INTONATIONAL ANALYSIS

AND PROSODIC ANNOTATION OF GREEK SPOKEN CORPORA[1]

Amalia Arvaniti

University of California, San Diego

and

Mary Baltazani

University of California, Los Angeles

4.1 INTRODUCTION

This paper provides an analysis of the prosodic and intonational structure of Greek within the autosegmental/metrical framework of intonational phonology (Pierrehumbert 1980; Pierrehumbert and Beckman 1988; Ladd 1996), and presents Greek ToBI (henceforth GRToBI), a system for the annotation of Greek spoken corpora based on this analysis. Both the analysis and the annotation system have largely been developed on the basis of a corpus of spoken Greek, especially collected for this purpose, and including data from several speakers and a variety of styles (read text, news broadcasting, interviews, spontaneous speech). The linguistic variety analysed—and the one for which GRToBI was conceived and designed—is Standard Greek as spoken in Athens. It is our hope that other varieties of Greek will be similarly analysed, and that eventually GRToBI will be adapted for the annotation of corpora in those varieties as well.

Greek is a stress accent language, described in traditional grammars as having "dynamic stress" (among others, Joseph and Philippaki-Warburton 1987). Stress is acoustically manifested as either longer duration or higher amplitude of the stressed syllable (or both), making *total amplitude* the most robust cue to stress (Arvaniti 1991, 2000).

Primary stress cannot be predicted from phonological structure, as there are no syllable weight distinctions in Greek and stress is not fixed. The main phonological limitation on the position of primary stress is that it fall on one of the last three syllables of the word (Joseph and Philippaki-Warburton 1987; Setatos 1974). In all other aspects, stress placement within this three-syllable window is largely determined by morphology (Drachman and Malikouti-Drachman 1999; Revithiadou 1998). At the lexical level, there are no other stresses in addition to primary stress.

The presence of *postlexical* rhythmic stresses, on the other hand, is disputed. Malikouti-Drachman and Drachman (1981), and Nespor and Vogel (1989) suggest that rhythmic stresses appear regularly to remedy stress lapses in Greek. This claim, however, is not supported either by acoustic evidence or by the native speakers' intuition (see Arvaniti 1994, and references therein). Thus, we assume here that at the postlexical level, as well as lexically, there is only one stressed syllable per word.

There is one regular exception to this stipulation, however: content words stressed on the antepenult (or the penult) and followed by one (or two) enclitics acquire an additional stress two syllables to the right of their lexical stress; e.g.

(1) a   /ˈfernodas to mu/ > [ˌfernoˈdastomu]

bringing it to-me

b   /to teˈtraðio mu/ > [to teˌtraðiˈomu]

the notebook my

For convenience, we will follow the practice of Greek grammarians and call the added stress of such sequences "enclitic stress," although it does not usually fall on the enclitic itself, as examples (1a) and (1b) show.

In contrast to words with enclitic stress, there are several disyllabic function words in Greek (e.g. (2a)-(2c)) which are normally uttered without stress, and thus form part of the following content word (with which they are syntactically, as well as prosodically, linked). Some of these words contrast with stressed homophones (cf. (2c) and (2d)).

(2) a   /aˈpo noˈris/ > [aponoˈris]

         since early

     b   /aˈna tin iˈfilio/ > [anatiniˈfilio]

         all-over the globe

         'everywhere'

     c   /kaˈta to ˈspiti/ > [katatoˈspiti]

         towards the house

     d   /kaˈta tu ˈɣamu/ > [kaˈta tuˈɣamu]

         against (the) marriage

With respect to rhythm, Greek has been described as syllable-timed, though Dauer (1983) places it somewhere in the middle of the continuum she proposes between stress- and syllable-timing. Arvaniti (1994), however, points out that by Dauer's own criteria (1983, 1987) Greek should be a prototypical syllable-timed language, virtually devoid of stress; such a description, however, is not supported either by acoustic data (e.g. Arvaniti 1991, 1992, 2000; Botinis 1989), or the fact that Greek has several minimal pairs and triplets distinguished solely by stress placement. Instead, Arvaniti (1994) proposes that rhythm in Greek (as in all languages) is based on the alternation of strong

and weak prosodic constituents. The difference between languages described as stress-timed, e.g. English, and languages described as syllable-timed, e.g. Greek, lies in the fact that the former tend to keep this alternation as even as possible, by using various strategies to eliminate stress clashes and lapses; in contrast, the latter languages seem to allow for less eurhythmic patterns, i.e. they tolerate clashes and lapses to a greater extent.

4.3 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. Therefore, it is not expected that every stressed syllable will be accented (see also section 4.4.1).

4.3.1 The pitch accents

Greek has five pitch accents, L*+H, L+H*, H*+L, H* and L*. By far the most frequently used pitch accent is L*+H, which is the predominant choice for pre-nuclear accented syllables. Because of its distribution, L*+H has been described as the "pre-nuclear" accent of Greek (Arvaniti, Ladd and Mennen 1998; Baltazani and Jun 1999). In our corpus, however, this accent was frequently attested in nuclear position, in calls, imperatives, negative declaratives and wh-questions.

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 L*+H accents, the L is aligned at the very beginning or slightly before the onset of the accented syllable, and the H early in the first post-accentual vowel (Arvaniti and Ladd 1995; Arvaniti *et al*. 1998). 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, Ladd and Mennen 2000, for a discussion of the problems that the alignment of L*+H poses for the notion of starredness in intonational phonology). In GRToBI this accent is analysed as L*+H, because our corpus and other quantitative data (Arvaniti *et al.* 1998; Arvaniti, Ladd and Mennen ms.) showed that it is in contrast with another accent which can be unambiguously described as L+H*.

Specifically, as illustrated in Figure 4.1, in L+H* the H tone appears roughly in the middle of the accented vowel, unlike L*+H which shows late alignment of the H tone. L+H* often used to signal narrow focus, as in Figure 4.9 (Arvaniti *et al*. ms.; Baltazani and Jun 1999; Botinis 1998).

///////////////////////// FIGURE 4.1 ABOUT HERE /////////////////////////

The two bitonal accents are in contrast with the monotonal H* accent. As can be seen in Figure 4.2, H* lacks the initial dip associated with the L tone of L+H* and L*+H. Rather, there is a declining plateau between the H tone of the L*+H accent and the nuclear H* (hence the use of the downstep diacritic, !, which however *does not* denote a distinct phonological category; see sections 4.3.2 and 4.5.1 for discussion). In Greek, this plateau does not exhibit the "sagging" posited by Pierrehumbert (1981) as a possible type of interpolation between successive H*s in American English. When H* is used as the nuclear accent in a declarative utterance, as in Figure 4.2, it signals broad focus, and thus contrasts with L+H*, which signals narrow focus in the same context.

//////////////////////// FIGURE 4.2 ABOUT HERE ////////////////////////

In nuclear position in declaratives, the H* also contrasts with H*+L. From a pragmatic point of view, H*+L conveys a more nonchalant (or even wearied) attitude on the part of the speaker than H*. Phonetically, it is realised as a fall from high pitch, with the fall being completed by the end of the accented syllable. It is useful to compare on this point the low stretch of the two contours shown in Figure 4.3, which illustrates the difference between H*+L and H*: as can be seen, when the pitch accent is H*+L the bottom of the speaker's voice is reached at the end of the accented syllable; in contrast, when the pitch accent is H*, this lowest F0 point is reached at the end of the first postnuclear syllable and the fall is realised in three discernible steps.

