

# Greek ToBI: A System For The Annotation Of Greek Speech Corpora

Amalia Arvaniti<sup>†</sup>, Mary Baltazani<sup>‡</sup>

<sup>†</sup>Department of Foreign Languages and Literatures, University of Cyprus, P.O. Box 20537, Nicosia 1678, Cyprus  
[amalia@ucy.ac.cy](mailto:amalia@ucy.ac.cy)

<sup>‡</sup>Department of Linguistics, UCLA, 405 Hilgard Avenue, Los Angeles, CA 90095-1543, USA

## Abstract

Greek ToBI is a system for the annotation of (Standard) Greek spoken corpora, that encodes intonational, prosodic and phonetic information. It is used to develop a large and publicly available database of prosodically annotated utterances for research, engineering and educational purposes. Greek ToBI is based on the system developed for American English (ToBI), but includes novel features (“tiers”) designed to address particularities of Greek prosody that merit annotation, such as stress and juncture. Thus Greek ToBI includes five tiers: the *Tone Tier* shows the intonational analysis of the utterance; the *Prosodic Words Tier* is a phonetic transcription; the *Break Index Tier* shows indices of cohesion; the *Words Tier* gives the text in romanization; the *Miscellaneous Tier* is used to encode other relevant information (e.g., disfluency or pitch-halving). The development of GRTToBI is largely based on the transcription and analysis of a corpus of spoken Greek, that includes data from several speakers and speech styles, but also draws on existing quantitative research on Greek prosody.

## Introduction

Greek ToBI (henceforth GRTToBI) is a system for the annotation of (Standard) Greek spoken corpora, that encodes intonational, prosodic and phonetic information. GRTToBI is used to develop a large and publicly available database of prosodically annotated utterances for research, engineering and educational purposes. It is based on the system developed for American English (ToBI; Silverman *et al.* 1992; Beckman & Ayers-Elam, 1997), but includes novel features (“tiers”) designed to address particularities of Greek prosody that merit annotation. GRTToBI has largely been developed on the basis of a corpus of spoken Greek (much of which was especially collected for this purpose), that included data from several speakers and a variety of styles (read text, news broadcasting, interviews, spontaneous speech). The linguistic variety for which GRTToBI was conceived and designed is Standard Greek as spoken in Athens. It is our hope that GRTToBI will eventually be adapted for the annotation of corpora in other varieties of Greek, such as those of Thessaloniki and Cyprus.

This database fills a void, since in Greek there is a lack both of spoken corpora and of a systematic way of annotating them. GRTToBI is, to our knowledge, the first attempt at the prosodic annotation of Greek and the first systematic description of Greek prosody. Prosodically annotated corpora are very useful resources for research in a wide range of disciplines. From the linguistics point of view, they can be used for the better understanding of prosody, the importance of which in speech production, speech perception, and language acquisition has now begun to emerge (e.g., Jusczyk, 1997). From the engineering viewpoint, they are invaluable for the development of more natural speech synthesis and more efficient speech recognition systems. From a practical point of view, the new tiers proposed here for GRTToBI show the versatility of ToBI systems, which can be adapted to the prosodic organization of different languages, and to the research needs of particular sites by the addition of site-specific tiers.

GRTToBI currently works in conjunction with Waves<sup>+</sup> running on UNIX, and consists of (a) the waveform of the utterance, (b) its F0 track and (c) five annotation tiers: the *Tone Tier* that gives the intonational analysis of the utterances; the *Prosodic Words Tier*, which is a fairly narrow phonetic transcription; the *Words Tier* that gives the text in romanization; the *Break Index Tier* that shows indices of cohesion; and finally a *Miscellaneous Tier* in which other information may be entered. The Tones and Break Indices Tiers in particular are designed on the basis of our analysis of the prosodic and intonational structure of Greek within the framework of intonational phonology (Pierrehumbert & Beckman, 1988; Ladd, 1996). A brief overview of this analysis is given below. The analysis is followed by a detailed description of each annotation tier.

## Prosodic And Intonational Analysis of Greek: An Overview

### Stress

Greek is a stress accent language in which stress is acoustically manifested as *total amplitude* (Arvaniti, 1991; Arvaniti, 1994; Arvaniti, *subm.*). Primary stress is lexically determined and falls on one of the last three syllables of a word (e.g. Joseph & Philippaki-Warburton, 1987). There is one regular exception to this stipulation: content words stressed on the antepenult (or the penult) and followed by one (or two) enclitics have two stressed syllables; e.g. /fernodas to mu/ > [ferno'dastomu] “bringing it to-me”; for convenience we will follow the practice of Greek grammarians and call the added stress of such sequences “enclitic stress.”

### Intonational Phonology

For the intonational analysis of Greek we recognize three types of tonal events: *pitch accents*, which associate with stressed syllables, and two types of phrasal tones, *phrase accents* and *boundary tones*, which associate with the boundaries of intermediate and intonational phrases respectively. In contrast to stress, which as mentioned is lexically determined, the tones are morphemes that encode pragmatic information.

