the WH-word, and normally have a falling nuclear and a final boundary rise), such that the interval between the stressed syllable of the WH-word and that of the following content word was varied (0, 2, or 3 unstressed syllables). For example:

H* L LH%
'pu 'menune 'where do they live?' (zero intervening syllables)

H* L LH%
'pu me peri'menune 'where are they waiting for me?' (3 intervening syllables)

We showed that the location of the "elbow" in the F0 contour - the point at which the steep fall from the nuclear F0 peak ends and the contour flattens out - is influenced by the location of the post-nuclear stressed syllable: it is earliest when there are no unstressed syllables intervening, and latest when there are three. Similar effects were observed for the location of the final elbow, i.e. the beginning of the final rise. This suggests strongly (i) that the elbow represents an independent L tone, as

Pierrehumbert originally suggested in her analysis of English, and (ii) that the L tone preferentially associates with lexically stressed syllables.

Finally, we should mention Expt. 3p, which dealt with the distinction between emphatic statement intonation and yes/no question intonation. To the English-speaking listener, the striking thing about the H phrase tone in a Greek yes-no question is that it sounds like an emphatic statement accent. In practice there is little likelihood of confusion, because in the emphatic statement the H peak is the nuclear accent, whereas the H peak in questions is post-nuclear. (So for example

L* H L%

θa 'fiyun ap to 'livano 'Are they leaving Lebanon?'
would never be heard as a statement because of the low-pitched nuclear accented material preceding the H accent; conversely

H* L L%

sto 'livano 'in Lebanon!'

would never be heard as a question because the preceding low-pitched material is absent.) Nevertheless, Greek native speakers are unaware of any ambiguity, and if their attention is drawn to it they insist that the two peaks sound different. On the basis of data from Expt. 3 (which included emphatic statements, although these are not included in the write-up submitted to JASA in order not to make the paper too diffuse), we concluded that the nuclear accent peaks of emphatic statements (ES) are distinguished from the phrase tone peaks of yes/no questions (YNQ) primarily through differences of alignment (some individual speakers also have consistent differences in the scaling of the peak, but for some of these speakers ES have higher peaks while for others YNQ have higher peaks). Specifically, there are two key points: (1) the peak in ES is generally slightly earlier than that in YNQ, although it is difficult to be sure that this is the correct interpretation of the data, because the onset consonants of the accented syllables tend to be longer in ES than in YNQ; (2) the fall from the peak in ES is more rapid than in YNQ. Both these points are consistent with the putative phonological analysis, in which the final fall in YNQ consists of a H phrase tone followed by a L final boundary tone, whereas the final fall in ES consists of a nuclear H accent tone followed by two L tones, a phrase tone and a final boundary tone. The existence of a phonological difference based on F0 rather than duration is confirmed by Expt. 3p, in which we asked people to judge whether a sentence fragment like me ton Virona ('with Vironas') was an ES or the end of a YNQ. We showed that artificially "monotonised" naturally spoken ES and YNQ could not be distinguished by native listeners, whereas naturally spoken contours resynthesised with their original F0 contours were correctly identified as ES or YNQ much more often than would be expected by chance.

5. Activities

Much of the work of the project was designing, running, and analysing the five production experiments: creating the speech materials, making the recordings, digitising and making measurements on the speech files, and statistically analysing the results. This work is inherently slow and labour-intensive, especially the acoustic measurement phase, and it continued throughout the duration of the project,

somewhat longer than expected because of the unplanned experiment on segmental conditioning of alignment (Expt. 2a). The following is an approximate timetable:

[omitted]

We encountered no noteworthy difficulties in the overall running of the project. In particular, we encountered no difficulties in collaborating at long distance, with Arvaniti in Cyprus and Ladd and Mennen in Edinburgh. A great deal of detailed planning and discussion (e.g. details of experimental design and speech materials) was carried out smoothly by e-mail, supplemented by fax. We encourage the ESRC to see long-distance collaborations as potentially viable.