//////////////////////// FIGURE 4.3 ABOUT HERE ////////////////////////

It should further be noted that in an earlier version of GRToBI we had analysed H*+L as !H* (Arvaniti and Baltazani 2000). We have now revised our position, to bring the use of the downstep feature more in line with that in other ToBI systems, in which downstep—for obvious reasons—cannot be used utterance-initially (as we would be forced to do in Figure 4.3). Furthermore, it is clear that H* and H*+L are distinct both in terms of meaning and in terms of realisation, and thus analysing H*+L as a "scaled down" version of H* is not well motivated. Finally, it appears that !H* is needed for cases of "scaled down" H* pitch accents, like the one shown in Figures 4.2 and 4.8. Despite these indications against our original analysis, it is not the case that the analysis of this accent as H*+L is entirely unproblematic; for example, it is not obvious why in

Figure 4.3, H*+L is scaled lower than the utterance beginning. Clearly then the utterances in Figure 4.3 and similar examples do not provide sufficient evidence on which to choose one alternative over the other. Rather, controlled experiments are needed to provide a definitive answer to this problem.

Finally, the L* accent is typically realised as a low plateau, as shown in Figure 4.4. The L* appears as the nuclear accent before a "continuation rise" (Baltazani and Jun 1999), in yes-no questions (Arvaniti *et al*. ms.; Baltazani and Jun 1999) and in the calling contour we term "suspicious" (see Table 4.1).

//////////////////////// FIGURE 4.4 ABOUT HERE ////////////////////////

4.3.2 Downstep and the phonetic realisation of pitch accents

The descriptions presented above provide a phonological analysis of the Greek accentual system and a sketch of the phonetics of the pitch accents under canonical conditions. However, pitch accents show significant contextual variability as regards both the scaling and the alignment of their targets, with tonal crowding and downstep being the main influences.

Concretely, L*+H and L* exhibit noticeable variability in contexts of tonal crowding, i.e. when several tones must be realised within a short segmental stretch. Previous research (Arvaniti 1994; Arvaniti *et al*. 1998, 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 L*+H accent; (b) realise the first accent earlier than normal and undershoot the second one (as in Figures 4.5 ands 4.10); (c) realise the first accent

earlier and the second one later than normal. Similarly, in cases of tonal crowding, L*

accents are realised as rising from a low point, as in [mo'ro] in Figure 4.10.

/////////////////////// FIGURE 4.5 ABOUT HERE ///////////////////////

There are two complementary explanations for these resolutions of tonal

crowding. First, Greek favours the undershooting of all underlying tones to the

truncation of some of them (for the distinction between undershooting and truncating

languages, see Grice, D'Imperio, Savino and Avesani this volume; Ladd 1996). It

follows that L*+H, which requires at least two syllables for its canonical alignment, will

be the accent most prone to undershoot. Second, it appears that in Greek the

undershooting of L tones is preferred to the undershooting of H*s*. Some support for this

hypothesis comes from similar evidence on the undershooting of L% in Japanese

(Venditti, this volume), suggesting that different realisation constraints may apply more

generally to L and H targets. Regardless of the underlying reasons, the resolution of

tonal crowding in Greek is still not entirely understood; e.g. it is not clear whether the

strategies mentioned above are under the speaker's choice or depend on prosodic

factors, such as phrasing and the relative metrical strength of the accented syllables.

To our knowledge the role of downstep in Greek intonation has not been

investigated before. Our data, however, allow us to make certain observations. First, it is

clear that in Greek downstep is *not* triggered by the presence of bitonal accents. The

hypothesis linking downstep to bitonal accents was first advanced in Pierrehumbert

(1980) for English, and was further refined and extended in Beckman and

Pierrehumbert (1986), who took both English and Japanese data into account. This

hypothesis (which has its origins in studies of tonal phonology in African languages) is

8

often taken to reflect a universal tendency of tonal implementation (e.g. Goldsmith 1999:4). As mentioned, however, in Greek the most frequently attested pitch accent in prenuclear position is the bitonal L*+H. As most content words are accented in Greek, sentences with consecutive L*+H accents but no downstep, such as those in Figures 4.1 and 4.10, are quite common. This lack of scaling interaction among tones in Greek is further supported by similar data from the scaling of phrase accents and boundary tones, discussed at some length in sections 4.3.3 and 4.3.4 (for additional evidence against the universality of the downstep trigger, see Yip 1996 and references therein).

In addition, our corpus suggests that certain scaling differences probably reflect phonetic regularities, and thus need not be part of the phonological description of Greek intonation. We refer, in particular, to the widespread lower scaling of the nuclear accent relative to previous accents, illustrated in Figures 4.2 and 4.5. One possible reason for this type of lower scaling could be *final lowering*, i.e. the progressive lowering of overall pitch range within the last 250 ms or so of an utterance (Liberman and Pierrehumbert 1984). However, evidence like that presented in Figure 4.5—in which the nuclear L+H* accent is clearly downstepped relative to the previous pitch accents, but the following H% is fully scaled—suggests that final lowering cannot be the only reason for the observed scaling effects. Although this type of downstep is regular and does not appear to have pragmatic significance—reasons for which we assume that it has no phonological bearing—it is clear that further research is necessary before final conclusions about the role and operation of downstep in Greek can be drawn.

4.3.3 The phrase accents

There are three types of phrase accent in Greek, H-, L- and !H-. The scaling of the H- and L- phrase accents does not appear to be influenced by the identity of neighbouring

tones. This contrasts with the situation observed in other languages, such as English and German, in which the scaling of phrase accents is influenced by preceding or following tones, resulting in upsteps and downsteps (Beckman and Ayers-Elam 1997, and Pierrehumbert and Hirschberg 1990, on American English; Grice and Benzmüller 1995, on German). Because of this difference between Greek and other languages, falls or rises to mid pitch cannot be attributed to the upstepping or downstepping influence of a neighbouring tone and are thus represented by !H-.

4.3.4 The boundary tones

Greek has three types of boundary tone, H%, L% and !H%. As with the phrase accents, !H% is used to represent mid-level pitch, since phrase accents do not trigger the upstep or downstep of boundary tones. Note, for example, that both L-H% and L-!H% are attested in Greek; that is, L- cannot be seen as the trigger of the downstep of !H% in the latter case, since it does not affect the scaling of H% in the former. Furthermore, no upstep of L% boundary tones due to a preceding H tone has been attested in the GRToBI corpus (see e.g. Figure 4.4).

The three boundary tones combine with the phrase accents in eight different configurations that appear to have specific pragmatic functions. The possible phrase accent-boundary tone combinations and their typical usage are shown in Table 4.1.

/////////////////////// TABLE 4.1 ABOUT HERE ///////////////////////

The L-L% configuration is manifested as a low plateau at the end of an utterance, illustrated in Figures 4.2, 4.3, 4.9 and 4.10. The L-L% is preceded by L+H*,

H* or !H* in declaratives, by L+H* in imperatives, and negative declaratives and by L*+H in wh-questions (Arvaniti 2001; Arvaniti and Ladd ms.).

The L-H% is manifested phonetically as a dip and then a rise to a high F0 value. The L-H% is often found after a L+H* pitch accent, with the whole configuration (L+H* L-H%) suggesting an "involved" type of continuation rise, that raises the expectations of the hearer about what is going to follow. An example of this type of rise is shown in Figure 4.5. The L-H% is preceded by L* in the "suspicious" calling contour mentioned in 4.3.1.