## The pitch accents

Greek has five pitch accents, L\*+H, L+H\*, H\*, !H\* and L\*. By far the most frequently used pitch accent is L\*+H. Phonetically, the L\*+H is manifested as a gradual rise from a trough (the L tone) to a peak (the H tone). In canonical conditions, that is if there are at least two unstressed syllables between consecutive accents, the L of a L\*+H is aligned at the very beginning or slightly before the onset of the accented syllable, and the H at the beginning of the first post-accentual vowel (Arvaniti *et al.*, 1998). An example is shown in Figure 1.

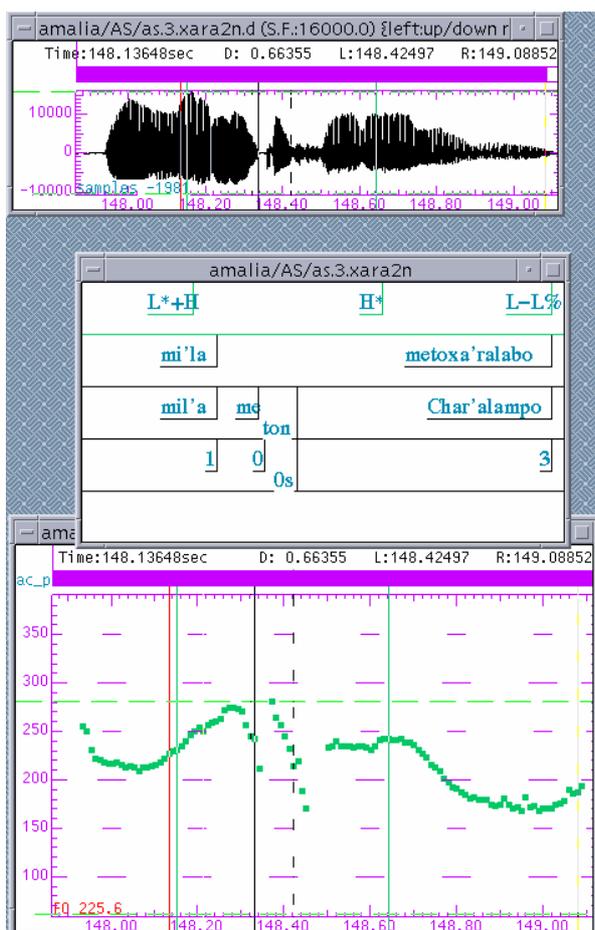


Figure 1: This example (gloss: “S/he’s talking to Charalambos”) shows typical L\*+H and H\* pitch accents. Note the alignment of the L\*+H, and the shape of the H\*. Details in the text.

The rather atypical alignment of the tones in the L\*+H accent has given rise to a great deal of fluctuation in its description (see also Arvaniti *et al.*, 2000, on the problems that the alignment of L\*+H may pose for the notion of starredness). Here we analyze this accent as L\*+H, because it is in contrast with another type of accent that can be unambiguously described as L+H\*. The difference between the two pitch accents lies in the alignment of the H tone, which in the L+H\* appears roughly in the middle of the accented vowel, as shown in Figures 2 and 3.

The L+H\* also contrasts with H\*. As can be seen in Figure 1, the H\* accent is realized as a peak on the accented syllable, but lacks the initial dip associated with

the L tone of the L+H\*. The H\* also contrasts with !H\*<sup>1</sup>, which is realized as a fall throughout the accented syllable, as illustrated in Figure 4. Finally, the L\* accent is typically realized as a low plateau, as shown in Figure 2 (see also Arvaniti *et al.*, ms.; Baltazani & Jun, 1999).

## The phrase accents

There are three types of phrase accent in Greek, H-, L- and !H-. The H- and L- phrase accents are always fully scaled, in contrast to other languages, such as English, in which the scaling of phrase accents is influenced by preceding or following tones, resulting in upsteps and downsteps (e.g. Beckman & Ayers-Elam, 1997). Because of this, falls or rises to mid pitch are represented in the Greek intonational system as !H- (and !H% as shown below).

## The boundary tones

Greek has three types of boundary tone, H%, L% and !H%. These boundary tones combine with the phrase accents in configurations that appear to have specific functions. These are schematically presented in Table 1 overleaf.

## Prosodic structure

The prosodic hierarchy we propose for Greek includes three levels at and above the word: the prosodic word, the intermediate phrase and the intonational phrase.

### The Prosodic Word

A prosodic word (PrWd) consists of a content word and its clitics. The term “clitic” here includes all items that in a given utterance lose their stress and form one PrWd with a host. In Greek this happens to many function words, including disyllabic ones which are not usually considered to be clitics; e.g. /a'po no'ris/ > [apono'ris] “since early.”

A PrWd has only one stress and thus it can only have one pitch accent. Prosodic words with enclitic stress, however, may have two pitch accents, one on the lexically stressed syllable of the host and one on the enclitic stress, although an accent on the former is not necessary. Despite the presence of two accents (or at least of two stressed syllables), such sequences are felt by native speakers to form one PrWd. PrWds are also the domain of several types of sandhi (juncture). Eight sandhi rules, some of which have previously been described and some of which are based on our corpus, are presented in Appendix I.