In addition, of course, Arvaniti made several visits to Edinburgh, as follows:

- (1) July 95, approx 4 weeks: Start-up discussions. Identifying appropriate speakers. Design of speech materials for Expt. 1, and first recordings and measurements. Supervising various aspects of Mennen's early work on the project.
- (2) January 1996, approx 10 days: Further analysis of Expt. 1 and design and recording of Expt. 2a. Work on Laboratory Phonology paper.
- (3) June 1996, approx 2 weeks: Final analysis of Expts. 2a and 2b. Final preparation for Laboratory Phonology conference. Write-up and submission of Journal of Phonetics paper. Design of speech materials for Expt. 3 and first recordings.
- (4) July-August 1996, approx 4 weeks: Write-up and submission of written version of Laboratory Phonology paper. First measurements and analyses of Expt. 3. Preliminary discussion of Expts. 3p and 4.
[note that visits (3) and (4) counted as a single visit for administrative purposes, as Arvaniti did not return to Cyprus in between these visits, but rather attended the Laboratory Phonology conference.]
- (5) April 1997. approx one week: Attending meeting of Linguistics Association of Great Britain. Discussion of results of Expt. 3 and preliminary results of Expt. 4.
- (6) August 1997, approx 2 weeks: final write-up of JASA and Phonology papers. Preparation of final report.

We attended the following conferences and presented the following talks as part of the activity of the grant:

- Arvaniti attended the Second International Congress on Greek Linguistics in Salzburg, Austria in September 1995; her paper (on emphatic accent in Greek) included references to the project work, which had just begun at the time.
- Ladd gave a talk in Utrecht in February 1996, sketching the findings of Expts. 1 and 2a.
- Arvaniti and Ladd attended the Fifth Conference on Laboratory Phonology at Northwestern University in Evanston, Illinois, USA, in July 1996. Arvaniti presented our paper on the phonological implications of Expts. 1, 2a and 2b.

- Arvaniti, Ladd and Mennen all attended the Spring meeting of the Linguistics Association of Great Britain in Edinburgh in April 1997. Ladd presented our paper on phrase accents.
- Ladd presented the phrase accent work at a one-day workshop on prosody at the University of Debrecen, Hungary, in April 1997.
- Ladd presented the phrase accent work at a one-day workshop on prosody at Cornell University, USA, in July 1997, in conjunction with the Linguistic Society of America's biennial summer institute.
- Arvaniti will attend the Conference on the Phonological Word in Berlin in October 1997, presenting a paper on Greek clitics which builds on some of our findings.
- Ladd will attend a workshop on accent, focus, and prosodic structure in Schloss Maurach, Germany, in December 1997. He will once again present the phrase accent work, this time to an audience of a dozen or so leading scholars from Europe and North America (Bruce, Gussenhoven, Selkirk, Drescher, and others) who represent the core specialist audience for our theoretical proposals.

We set up a small website for the project in March 1997, with a brief sketch of the project and a publication list, abstracts of papers, and links to Ladd's and Arvaniti's homepages. The website's URL is <http://www.ling.ed.ac.uk/~bob/greek/>.

6. Outputs

The most important output of the project is expected to consist of five refereed papers, four in major journals and one in an influential book series. The first (included with this report) has been formally accepted and scheduled for publication; the second will presumably be accepted but the editorial process for the volume as a whole has been delayed for various reasons beyond our control; the third and fourth, which are intended as companion papers, are nearly complete and will be submitted shortly; and the fifth (included with this report in draft form) will be submitted in January or February 1998, after Ladd's attendance at the workshop mentioned in the previous section. The five are as follows:

1. Arvaniti, Ladd, and Mennen. Stability of tonal alignment: the case of Greek prenuclear accents. Accepted for publication in Journal of Phonetics, to appear in January 1998. This describes the discovery of segmentally-anchored alignment in Expts. 1, 2a, and 2b.
2. Arvaniti, Ladd, and Mennen. What is a starred tone? Evidence from Modern Greek. Submitted for the fifth Laboratory Phonology volume, published by Cambridge University Press. This discusses the phonological implications of fixed tonal alignment outside the accented syllable for the theoretical notions of "association" and "starred tone".
- 3 and 4. Arvaniti, Ladd, and Mennen. Acoustic properties of tonal targets in Modern Greek question intonation I: Polar questions. and Acoustic properties of tonal targets in Modern Greek question intonation II: WH-questions. Companion papers to be submitted shortly to the Journal of the Acoustical Society of America (JASA). These two papers describe the findings of Expt. 3 and Expt. 4 (respectively) and their potential usefulness in speech synthesis. They both emphasise the important role of phonological theory in guiding our investigations.