The H-L% configuration is used in yes-no questions, in which the nuclear accent is invariably L*. The H- accent in this case, shows two distinct patterns of alignment, depending on the position of the nucleus (Arvaniti in press; Arvaniti *et al.* ms.; Baltazani and Jun 1999; Grice, Ladd, and Arvaniti 2000). Specifically, if the nucleus of the question is *not* on the final word of the utterance, the H- aligns with the stressed syllable of the final word. If the nucleus is on the final word, the H- and L% are realised at the right edge of the utterance. These two alignment patterns can be clearly observed in Figure 4.4.

The H-H% configuration is manifested as a smooth rise to a high F0, as shown in Figure 4.6. It is typically preceded by a L* in both "continuation rises" and in the questioning calling contour. The L-!H% is found in wh-questions, requesting imperatives, and negative declaratives that show reservation. All these types of utterance have similar intonational structure: a L*+H or L+H* nucleus followed by a low plateau (a spreading L-), and a small rise (the !H%). As can be seen in Figure 4.7 that illustrates this contour, the !H% remains approximately in the middle of the speaker's range.

/////////////////////// FIGURES 4.6 AND 4.7 ABOUT HERE ///////////////////////

The stylised configurations—H-!H%, !H-!H%, !H-H%—are used less often than the rest and for a limited number of pragmatic purposes. The H-!H%, illustrated in Figure 4.8, is realised as a rise to a high plateau. Our corpus suggests that it is employed mostly when the speaker wants to hold the floor while preparing his/her next utterance, but s/he is unwilling to use a (filled) pause. The !H-!H% is the mirror image of H-!H%, that is a fall to a plateau in the middle of the speaker's range; it is used in the *vocative chant* after L*+H, and also in incredulous questions after L+H*, as illustrated in Figure 4.1. Finally, the !H-H% is similar to !H-!H%. The difference between the two is the small rise at the end of the plateau in !H-H%, which makes the utterance sound more tentative or polite.

/////////////////////// FIGURE 4.8 ABOUT HERE ///////////////////////

4.4 PROSODIC STRUCTURE

In the analysis adopted here, we assume that Greek has only three prosodic constituents at and above the word: the prosodic word, the intermediate phrase and the intonational phrase. As we show below, there is ample stress, tonal and sandhi evidence in support of these three prosodic levels.

4.4.1 The Prosodic Word

A prosodic word (henceforth PrWd) consists of a content word and its clitics. The term "clitic," as used here, includes all items that in a given utterance lose their stress and form one PrWd with a host. As mentioned in section 4.2, in Greek this is the common

fate of most function words, including disyllabic ones (which are *not* traditionally considered to be phonological clitics).

A PrWd is expected to have only one stress; consequently it may bear only one pitch accent, though it is also possible for a PrWd not to be accented at all; this is, for example, the case with postnuclear PrWds (see Figure 4.7 and 4.9). PrWds with enclitic stress, on the other hand, may have two pitch accents, one on the lexically stressed syllable of the host and one on the syllable with enclitic stress (PrWds with enclitic stress may of course be de-accented; however, if they have only *one* pitch accent, this will necessarily fall on the syllable with enclitic stress).

//////////////////////// FIGURE 4.9 ABOUT HERE ////////////////////////

In addition to stress and tonal cues, PrWds are the domain of at least eight types of sandhi, presented below. Some of these—stop-voicing, /n/-deletion, /s/-voicing, vowel degemination, and vowel deletion—have been reported in Kaisse (1985), Nespor and Vogel (1986), and Condoravdi (1990), though their descriptions do not always match our data (see section 4.4.3). The other three rules emerged from our corpus.

- 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.

(3)   /tin ˈpoli/ > [tiˈboli] or [tiˈmboli]

the town ACC.

- /n/-deletion before sonorants and fricatives; e.g.

(4)   /ton laˈo/ > [tolaˈo]

the people ACC.

- /s/-voicing before sonorants; e.g.

13

(5)   /oˈjos mu/ > [oˈjozmu]

       'my son'

●   vowel degemination; e.g.

(6)   /ta ˈatoma/ > [ˈtatoma]

       the individuals

●   deletion of one of non-identical vowels; e.g.

(7)   /to ˈatomo/ > [ˈtatomo]

       the individual

●   /n/-resyllabification before a word-initial vowel (in accented syllables /n/-

       resyllabification is evident from tonal alignment); e.g.

(8)   /ˈo.tan. ˈe.fta.se/ > [o.ta.ˈne.fta.se]

       when s/he arrived

●   consonant degemination; e.g.

(9)   /oˈjos su/ > [oˈjosu]

       'your son'

●   diphthongisation of non-identical vowels; e.g.

(10)   /o. ˈi.ʎos/ > [ˈoi.ʎos]

       the sun

4.4.2 *The intermediate phrase and the intonational phrase*

The two levels of phrasing above the PrWd are 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

14

accent (H-, L- or !H-) 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.

There is abundant evidence for these two levels of phrasing in Greek. First, the tones associated with ip*s* show a simple F0 movement, such as a fall or a rise, unlike the right edges of IP*s* which often show more complex pitch configurations (see 4.3.4). 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 4.6. This observation is supported by quantitative data: after the level of phrasing had been assigned (independently of scaling) and agreed upon by the authors, a systematic comparison was made of the difference in Hz between the lowest and highest F0 point in L* H- and L* H-H% configurations in the data of four speakers reading *The North Wind and the Sun*. The results showed that this difference was smaller in L* H- than in L* H-H%, i.e. the peak was scaled lower in the former case by 30 Hz on average.

On the other hand, in IP*s* with complex final movement, such as a rise-fall, the two tones align independently (as noted, e.g., in the description of H-L% and L-!H% in section 4.3.4 and illustrated in Figures 4.4 and 4.7 respectively). This clearly shows that we are not dealing with bitonal boundary tones; if that were the case, then the individual tones would align together at the edge of the relevant phrase. Rather, in the L-!H% melody the L- spreads, while the !H% aligns with the last vowel of the utterance (Arvaniti 2001; Arvaniti and Ladd ms.); in the H-L% melody the H- aligns with a stressed vowel, if one is available, while the L% always aligns with the last vowel (Arvaniti, in press; Arvaniti *et al.* ms.).

Grice *et al.* (2000) reviewed these Greek data (as well as related data from German, Hungarian, Romanian, and English), and concluded that this behaviour of phrase accents can be accounted for if we view phrase accents as phrasal tones with a

secondary association to a specific tone bearing unit. This position does not of course account for the fact that phrase accents, in Greek at least, appear to always align at the edge of a non-final ip, and move to their secondary association site only when there is a boundary tone following. One possible reason for this difference between non-final and final ip*s* is that in the former the delimitative function of the phrase accent takes priority, while in the latter, this function is assumed by the boundary tone and thus need not be fulfilled by the phrase accent itself (for a detailed analysis along these lines, see Grice and Truckenbrodt ms.)

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 case of sandhi relates to vowel hiatus, which is resolved within ip boundaries but not across them (an observation confirmed by Pelekanou and Arvaniti in press). On the other hand, Figure 4.6 illustrates another type of sandhi, consonant degemination, which does apply across an ip boundary (though not across IP boundaries). Finally, evidence for the two levels of phrasing comes from pauses: IP*s*, even non-final ones, may be followed by a lengthy pause, while pauses are rare after ip*s* and always very short.