### The Intermediate Phrase and the Intonational Phrase

Above the PrWd level, we distinguish two levels of phrasing, the intermediate and the intonational phrase (*ip* and *IP* respectively). An *ip* must include at least one pitch accent (i.e. there are no headless phrases in Greek), and is tonally demarcated by the presence of a phrase accent, either a H- or a L-, at its right edge. An *IP* must include at least one *ip* and is tonally demarcated by the presence of a boundary tone (H%, L% or !H%) at its right edge.

<sup>1</sup> In deciding to use downstep (!) in the intonational analysis adopted in GRToBI, we follow the example of (American English) ToBI, in which downstep is marked independently of a trigger, so as to facilitate future research on the phonological analysis of downstep.

| Configuration | Schematic representation  | Usage   |
|---------------|---|---|
| L-L%          | _____   | Declaratives, negative declaratives, imperatives, wh-questions                                  |
| L-H%          |  | <i>Involved</i> continuation rise, <i>suspicious</i> calling contour                            |
| H-L%          |  | Yes-no questions, <i>requesting</i> calling contour   |
| H-H%          |  | Continuation rise, <i>questioning</i> calling contour   |
| L-!H%         |  | <i>Involved</i> wh-questions, requesting imperatives, negative declaratives showing reservation |
| H-!H%         |  | <i>Stylized</i> continuation rise   |
| !H-!H%        |  | <i>Stylized</i> calling contour, incredulous questions  |
| !H-H%         |  | Polite <i>stylized</i> calling contour  |

Table 1: Possible combinations of phrase accent and boundary tone and their usage.

There is abundant evidence for these two levels of phrasing in Greek. First, the tones associated with ips show a simple F0 movement, unlike the right edges of IPs that often show more complex pitch configurations. In the cases where the pitch movement is of the same type (e.g. a rise), ips and IPs show a difference in scaling, as illustrated in Figure 2 (in these and other similar data, the level of phrasing had been assigned and agreed upon by two transcribers independently of the scaling of the rise).

On the other hand, in IPs with complex final movement, such as a rise-fall, the two tones align independently. In the L-H% melody, for instance, the L- spreads, while the H% aligns with the last vowel of the utterance (Grice *et al.*, in press). This clearly shows that the tones do not form bitonal boundary tones; if that were the case, then the individual tones would align together at the edge of the relevant phrase.

Further, IPs (even non-final ones) may be followed by a lengthy pause, while pauses are rare after ips and always very short. It also seems likely that the IP (but not the ip) is the domain of downstep, pitch reset and final lengthening, though this is at present a tentative conclusion that requires further research.

In addition to the tonal evidence, our corpus suggests that at least some types of sandhi take place within ip boundaries but not across them. One such rule is consonant degemination, illustrated in Figure 2. At present, however, there is no coherent description of the sandhi rules of Greek and of their domains of application, especially as far as the rules above the PrWd level are concerned (but see below for a discussion).

## The GRTToBI Annotation System

### The Tone Tier

As mentioned, the Tone Tier presents the intonational structure of the utterance, using the analysis and criteria presented above. In addition to the pitch accents, phrase accents and boundary tones, some diacritics are also used

in the GRTToBI annotation system. These are largely employed to provide a more detailed description of the phonetic realization of the pitch accents.

Concretely, although the phonetic realization of the H\*, !H\* and L+H\* accents is relatively stable, that of the L\*+H and L\* accents shows variability. This is particularly noticeable in contexts of tonal crowding, i.e. when several tones must be realized within a short segmental stretch. A probable reason for this is that Greek favors the undershot realization of underlying tones to the truncation of some of them (for the distinction between undershooting and truncating languages, see Ladd, 1996). It follows that the L\*+H, which requires at least two syllables for its canonical alignment, will be the accent most prone to undershoot. Previous research (Arvaniti, 1994; Arvaniti *et al.*, 1998; Arvaniti *et al.*, 2000) and the data of our own corpus show that the speakers adopt mainly three strategies to cope with the tonal crowding of consecutive L\*+H accents. Specifically, they may (a) undershoot the L tone of the second of the two L\*+H accents; (b) realize the first accent earlier than normal and undershoot the second one; (c) realize the first accent earlier and the second one later than normal. Despite previous research on this issue, the realization of tonal targets under tonal crowding is not entirely understood; e.g. it is not clear whether the strategies mentioned above are a matter of free choice or depend on other prosodic factors, such as phrasing and metrical strength. Since this is still an open research question, we have decided to mark L\*+H pitch accents in tonal crowding contexts using three diacritics:  $wL^*+H$  is used when the L tone is undershot, as in [malone] in Figure 4;  $>L^*+H$  is used when the accent is realized earlier than typically expected, as in [ðalið̥a] in the same figure; and  $<L^*+H$  is used when the accent is realized later than typically expected.