5. Ladd, Arvaniti, and Mennen. On the place of phrase tones in intonational phonology. To be submitted to Phonology. This discusses the phonological implications of Expts. 3, 3p, and 4, together with data from other languages. Further planned outputs include: (1) a chapter by Ladd in the festschrift for Gösta Bruce, due to be published by Kluwer next year, which deals with a variety of issues and includes some discussion of the project's findings; (2) Arvaniti's chapter on Greek pronunciation and intonation in the "Threshold" series of language manuals (see Impacts below); and (3) a technical report describing the data in more quantitative detail, which will be listed on the project website and available from the Linguistics Department here. The first two are essentially complete and are expected to be published in 1998; the third will have to be produced during the coming academic year.

In addition, we have presented project work at conferences and in invited talks, and on the project website, as described above under Activities. Work from the project is also discussed in Ladd's 1996 book Intonational Phonology (CUP).

7. Impacts

It is too early to tell what the impact of the discovery of segmentally-anchored alignment may be, but we have definitely aroused the interest of several researchers, for both theoretical and practical reasons. As we say in the Journal of Phonetics paper, we feel that our findings have important theoretical implications for the long-standing issue of whether pitch targets or pitch movements are the primes of intonation; this claim, if we can judge from the reaction of one of the referees, seems likely to stimulate further research and lead to further published discussion. As for practical applications, our findings can be put to use in speech technology, especially speech synthesis (see further under Future Research Priorities below).

In this connection it is worth mentioning an experiment that has recently been completed by an M.Sc. student in the department, under Ladd's supervision. Daniel Faulkner, as the basis of his dissertation project, studied the effect of speech rate on the "rise duration" of prenuclear accents in English. His results show that the duration of the rise correlates strongly with the duration of the segmentals as rate increases, which is exactly what one would expect if segmental anchoring applies in English in the same way as in Greek. Faulkner's materials were not set up to test a precise hypothesis about segmental anchoring, but Ladd expects to be able to supervise an undergraduate Honours dissertation during the coming academic year which will directly investigate segmental anchoring under changes of rate; if the hypothesis is confirmed, the results of the two dissertations together should be publishable. Rate effects on intonation are not modelled in any existing speech synthesis applications that we know of, and several speech technology researchers have responded with interest to Faulkner's investigation.

The impact of the findings about phrase accents is more theoretical than applied. Even more than in the case of segmentally-anchored alignment, it is too early to tell what the impact of this work may be; we note only the apparent interest with which this work has been received on the three occasions when it has been presented publicly (see Activities above).

Finally, it should be noted that the work of the project will shortly put to direct practical use, as Arvaniti has been commissioned to prepare the chapter on pronunciation, rhythm, and intonation for the "Threshold Level" manual on Greek. This manual is being prepared by the Centre for Greek Language of the Greek Ministry of Education and is one of a series published by the Council of Europe. The idea of each book is not to be a grammar or a method for teaching the foreign language but a description (within the communicative framework for second language learning) of what a competent learner of a foreign language should know (hence "Threshold"). The books are supposed to be used as a guide by teachers of the language, to test whether they have covered these key points. The original book in the series was about English, first published in the 1970s and subsequently revised; since then other similar books have been produced about other languages. Greek is the twentieth language to be covered in this way, but will be the first to have a section on intonation based on modern phonological theory and instrumental phonetic research.

8. Future Research Priorities

We believe that the discovery of segmentally-anchored alignment should be followed up in other languages, especially languages with important speech technology industries, because of the direct applicability of alignment data in speech synthesis. We are pleased to note that the ESRC and its reviewers apparently agree with this view, since they have supported an application by Ladd and Mennen (R000-23-7447) to pursue this line of research in English and Dutch. It is hoped that work on the new project will begin sometime in the next 3-6 months.

References

- Arvaniti, Amalia (1991). The Phonetics of Modern Greek Rhythm and its Phonological Implications. Ph.D. thesis, University of Cambridge.
- Arvaniti, Amalia (1994). Acoustic features of Greek rhythmic structure. Journal of Phonetics 22: 239-268.
- Arvaniti, Amalia; Ladd, D. Robert (1995). Tonal Alignment and the Representation of Accentual Targets. Proceedings of ICPHS 13 (Stockholm), vol 4, pp. 220-223.
- Beckman, Mary; Kingston, John (1990). Introduction to J. Kingston and M. Beckman (eds.), Papers in Laboratory Phonology I, Cambridge: Cambridge University Press, 1-16.
- Ladd, D. R. (1996). Intonational Phonology. Cambridge: CUP
- Pierrehumbert, Janet (1980). The Phonology and Phonetics of English Intonation. PhD dissertation, MIT.