4.4.3 Sandhi and prosodic phrasing

Although in the prosodic analysis presented here we assume only three constituent levels, in previous analyses of the prosodic structure of Greek, additional levels have been posited. Concretely, Nespor and Vogel (1986) propose that Greek prosodic structure includes the clitic group—the need for which has been generally disputed (e.g. Zec and Inkelas 1991)—and three phrasal constituents, the phonological phrase, the intonational phrase and the phonological utterance. An additional constituent, the minimal phrase, was later proposed by Condoravdi (1990).

Although a full discussion of these analyses is beyond the scope of this paper, it should be noted that evidence for these additional constituents comes exclusively from sandhi. However, many of the sandhi phenomena used to support these analyses have not been reliably described and analysed, resulting in disagreements between the phonological descriptions (such as Kaisse 1985, and Nespor and Vogel 1986) on the one hand, and naturally occurring data (such as those of Fallon 1994) on the other.

The examination of our own corpus allows us to make the following observations regarding sandhi. First, several types of sandhi apply across larger constituents than has previously been suggested. The sequence ['ota'xtips] in Figure 4.10 is a case in point: the adverb /'otan/ 'when' loses its final /n/ before the verb /'xtipise/ 'rang', although /'otan/ and /'xtipise/ form separate PrWds (e.g. both remain stressed). According to Nespor and Vogel (1986) however, /n/-deletion before fricatives applies only within PrWd boundaries (to be precise, within the Clitic Group, which corresponds to our PrWd). This example is not an isolated instance, and cannot be attributed to fast speech, as 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 realised as such, rather than as ['maḷne] and ['tlefono] respectively, as would be expected in fast casual speech.

Second, the application of some rules presented in Kaisse (1985) and Nespor and Vogel (1986) depends on the lexical items used, something rather unusual for postlexical rules (see, e.g., Arvaniti 1991, and Malikouti-Drachman and Drachman 1992, for a discussion on the above-mentioned rule of /n/-deletion). Third, sandhi does not appear to be obligatory at any level, as Nespor and Vogel suggest about certain rules; the speaker may choose to apply a particular rule, or she may not. Finally, it

17

appears that at least some of the rules involve gradient, rather than categorical, changes. This observation was supported by Pelekanou and Arvaniti (in press), who show that gradience holds particularly true of /s/-voicing and vowel-deletion (both described in 4.4.1): in many instances of /s/-voicing, the /s/ is only partially voiced, while complete deletion of a vowel under hiatus appears to be very rare; in most cases, audible and (spectrographically) visible evidence of the "deleted" vowel remains in the signal. These findings are not surprising, as they agree entirely with results reported in studies of similar phenomena in English and other languages (among several, Holst and Nolan 1995; Zsiga 1997; Ellis and Hardcastle 1999). Nevertheless, they strongly suggest the necessity of empirically re-examining the phonological descriptions of Greek sandhi in particular, and of the reliability of sandhi as a phrasing marker in general.

4.5 THE GRTOBI ANNOTATION SYSTEM

As mentioned, GRToBI is a system for the prosodic annotation of Greek spoken corpora. The system is based on the American English ToBI (Silverman *et al.* 1992; Beckman, Hirschberg and Shattuck-Hufnagel this volume), but it has been adapted so as to take into account additional facets of Greek prosody (such as extensive sandhi) that merit particular attention. GRToBI has five 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 romanisation; the *Break Index Tier*, showing indices of cohesion; and finally a *Miscellaneous Tier*, in which other information may be entered. Details on each tier are given below (for labelling conventions see Appendix I).

4.5.1 The Tone Tier

As mentioned, the Tone Tier presents the intonational structure of the utterance, using the analysis and criteria presented in section 4.3. In addition to the pitch accents, phrase accents and boundary tones described in that section, some diacritics are also used in the GRToBI annotation system. These are largely employed to provide a more detailed description of the phonetic realisation of the pitch accents, in order to shed light on the relation between the phonological representation of accents and their context-dependent phonetics.

Concretely, L*+H pitch accents in tonal crowding contexts, in which as mentioned earlier their realisation varies, are annotated using three diacritics: wL*+H ("w" for *weak*) is used when the L tone is undershot, as in [ˈmalone] in Figure 4.10; >L*+H is used when the accent is realised *earlier* than typically expected, as in [ðaliˈða] in the same figure; and <L*+H is used when the accent is realised later than typically expected. Similarly, undershot L* accents, usually realised as a low *point* rather than a plateau, are annotated as wL* (see Figure 4.10).

/////////////////////// FIGURE 4.10 ABOUT HERE ///////////////////////

Further, the downstep diacritic (!), may be used with any of the pitch accents with a H tone, if the transcriber feels that the accent is scaled lower than declination warrants. As mentioned in section 4.3.2, it is not clear what the role of downstep is in Greek. For this reason, we have decided to explicitly annotate downstep in the Tone Tier, even in cases in which we have reason to believe that the presence of downstep is phonetically determined (as is the case with the scaling of the nuclear accents in Figures

4.2 and 4.10). Explicit marking of downstep will facilitate further research, which can illuminate the scope and function of downstep.

4.5.2 The Prosodic Words Tier

The Prosodic Words Tier provides a detailed phonetic transcription of the utterances. Currently ASCII characters (with a fairly transparent relation to their IPA equivalents) are employed. 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.

This tier was deemed 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, since Greek orthography marks stress only on polysyllabic words. In a given utterance, however, a monosyllabic content word will most likely be stressed and accented, while a disyllabic function word will most often be cliticised. By coding and examining such cases we hope that a better understanding of the relation between stress and accent in Greek will emerge.

4.5.3 The Words Tier

At present the Words Tier provides a word-by-word romanisation 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 romanisation of Greek, we have followed some of the

more widely accepted conventions (such as *ch* for *χ*), and have devised means for transliterating the rest of the characters as transparently as possible. 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.

4.5.4 Break Index Tier

GRToBI 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 also be stressed that although the use of a particular index relies on the transcriber's judgment, indices correlate with specific stress, sandhi and tonal events, which the transcriber must take into consideration before reaching a decision.

BI 0 is used to mark boundaries within 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 labelled BI 0. However, little is as yet known of the intonational behaviour of such sequences (but see Arvaniti 1992; Botinis 1998). 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*), as shown in Figure 4.8.

21

Although, as noted, several types of sandhi take place across a BI 0 boundary, its presence is not a necessary condition for BI 0 to be used (as it is in the English ToBI for instance; Beckman and Ayers-Elam 1997). 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 takes 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. The presence of an accent should be considered crucial for deciding that an item is a distinct 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; they may also be flagged with "accdCL" (for *accented clitic*) in the Tone Tier. On the other hand, it should be stressed that the absence of accent does not constitute evidence that a given stretch is *not* a PrWd; for instance, in cases of early focus in an utterance, de-accenting of all PrWds following the nucleus is expected (Baltazani and Jun 1999; Botinis 1998), as illustrated in Figures 4.7 and 4.9.

BIs 2 and 3 mark ip*s* and IP*s* respectively.[2] 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 detail in section 4.4.2.

In addition to the break indices, four diacritics are used to provide more detail on the prosodic structure of the annotated utterances: "s" for *sandhi*, "m" for *mismatch*, "p" for *pause* and "?" for *uncertainty*.