The possible realizations of L\*+H implies that in Greek the undershooting of L tones is preferred to the undershooting of Hs. (Similar evidence on the undershooting of L% in Japanese [Venditti, to appear] suggests that different realization constraints may apply to

L and H targets universally.) The asymmetry between L and H tones is also supported by the variable realization of the L\* accent, which in cases of tonal crowding may be realized as rising, as shown in [mo'ro] in Figure 4. Similarly to the undershot L\*+H, we have decided to mark instances of undershot L\* as  $wL^*$ .

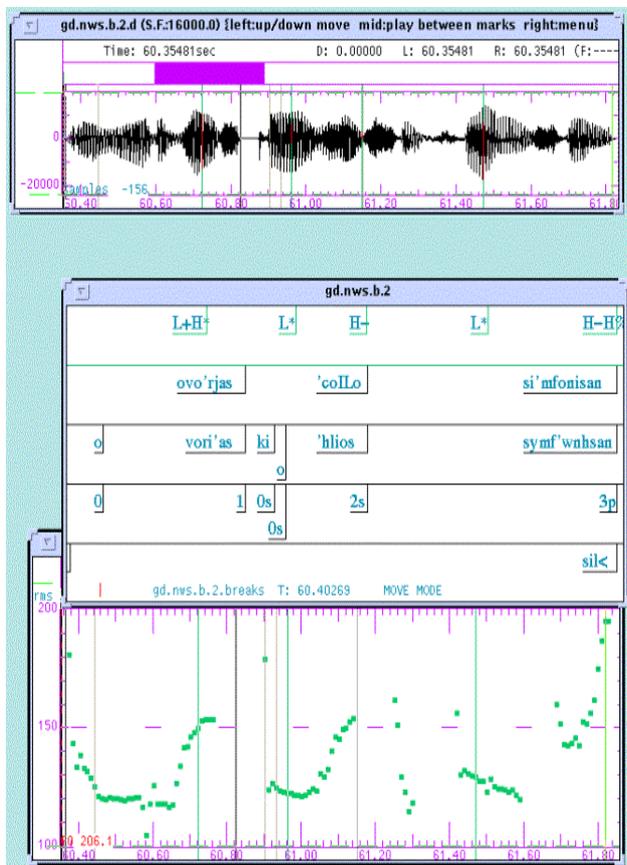


Figure 2: This example (gloss: “The North wind and the sun agreed...”), illustrates (a) canonical L\* accents (c.f. the undershot  $wL^*$  of Figure 4); (b) the difference in scaling between H-H% and H-; (c) consonant degeneration across ip boundaries (the break index labeled 2s).

### The Prosodic Words Tier

The Prosodic Words Tier provides a phonetic transcription of the utterances. Currently ASCII characters are used, but we hope that in the future the information on this tier will be presented in IPA notation (for the current conventions see Appendix II).

In this tier, each PrWd constitutes one label. The aim of the PrWords Tier is to provide the users of the database with information about the actual pronunciation of the utterances. To this purpose the transcription is phonetic rather than phonological, that is, it encodes stress, allophonic variation, phone deletions, assimilations and sandhi in general (but not, e.g., the precise quality of reduced vowels).

This tier is necessary for two reasons. First, it facilitates the analysis of sandhi and fast speech rules, which abound in Greek, by encoding their outcome. Second, it provides information about stress. This information cannot be

deduced from the transliteration (or from Greek spelling conventions for that matter), since in Greek only polysyllabic words are marked for stress. In a given utterance, however, a monosyllabic content word may be accented, while a disyllabic function word may be cliticized (i.e. lose its accent). By coding and examining such cases we hope to shed light on the relation between stress and accent in Greek.

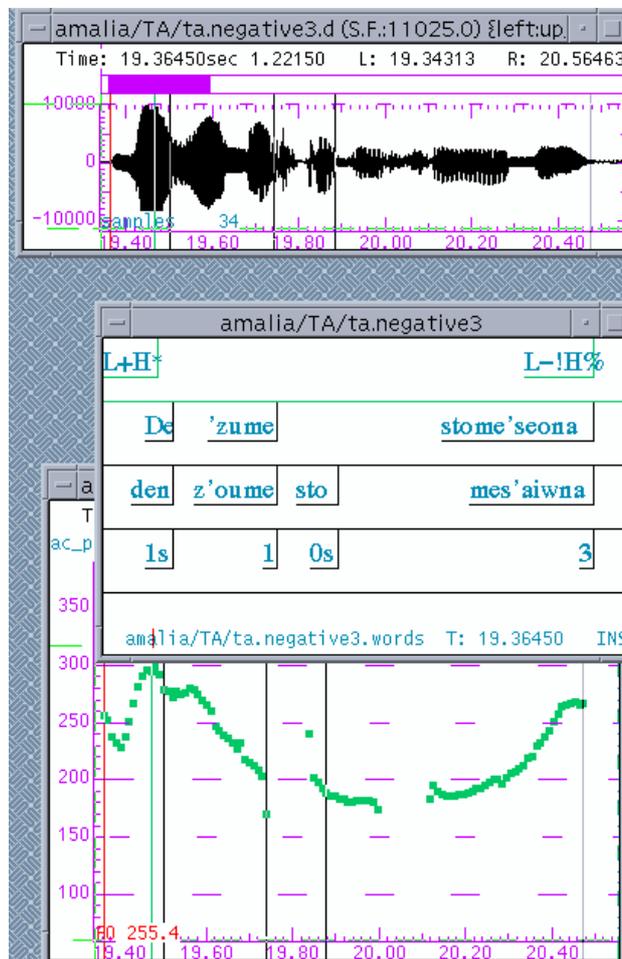


Figure 3: This example (gloss: “We do not live in the Middle Ages”) illustrates (a) a typical L+H\* accent; (b) the scaling of !H% (c.f. H% in Figure 2, relative to the other H tones in the same utterance); (c) sandhi between /den/ and /zume/. For details see text.

### The Words Tier

At present the Words Tier provides a word-by-word romanization of the text, although our long-term goal is to present this information in Greek orthography. In the absence of a generally agreed system for the romanization of Greek, we have followed some of the more generally accepted conventions (such as *ch* for  $\chi$ ) and have devised means for transliterating the rest of the characters. Our aim has been to represent each Greek letter and combination of letters with a unique roman character or set of characters, so that (a) searches of the Words Tier in the database yield unambiguous results and (b) the future algorithmic conversion to the Greek alphabet is possible. The full set of transliteration conventions can be found in Appendix III.

## The Break Index Tier

### The break indices

GRTToBI uses four levels of break indices, 0, 1, 2 and 3. These levels correspond to a *subjective* sense of increasing disjuncture between words. By *word* here we mean any item that is separated by spaces in the orthography of Greek; *orthographic* words often form but part of a *prosodic* word. It should be noted that although the use of a particular index relies on the transcriber's judgment, each index usually correlates with specific stress and tonal events.

BI 0 is used for a sequence of orthographic words that show total cohesion of the type typically expected between items that form one PrWd. Thus, we assume that a sequence of orthographic words separated by BI 0 corresponds to a PrWd that has only one stressed syllable and may bear only one pitch accent. As noted, cases with two accents due to enclitic stress are also felt to form one PrWd. Because of this sense of cohesion, the boundaries between hosts and enclitics are labeled BI 0. However, little is as yet known of the intonational behavior of such sequences (but see Arvaniti, 1992). Since this is still an open research question, we decided to flag the second accent in these cases by adding a label to it, namely "enclA" (for *enclitic accent*).

Although, as noted, several types of sandhi take place across a BI 0 boundary (see Appendix I for details), its presence is not a necessary condition for BI 0 to be used. For example, several forms of the Greek verbs include the proclitic particles, /θa/ or /na/; when the following verb stem begins with a consonant, no sandhi can take place between the particle and the verb; however, native speakers feel that these particles cannot be conceived but as part of the verb form. For this reason, BI 0 is marked in such cases.

BI 1 marks boundaries between PrWds. Tonally, items separated by BI 1 should carry one pitch accent each (or two, in cases of enclitic stress); PrWrds, however, need not always be accented. For instance, in cases of early focus in an utterance, de-accenting of all PrWds following the nucleus is expected, as illustrated in Figure 3 (see also Baltazani & Jun, 1999). Although the absence of accent does not constitute evidence that a given stretch is *not* a PrWd, the presence of an accent should be considered crucial for deciding that an item is a PrWd. Thus, when articles (which are normally proclitics) are accented, as often happens in media-speech (Arvaniti, 1997), then they are separated by BI 1 from the nouns that would normally be their hosts, and flagged with "accdCL" (for *accented clitic*) in the Miscellaneous Tier.

BIs 2 and 3 mark ips and IPs respectively. The arguments for these two levels of phrasing and a description of the tonal and other prosodic cues that accompany each of them are presented in the Prosodic and Intonational Analysis section.

### Diacritics for the Break Index Tier

In addition to the break indices, four diacritics are used to provide more detail on the prosodic structure of the annotated utterances. By far the most important diacritic for GRTToBI is *s*, which is used with all break indices

when there is evidence of sandhi (examples of sandhi across various constituents can be seen in all Figures).