The diacritic *s* is used to flag *all* instances of sandhi at all prosodic levels, independently of whether sandhi rules operating at this level have previously been described for Greek or not. We hope that by investigating a large corpus of spoken data

thus marked, a better understanding of the relation between sandhi, phrasing and prosodic structure will be reached. Already the acoustic investigation of instances of sandhi in the GRToBI corpus (Pelekanou and Arvaniti in press) has largely confirmed the more informal observations presented in section 4.4.3, and has provided a preliminary distinction between categorical postlexical rules and rules of gradient phonetic implementation.

The *m* diacritic flags two types of mismatch between a given break index and the prosodic or tonal cues for it. The *m* 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. For example, when a sequence such as /tin ˈkori/ 'the daughter' ACC. is pronounced [tiˈgori] or [tiˈŋgori], then the boundary between /tin/ and /ˈkori/ is labelled *0s*; when the same sequence is pronounced [tin ˈkori] then it is labelled *0m*. In contrast, the sequence /i ˈkori/ 'the daughter' NOM., in which sandhi is not possible, is labelled simply *0*. The *m* diacritic is used with BI 1, 2 and 3, to mark cases in which the transcriber feels that a certain boundary is present, yet the tonal events that normally accompany it are not in place. For example, when the transcriber feels that a sequence which does not end with a phrase accent nevertheless forms a separate ip, then the boundary between this and the following ip should be labelled *2m*.

Finally, *p* is used to mark pause at a given boundary, and *?* 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; if this is not possible (i.e. if the transcriber does not find the tonal cues that normally accompany a particular break index), then *m* should also accompany the break index.

4.5.5 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, speaking rate, or pitch-halving (illustrated in Figure 4.9) are marked in this tier.

4.6 DISCUSSION AND CONCLUSION

In this paper we have presented a prosodic analysis of Greek, concentrating mostly on intonation and phrasing, but also dealing (albeit to a lesser extent) with stress and rhythm. It transpires from this analysis that the prosody of Greek is by and large understood, though certain issues remain unresolved. Among them are the phonology and phonetics of downstep, the phonetic realisation of accents under tonal crowding, and the phonological representation of the pitch accent currently analysed as H*+L.

Further, what emerges from this analysis is that the prosody of Greek has certain characteristics that merit further consideration. One of these relates to the phonological relevance and modelling of downstep. This issue has been extensively discussed in the literature (for a review, see Ladd 1996), but it is still far from being understood. This is, after all, the reason why downstep is explicitly annotated in many ToBI systems, in contrast to the theoretical works on which these systems are based (cf. Silverman et al. 1992, and Beckman and Pierrehumbert 1986 respectively). Yet, certain aspects of downstep are taken for granted, such as that downstep is triggered by bitonal accents. The Greek data clearly show that this is not a universal tendency, and point towards an analysis in which downstep is seen as an independent—rather than predictable—intonational feature (supporting Ladd's 1996 view on this issue).

24

A related issue is that of scaling influences between phrasal tones: in most intonational systems H- phrase accents upstep L% boundary tones, while L- phrase accents downstep H% boundary tones. Again, Greek does not exhibit this tendency. This means that the description of Greek intonation requires the use of a mid level, namely the tones we analyse as !H- and !H%.[3] This is problematical, given that the autosegmental/metrical framework assumes that intonational patterns can be adequately modelled using only two tones, H and L. However, if we were to limit ourselves to H and L for the description of Greek intonation, we would have to resort to an analysis in which the donwstepped patterns involved sequences of abstract L tones, which are not phonetically realised but serve to downstep others (as in Beckman and Pierrehumbert 1986). Apart from the fact that such as analysis would be highly abstract and unmotivated, at present there appears to exist strong evidence against it, since both non-downstepped L- H% and non-upstepped H-L% sequences are attested in Greek.

A third point that emerges from the Greek data is the asymmetric behaviour of phrase accents in final and non-final ip*s*. As noted, phrase accents assume their secondary association only in the former case, and one possible interpretation of this behaviour is that in non-final ip*s* phrase accents have to fulfil their delimitative function, something that is not necessary in final ip*s* (in which boundary tones assume this role). The secondary association of phrase accents is extensively discussed in Grice *et al.* (2000), but it is clear that the puzzle of the phrase accents' behaviour is far from being solved. Data from more languages and possibly from the less common melodies of the languages already studied could provide a better understanding of this issue.

A final point that is worth commenting on is that of phrasing and the prosodic hierarchy. The three-level hierarchy adopted in the GRToBI analysis is the one typically assumed in intonational phonology; at the same time it differs dramatically from the

richer hierarchy assumed in prosodic phonology. The reason for the discrepancy may lie in the evidence used: in intonational studies evidence for phrasing comes mainly from tonal patterns, while work in prosodic phonology relies more heavily on sandhi. Greek in this respect may be quite unusual in having a large number of sandhi rules and requiring few levels to account for stress and intonational patterns, thereby bringing to the fore the asymmetries between the two hierarchies. It is still uncertain whether the more elaborate prosodic structures that have been postulated in the past will turn out to be necessary. It this respect, however, it is clear that GRToBI can make a real contribution by providing natural data on sandhi, a phenomenon that is not easily amenable to laboratory testing. This holds also for other aspects of Greek prosody, such as downstep, which will certainly benefit mostly from the examination of prosodically annotated corpora like the GRToBI database.

To conclude, we hope that the prosodic analysis of Greek presented here will serve as the basis for further research, leading to the examination of the problems that emerged during the preparation of GRToBI and indicated throughout the paper, as well as to the re-evaluation of certain assumptions currently made in the cross-linguistic study of prosody.

REFERENCES

Arvaniti, A. (1991), 'The Phonetics of Modern Greek Rhythm and its Phonological Implications', Ph.D. dissertation (Cambridge University).

____ (1992), 'Secondary Stress: Evidence from Modern Greek', in G. J. Docherty and D. R. Ladd (eds.), *Papers in Laboratory Phonology II: Gesture, Segment, Prosody* (Cambridge: Cambridge University Press), 398-423.

____ (1994), 'Acoustic Features of Greek Rhythmic Structure', *Journal of Phonetics*, 22: 239-268.

____ (1997), 'Greek "Emphatic Stress": A First Approach', *Greek Linguistics 95* (*Proceedings of the 2nd International Conference on Greek Linguistics*), vol. I (Salzburg: The Department of Linguistics, University of Salzburg), 13-22.

____ (2000), 'The Acoustics of Stress in Modern Greek', *Journal of Greek Linguistics*, 1: 9-39.

____ (2001), 'The Intonation of Wh-Questions in Greek', *Studies in Greek Linguistics 21* (Thessaloniki), 57-68.

____ (in press), 'The Intonation of Yes-No Questions in Greek', *Proceedings of 14th International Symposium on Theoretical and Applied Linguistic*s (Thessaloniki).

____ and Baltazani, M. (2000), 'Greek ToBI: A System for the Annotation of Greek Speech Corpora', *Proceedings of Second International Conference on Language Resources and Evaluation* (*LREC2000*), vol. 2 (Athens), 555-562.

____ and Ladd, D. R. (1995), 'Tonal Alignment and the Representation of Accentual Targets', *Proceedings of the XIIIth International Congress of Phonetic Sciences*, vol. 4 (Stockholm), 220-223.

____ and Ladd, D. R. (ms.), 'Scaling and Alignment of Pitch Targets in Modern Greek Wh-Question Intonation', University of Cyprus and University of Edinburgh.

____ Ladd, D. R., and Mennen, I. (1998), 'Stability of Tonal Alignment: The Case of Greek Prenuclear Accents', *Journal of Phonetics*, 26: 3-25.

____ Ladd, D. R., and Mennen, I. (2000), 'What Is a Starred Tone? Evidence from Greek', in M. Broe and J. Pierrehumbert (eds.), *Papers in Laboratory Phonology V: Acquisition and the Lexicon*, (Cambridge: Cambridge University Press), 119-131.

____ Ladd, D. R., and Mennen, I. (ms.), 'Quantitative *F*0 Descriptions Require Reference to Local Targets: The Intonation of Greek Polar Questions and Emphatic Statements', University of Cyprus and University of Edinburgh.

Baltazani, M. and Jun, S. (1999), 'Focus and Topic Intonation in Greek', *Proceedings of the XIVth International Congress of Phonetic Sciences*, vol. 2 (San Francisco), 1305-1308.

Beckman, M. E. and Ayers-Elam, G. (1997), 'Guidelines for ToBI Labeling' (Columbus, OH: The Ohio State University Research Foundation).

____ and Pierrehumbert, J. B. (1986), 'Intonational Structure in Japanese and English', *Phonology Yearbook*, 3: 255-310.

____ Hirschberg, J., and Shattuck-Hufnagel, S. (this volume), 'The Original ToBI System and the Evolution of the ToBI Framework'.

Botinis, A. (1989), *Stress and Prosodic Structure in Greek* (Lund: Lund University Press).

____ (1998), 'Intonation in Greek', in D. Hirst and A. DiCristo (eds.), *Intonation Systems* (Cambridge: Cambridge University Press), 288-310.

Condoravdi, C. (1990), 'Sandhi Rules of Greek and Prosodic Theory', in S. Inkelas and D. Zec (eds.), *The Phonology-Syntax Interface* (Chicago: The University of Chicago Press), 63-84.

Dauer, R. (1983), 'Stress-Timing and Syllable-Timing Reanalysed', *Journal of Phonetics*, 11: 51-62.

____ (1987), 'Phonetic and Phonological Components of Language Rhythm', *Proceedings of the XIth International Congress of Phonetic Sciences*, vol. 5, 447-450.

Drachman, G. and Malikouti-Drachman, A. (1999), 'Greek Word Stress', in H. van der Hulst (ed.), *Word Prosodic Systems in the Languages of Europe* (Berlin and New York: Mouton de Gruyter), 897-945.

Ellis, L. and Hardcastle, W. J. (1999), 'An Instrumental Study of Alveolar to Velar Assimilation in Fast and Careful Speech', *Proceedings of the XIVth International Congress of Phonetic Sciences*, vol. 3 (San Francisco), 2425-2428.

Fallon, P. (1994), 'Naturally Occurring Hiatus in Modern Greek', in I. Philippaki-Warburton, K. Nicolaidis, and M. Sifianou (eds.), *Themes in Greek Linguistics* (London: John Benjamins Publishing Co), 217-224.

Goldsmith, J. A. (1999), *Phonological Theory: The Essential Readings* (Oxford: Blackwell Publishers).

Grice, M. and Benzmüller, R. (1995), 'Transcription of German Intonation Using Tobi-Tones – The Saarbrücken System', *Phonus*, 1: 33-51.

____ D'Imperio, M., Savino, M., and Avesani, C. (this volume), 'Towards a Strategy for Labelling Varieties of Italian'.

____ Ladd, D. R., and Arvaniti, A. (2000), 'On the Place of Phrase Accents in Intonational Phonology', *Phonology*, 17: 143-185.

____ and Truckenbrodt, H. (ms.), 'Hybrid Tones in Optimality Theory', University of the Saarland and Rutgers University.

Holst, T. and Nolan, F. J. (1995), 'The Influence of Syntactic Structure on [s] to [ʃ] Assimilation', in B. Connell and A. Arvaniti (eds.), *Phonology and Phonetic Evidence: Papers in Laboratory Phonology IV* (Cambridge: Cambridge University Press), 315-333.

Joseph, B. D. and Philippaki-Warburton, I. (1987), *Modern Greek* (London: Croom Helm).

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

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

Liberman, M. and Pierrehumbert, J. (1984), 'Intonational Invariance Under Changes in Pitch Range and Length', in M. Aronoff and R. Oehrle (eds.), *Language Sound Structure: Studies in Phonology Presented to Moris Halle by his Teacher and Students* (Cambridge, MA: The MIT Press), 157-233.

Malikouti-Drachman, A. and Drachman, G. (1981), 'Slogan Chanting and Speech Rhythm in Greek', in W. Dressler, O. Pfeiffer, and J. Rennison (eds.), *Phonologica 1980* (Innsbruck), 283-292.

____ and Drachman, G. (1992), 'Greek Clitics and Lexical Phonology', in W. U. Dressler, H. C. Luschützky, O. E. Pfeiffer, and J. R. Rennison (eds.), *Phonologica 1988* (Cambridge: Cambridge University Press), 197-206.

Nespor, M. and Vogel, I. (1986), *Prosodic Phonology* (Dordrecht: Foris).

____ and Vogel, I. (1989), 'On Clashes and Lapses', *Phonology*, 6: 69-116.

Pelekanou, T. and Arvaniti, A. (in press), 'Postlexical Rules and Gestural Overlap in a Greek Spoken Corpus', *Proceedings of the 5th International Conference on Greek Linguistics*, Paris, 13-15 September 2001.

Pierrehumbert, J. (1980), 'The Phonology and Phonetics of English Intonation', Ph.D. dissertation (MIT).

____ (1981), 'Synthesizing Intonation', *Journal of the Acoustical Society of America*, 70: 985-995.

____ and Beckman, M. (1988), *Japanese Tone Structure* (Cambridge, MA: The MIT Press).

____ and Hirschberg, J. (1990), 'The Meaning of Intonational Contours in the Interpretation of Discourse', in P. R. Cohen, J. Morgan, and M. E. Pollack (eds.), *Intentions in Communication* (Cambridge, MA: The MIT Press), 271-311.

Revithiadou, A. (1998), *Headmost Accent Wins: Head Dominance and Ideal Prosodic Form in Lexical Accent Systems* (The Hague: Holland Academic Graphics).

Setatos, M. (1974), Fonologia tis Koinis Neoellinikis [*Phonology of Standard Modern Greek*] (Athens: Papazisis).

Silverman, K. E. A., Beckman, M., Pitrelli, J. F., Ostendorf, M., Wightman, C., Price, P., Pierrehumbert, J., and Hirschberg, J. (1992), 'TOBI: A Standard for Labeling English Prosody', *Proceedings of the 1992 International Conference on Spoken Language Processing*, vol. 2 (Banff, Canada), 867-870.

Venditti, J. J. (this volume), 'The J_Tobi Model of Japanese Intonation'.

Yip, M. (1996), 'Tone in East Asian Languages', in J. A. Goldsmith (ed.), *The Handbook of Phonological Theory* (Oxford: Blackwell Publishers), 476-494.

Zec, D. and Inkelas, S. (1991), 'The Place of Clitics in the Prosodic Hierarchy', *Proceedings of WCCFL*, 10 (Stanford: SLA).

Zsiga, E. C. (1997), 'Features, Gestures, and Igbo Vowel Assimilation: An Approach to the Phonology/Phonetics Mapping', *Language*, 73: 227-274.