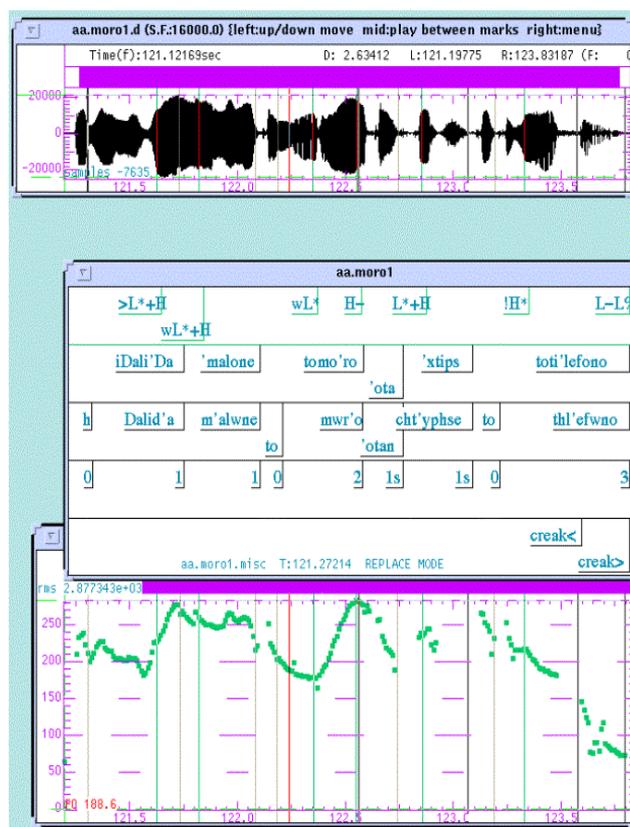


Figure 4: This example (gloss: “Dalida was scolding the baby when the phone rang”) illustrates (a) undershot realizations of L\*+H and L\* in tonal crowding; (b) a typical !H\*; (c) extensive sandhi. For details see text.

The reason why we decided to flag sandhi in this way is that, as noted, sandhi phenomena abound in Greek, but they are not coherently described and analyzed. Nespor & Vogel (1986) present an analysis of some sandhi rules of Greek (partly based on Kaisse, 1985). However, research based on naturally occurring speech and the native speakers' intuitions cast serious doubt on these analyses. First, the application of some rules presented in Kaisse (1985) and Nespor & Vogel (1986) depends extensively on the lexical items used in a given utterance (Arvaniti, 1991; Arvaniti, in prep.). More importantly, however, many rule descriptions do not tally with naturally occurring data, such as those in our corpus or those of Fallon (1994).

Thus, several types of sandhi that are supposed to be limited to lower level prosodic domains apply across larger constituents than postulated. A striking example of this is the sequence [ota'xtips] in Figure 4, in which the adverb /otan/ “then” loses its final /n/ before the verb /xtipise/ “rang”, although the two form separate PrWds (e.g. both remain stressed). According to Nespor & Vogel (1986) however, /n/-deletion before fricatives does not apply across PrWd boundaries. Figure 3 shows a similar case. In that figure the negative particle /den/, normally seen as a clitic, is accented, while the following verb /zume/ “we live” is also stressed (though unaccented);

thus /ðen/ and /zume/ form separate PrWds; yet, /n/-deletion again takes place across the PrWd boundary. Equally, the final /e/ of /xtipise/ in Figure 4 should not have been deleted for similar reasons. These utterances cannot be attributed to fast speech; for Figure 4, in particular, there is evidence that the utterance it is part of (and which was elicited under laboratory conditions) was rather carefully enunciated; this evidence comes from the words /malone/ and /ti'lefono/ which are realized as such, rather than as [ma|ne] and [tlefono] respectively, as would be expected in fast casual speech. Yet, such reduced pronunciations were happily used by other speakers in related laboratory recordings, suggesting that these processes of vowel reduction and sandhi are much more widespread in Greek than previously thought. The point we wish to stress, however, is that the presence of extensive sandhi in Greek, even across phrase boundaries (as in Figure 2), does not necessarily signal cohesion, as is often expected of sandhi.

For the above reasons, we decided to use the diacritic *s* to flag all instances of sandhi at all prosodic levels. We hope that by investigating a large corpus of spoken data thus marked, a better understanding of the possible environments for sandhi and of the prosodic constituents across which sandhi may apply can be reached. We anticipate that such research will have far reaching consequences in terms of the general understanding of the relation between sandhi, phrasing and prosodic structure.

The diacritic *m* flags two types of mismatch depending on the break index it is associated with. This diacritic is used with BI 0 to mark cases in which the context for sandhi at BI 0 exists but sandhi does not take place (i.e. the speaker chose not to apply the rule). The *m* diacritic can be used with BIs 1, 2 and 3 to mark cases in which the transcriber feels that a certain boundary is present, yet the stress or tonal events that normally accompany this boundary are not evident. For example, when the transcriber feels that a sequence which does not end in a phrase accent nevertheless forms an ip, then the boundary between this and the following ip should be labeled *2m*.

Finally, the other two diacritics for the Break Index Tier are *p*, which should be used to mark pause at a given boundary, and *?*, which is used to mark uncertainty about the strength of a boundary. In cases of uncertainty the highest of the two possible candidates is marked, together with a matching analysis in the Tone Tier.

### The Miscellaneous Tier

The purpose of the Miscellaneous Tier is to encode information about the utterance that is beyond the scope of the other tiers but may help the users in understanding the information encoded in those. Thus, comments such as disfluency, pitch-halving, speaking rate, and also the presence of accented clitics (the *accCL* label) are marked in this tier.