NOTES

[2] It her review Janet Fletcher suggested that the ip and IP BIs should be 3 and 4 respectively to bring GRToBI in line with English ToBI. We decided not to follow her suggestion, because skipping level 2 would make our annotation system less transparent. As it is, GRToBI is similar to other systems, such as Japanese ToBI (Venditti, this volume).

[3] Carlos Gussenhoven and Esther Grabe (pers. com.) have suggested we use a % marker for mid level pitch, i.e. that we analyse mid level as the absence of a H or L tone. Although this is a suggestion that should be explored further, it does not appear to alter the fact that Greek requires a phonologically contrastive mid level of pitch.

APPENDIX I: LABEL ALIGNMENT CONVENTIONS

- The labels for the L+H*, H* and H*+L pitch accents should be aligned with the highest non-spurious F0 point on 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, for which the canonical alignment of both tones is outside the accented syllable, a reliable point early in the accented vowel should be used.

- Phrase accents should be aligned with the relevant BI 2.

- Phrase accent and boundary tone combinations should be aligned with the relevant BI 3.

- The *enclA* and *accdCL* labels should be placed above or below the relevant accent in the Tone Tier.

- The transcriptions in the PrWords Tier should be aligned with the right edge of the sequence of orthographic items (presented in the Words Tier) that form one PrWd.

- Transliterated forms in the Words Tier are aligned at the (acoustic) right edge of the relevant word.

- Break indices are aligned at the (acoustic) right edge of the relevant constituent.

APPENDIX II: PHONETIC TRANSCRIPTION CONVENTIONS

IPA	ASCII	IPA	ASCII	IPA	ASCII
p	p	θ	th	ɲ / ɲ̥	N / NN
t	t	ð	D	l / ḷ	l / ll
k	k	s	s	r	r
c	c	z	z	ʎ / ʎ̥	L / LL
b / ᵐb	b / mb	ʃ	$	i	i
d / ⁿd	d / nd	ç	X	e	e
g / ᵑg	g /Ng	j	j	ɐ	a
ɟ / ⁿɟ	J / NJ	x	x	o	o
β	B	ɣ	G	u	u
f	f	m / m̥	m / mm	y	y
v	v	n / n̥	n / nn		

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

● Noticeably centralised vowels should be transcribed as @.

● Noticeably nasalised vowels should be transcribed with a following tilde; e.g. a~ for [ɐ̃].

● In cases of vowel coalescence, both vowels should be transcribed and joined by +; e.g. u+o for [o̩] resulting from a sequence of /u/ and /o/.

● Whispered vowels should be transcribed in brackets; e.g. (i) for [i̥].

● 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 capitalised; 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.)

GREEK	Romanisation	GREEK	Romanisation	GREEK	Romanisation
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 *αϊδόνι*.

● Spellings with double graphemes are transliterated in the same way; e.g. *θάλασσα* 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.

Table 4.1: Possible combinations of phrase accent and boundary tone and their usage.

Configuration	Schematic representation	Usage
L-L%		declaratives, negative declaratives, imperatives, wh-questions
L-H%		"involved" continuation rise, "suspicious" calls
H-L%		yes-no questions, requesting calling contour
H-H%		continuation rise, questioning calling contour
L-!H%		"involved" wh-questions, negative declaratives showing reservation, requesting imperatives
H-!H%		stylised continuation rise
!H-!H%		stylised call, incredulous questions
!H-H%		polite stylised call