### Conclusion

GRTToBI, as a tool for the annotation of Greek corpora, concentrates on three aspects of spoken Greek: intonation, sandhi and phrasing. Although we believe that the intonational structure of the language is by and large

understood, several issues remain outstanding, such as the use, domain and purpose of downstep, or the realization of accents under tonal crowding. We believe that although issues concerning phonological analysis may be handled more efficiently by running controlled laboratory experiments, a better understanding of issues like downstep will benefit mostly from the examination of GRTToBI transcribed corpora. Yet, we believe that the greatest contribution of GRTToBI may come from the examination of sandhi and phrasing, and the relation between stress and accent; these issues are generally not well understood, and Greek, due to its abundant use of such phenomena, offers a fertile ground for further research in these areas.

### References

- Arvaniti, A. (1991). The Phonetics of Modern Greek Rhythm and its Phonological Implications. Ph.D. Thesis. University of Cambridge.
- Arvaniti, A. (1992). Secondary stress: evidence from Modern Greek. In G.J. Docherty & D.R. Ladd (Eds.), *Papers in Laboratory Phonology II: Gesture, Segment, Prosody* (pp. 398--423). Cambridge University Press.
- Arvaniti, A. (1994). Acoustic features of Greek rhythmic structure. *Journal of Phonetics*, 22, 239--268.
- Arvaniti, A. (1997). Greek "emphatic stress": a first approach. In *Greek Linguistics 95* (pp. 13--22). Salzburg: The Department of Linguistics, University of Salzburg.
- Arvaniti, A. (subm.). The phonetics of stress in Greek.
- Arvaniti, A. (in prep.). Investigating prosodic structure by means of spoken corpora annotation. 7<sup>th</sup> Conference on Laboratory Phonology, 29/6-1/7 2000. Nijmegen.
- Arvaniti, A., Ladd, D.R. & Mennen, I. (1998). Stability of tonal alignment: the case of Greek prenuclear accents. *Journal of Phonetics*, 26, 3--25.
- Arvaniti, A., Ladd, D.R. & Mennen, I. (2000). What is a starred tone? Evidence from Greek. In M. Broe & J. Pierrehumbert (Eds.), *Papers in Laboratory Phonology V* (pp. 119--131). Cambridge: Cambridge University Press.
- Arvaniti, A., Ladd, D.R. & Mennen, I. (ms). Quantitative F0 descriptions require reference to local targets: the intonation of Greek polar questions and emphatic statements.
- Baltazani, M. & Jun, S. (1999). Focus and topic intonation in Greek. In *Proceedings of the XIVth International Congress of Phonetic Sciences* (pp. 1305--1308). San Francisco, CA.
- Beckman, M. E. & Ayers-Elam, G. (1997). Guidelines for ToBI Labelling. The Ohio State University Research Foundation.
- Fallon, P. (1994). Naturally occurring hiatus in Modern Greek. In I. Philippaki-Warbuton, K. Nicolaidis and M. Sifianou (Eds.), *Themes in Greek Linguistics* (pp. 217--224). London: John Benjamins Publishing Co.
- Grice, M., Ladd, D.R., & Arvaniti, A. (in press). On the place of phrase accents in intonational phonology. *Phonology*.
- Joseph, B.D. & Philippaki-Warbuton, I. (1987). *Modern Greek*. London: Croom Helm.
- Juszyk, P.W. (1997). *The Discovery of Spoken Language*. Cambridge, MA: The MIT Press.

Kaisse, E.M. (1985). *Connected Speech: The Interaction of Syntax and Phonology*. Academic Press.

Ladd, D.R. (1996). *Intonational Phonology*. Cambridge: Cambridge University Press.

Nespor, N. & Vogel, I. (1986). *Prosodic Phonology*. Dordrecht: Foris.

Pierrehumbert, J. & Beckman, M. (1988). *Japanese Tone Structure*. Cambridge, MA: The MIT Press.

Silverman, K.E.A., Beckman, M., Pitrelli, J.F., Ostendorf, M., Wightman, C., Price, P., Pierrehumbert, J. & Hirschberg, J. (1992). ToBI: A standard for labeling English prosody. In *Proceedings of the 1992 International Conference on Spoken Language Processing*, (pp. 867--870). Banff, Canada.

Venditti, J. J. (to appear). The J\_ToBI model of Japanese intonation.

## Acknowledgments

GRTToBI was first presented at the XIVth ICPHS Satellite Meeting *Workshop on "Intonation: Models and ToBI labeling."* We are grateful to Sun-Ah Jun, Mary Beckman, Julie McGory, Shu-hui Peng, Amanda Miller-Ockhuizen and Mariapaola D'Imperio for their support and input, and for long distance technical help. Thanks are also due to the audience of Mary Beckman's and July McGory's ToBI course for their feedback at a first presentation of GRTToBI, and to Bob Ladd for his comments on a first draft. We are also grateful to Jenny and Peter Ladefoged for their kind hospitality to the first author during her stay in Los Angeles. The development of GRTToBI largely took place while the first author was on sabbatical leave at the Ohio State University Linguistics Laboratory. This leave is gratefully acknowledged.

## Appendix I: Sandhi Rules Within Prosodic Words

The following are the types of sandhi that unequivocally take place across a BI 0 (i.e. across orthographic words that form one PrWd). (Some of these types are presented in Nespor & Vogel, 1986; others are based on the data of our corpus.)

- /n/-resyllabification before a word-initial vowel; e.g. /o.tan. 'e.fta.se/ > [o.ta.'ne.fta.se] "when s/he arrived"; in accented syllables /n/-resyllabification is evident from tonal alignment.
- Stop-voicing after a word-final nasal (the nasal is usually deleted; if not, it assimilates for place of articulation to the stop); e.g. /tin 'poli/ > [ti'boli] or [ti'mboli] "the town" ACC.
- /n/-deletion before sonorants; e.g. /ton la'o/ > [to'la'o] "the people" ACC.
- /s/-voicing before sonorants; e.g. /jos mu/ > [jozmu] "my son."
- degemination of identical consecutive consonants; e.g. /jos su/ > [josu] "your son."
- degemination of identical consecutive vowels; /ta 'atoma/ > [tatoma] "the individuals."
- diphthongization of non-identical vowels, e.g. /o. 'i.ɫos/ [oi.ɫos] "the sun."
- deletion of one of non-identical vowels; e.g. /to 'atomo/ [tatomo] "the individual."

## Appendix II: Phonetic Transcription Conventions

| IPA                | ASCII  | IPA   | ASCII  | IPA    | ASCII  |
|--------------------|--------|-------|--------|--------|--------|
| p                  | p      | v     | v      | n / ɳ  | n / nn |
| t                  | t      | θ     | th     | ɲ / ɲ̃ | N / NN |
| k                  | k      | ð     | D      | l / ɭ  | l / ll |
| c                  | c      | s     | s      | r      | r      |
| b / <sup>m</sup> b | b / mb | z     | z      | ɭ / ɭ̃ | L / LL |
| d / <sup>n</sup> d | d / nd | ç     | X      | i      | i      |
| g / <sup>ŋ</sup> g | g / Ng | ʃ     | j      | e      | e      |
| ʃ / <sup>ɲ</sup> ʃ | J / NJ | x     | x      | ɐ      | a      |
| β                  | B      | ɣ     | G      | o      | o      |
| f                  | f      | m / ɱ | M / mm | u      | u      |

In addition to the above symbols, the following conventions should be used:

- Noticeably centralized vowels should be transcribed as @.
- Whispered vowels should be transcribed in brackets.
- Vowels that phonologically form separate syllables but are phonetically manifested as a rising diphthong (on the basis, e.g., of tonal alignment evidence), should be transcribed with the second vowel capitalized; stress should be placed before the diphthong.
- Stress should be marked before the consonant(s) of the stressed syllable, following IPA conventions. (At present we are agnostic as to syllabification, so we suggest that transcribers mark maximal onsets, unless tonal alignment or their own intuitions suggests otherwise.)

### Appendix III: Romanization Conventions

| GREEK | Romanization | GREEK | Romanization | GREEK   | Romanization |
|-------|--------------|-------|--------------|---------|--------------|
| α     | a            | ν     | n            | αι      | ai           |
| β     | v            | ξ     | x            | ει      | ei           |
| γ     | g            | ο     | o            | οι      | oi           |
| δ     | d            | π     | p            | ου      | ou           |
| ε     | e            | ρ     | r            | αυ      | ay           |
| ζ     | z            | σ     | s            | ευ      | ey           |
| η     | h            | τ     | t            | μπ      | mp           |
| θ     | 0            | υ     | y            | ντ      | nt           |
| ι     | i            | φ     | f            | γγ / γκ | gg / gk      |
| κ     | k            | χ     | ch           | τσ      | ts           |
| λ     | l            | ψ     | ps           | τζ      | tz           |
| μ     | m            | ω     | w            | ντζ     | ntz          |

- When the grapheme combinations that usually represent one vowel (e.g., αι) represent two separate vowels, the graphemes are separated by full stops; e.g. *a.i.d'oni* for *αἰδόνι*.
- Spellings with double graphemes are transliterated in the same way; e.g. *th'alassa* is transliterated as *th'alassa*.
- In words with more than one syllable, stress is marked as an apostrophe before the stressed vowel. Monosyllables bear no stress mark in the Words Tier.
- Initials capitalized in Greek orthography should be transliterated with capital letters as well.

### Appendix IV: Label Alignment Conventions

- The labels for the L+H\*, H\* and !H\* pitch accents should be aligned with the highest non-spurious F0 point of the accented vowel.
- For the L\* accent the lowest F0 point on the accented vowel should be chosen for alignment.
- For the L\*+H pitch accent, the canonical alignment of which is outside the accented syllable, a reliable point early in the accented vowel should be used instead.
- Phrase accents should be aligned with the right boundary of the relevant ip.
- Phrase accent and boundary tone combinations should be aligned with the right boundary of the relevant IP.
- The *enclA* label should be placed above or below the relevant accent in the Tones Tier.
- The transcriptions in the PrWords Tier should be aligned with the right edge of the whole sequence of orthographic items that form one PrWd.
- Transliterated forms are aligned at the right edge of words.
- Break indices are aligned at the right edge of relevant constituents.
- The *accCL* label should be aligned with the relevant accent, but should be marked in the Miscellaneous Tier.