Figure 4.1

38
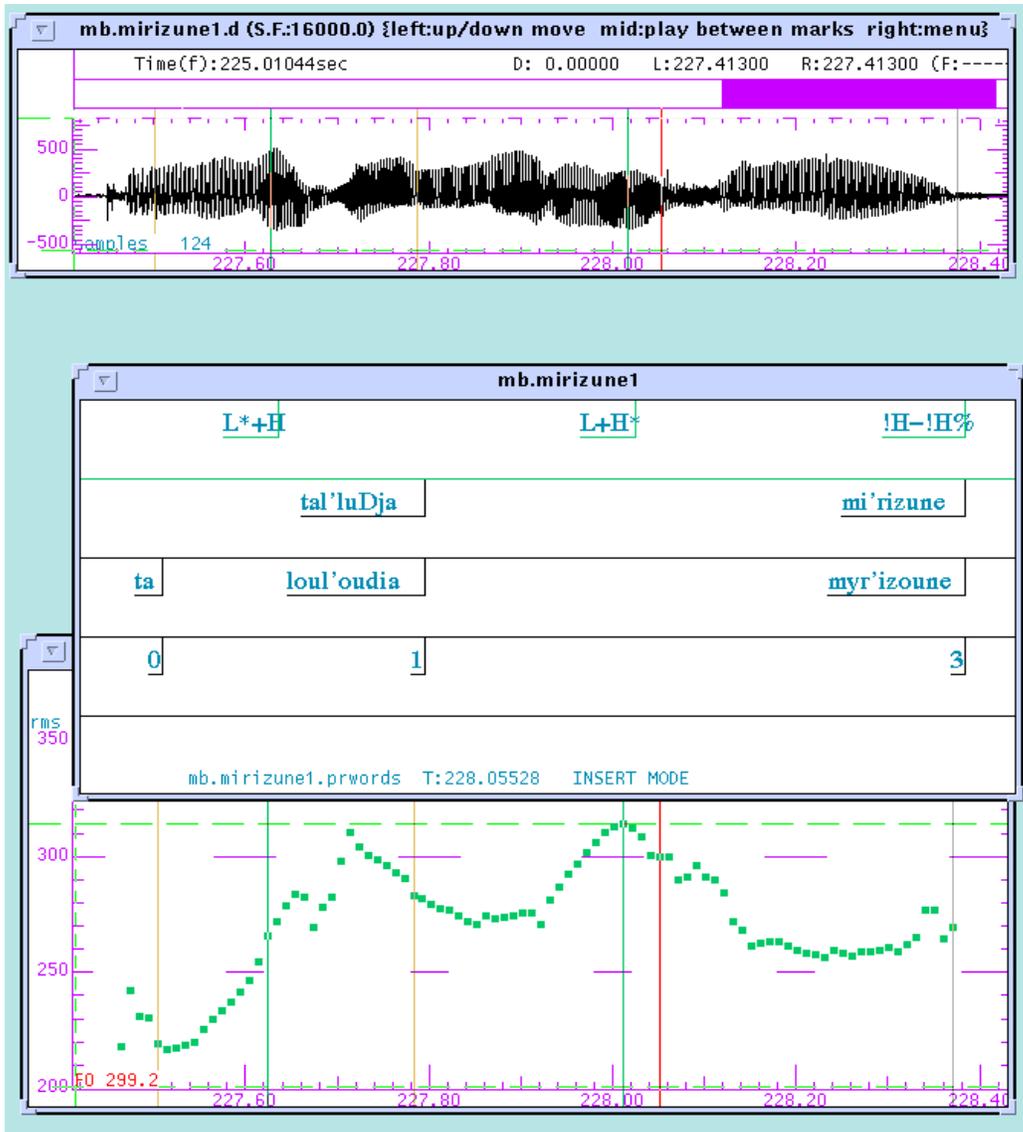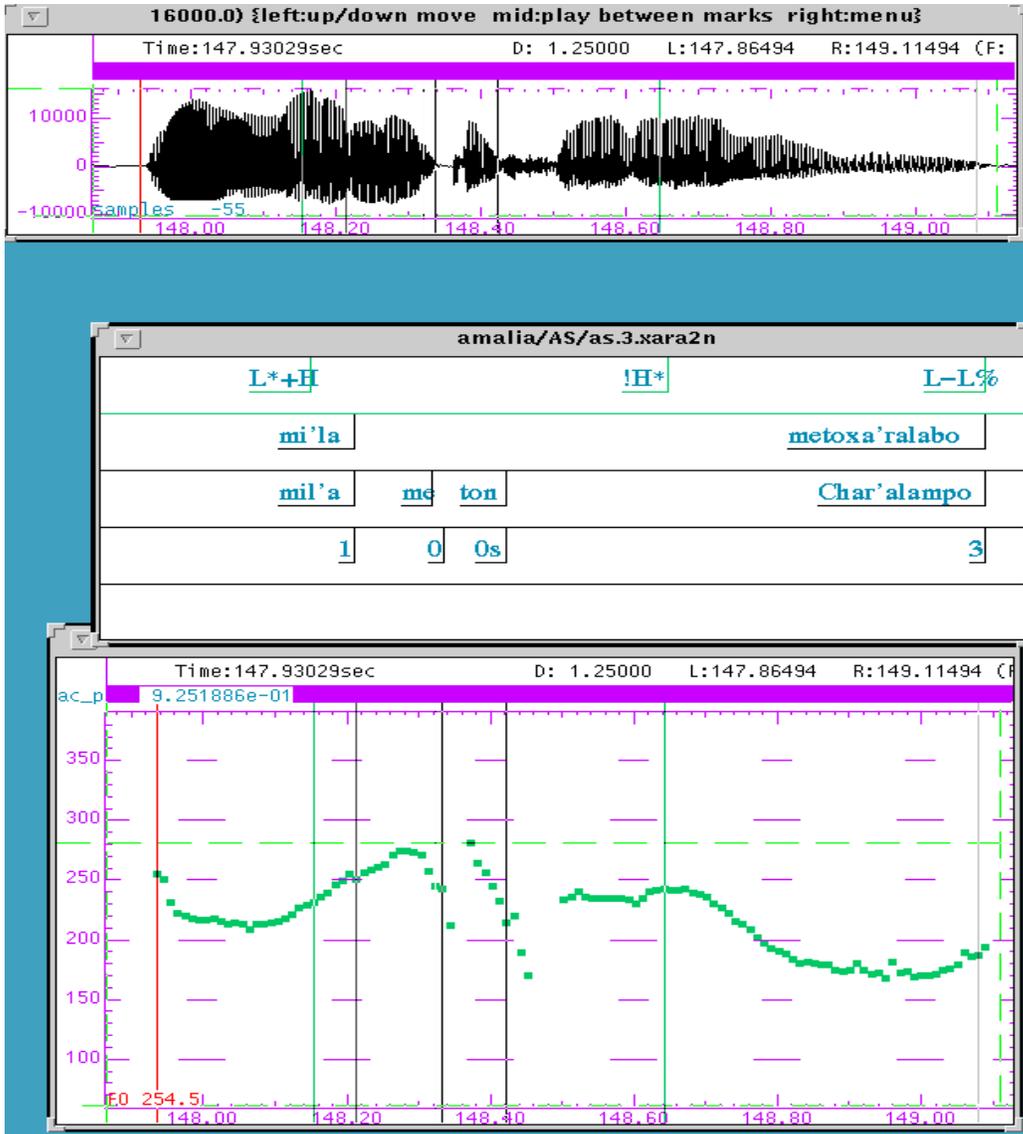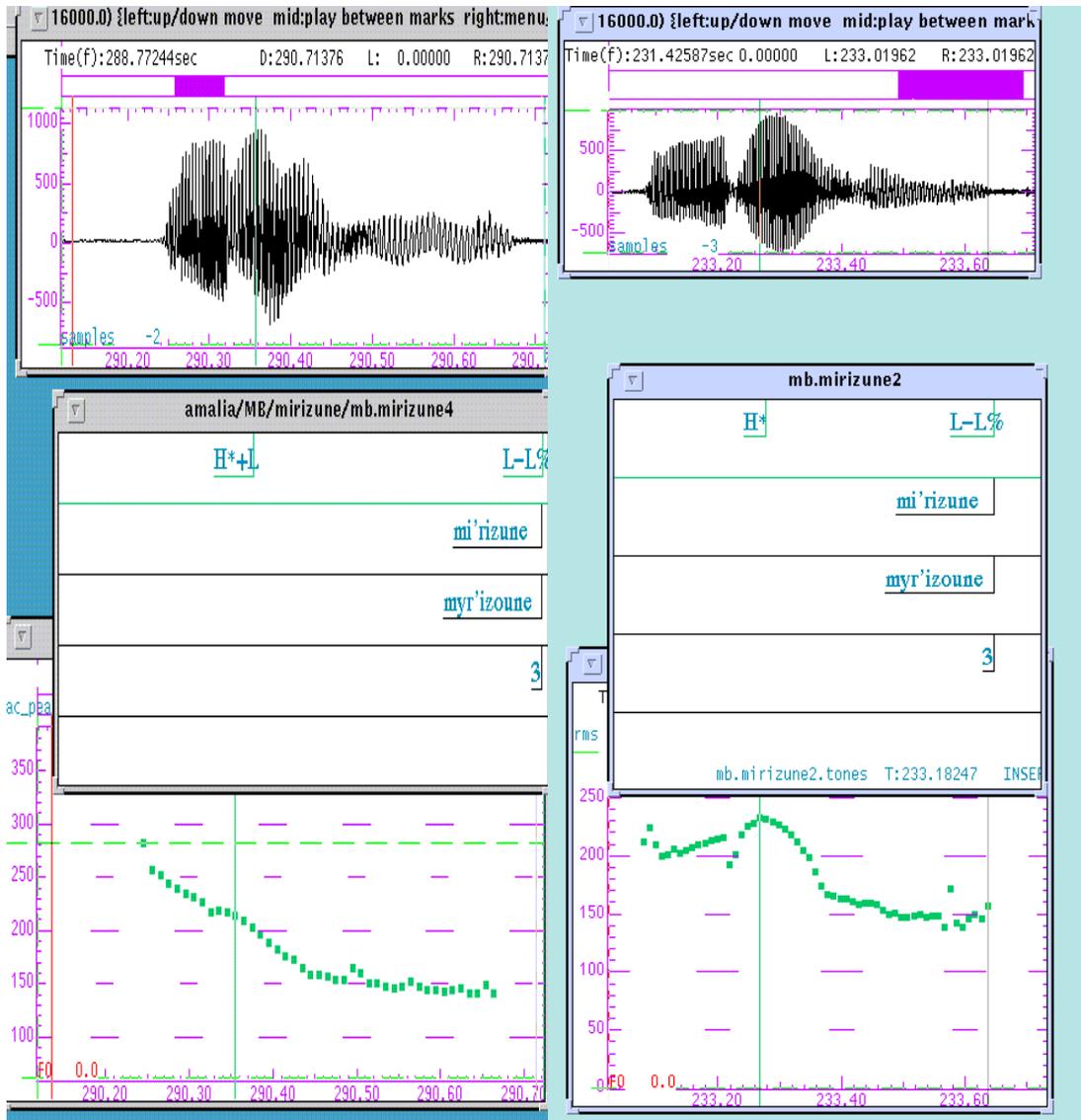
Figure 4.2

Figure 4.3

Figure 4.4

41
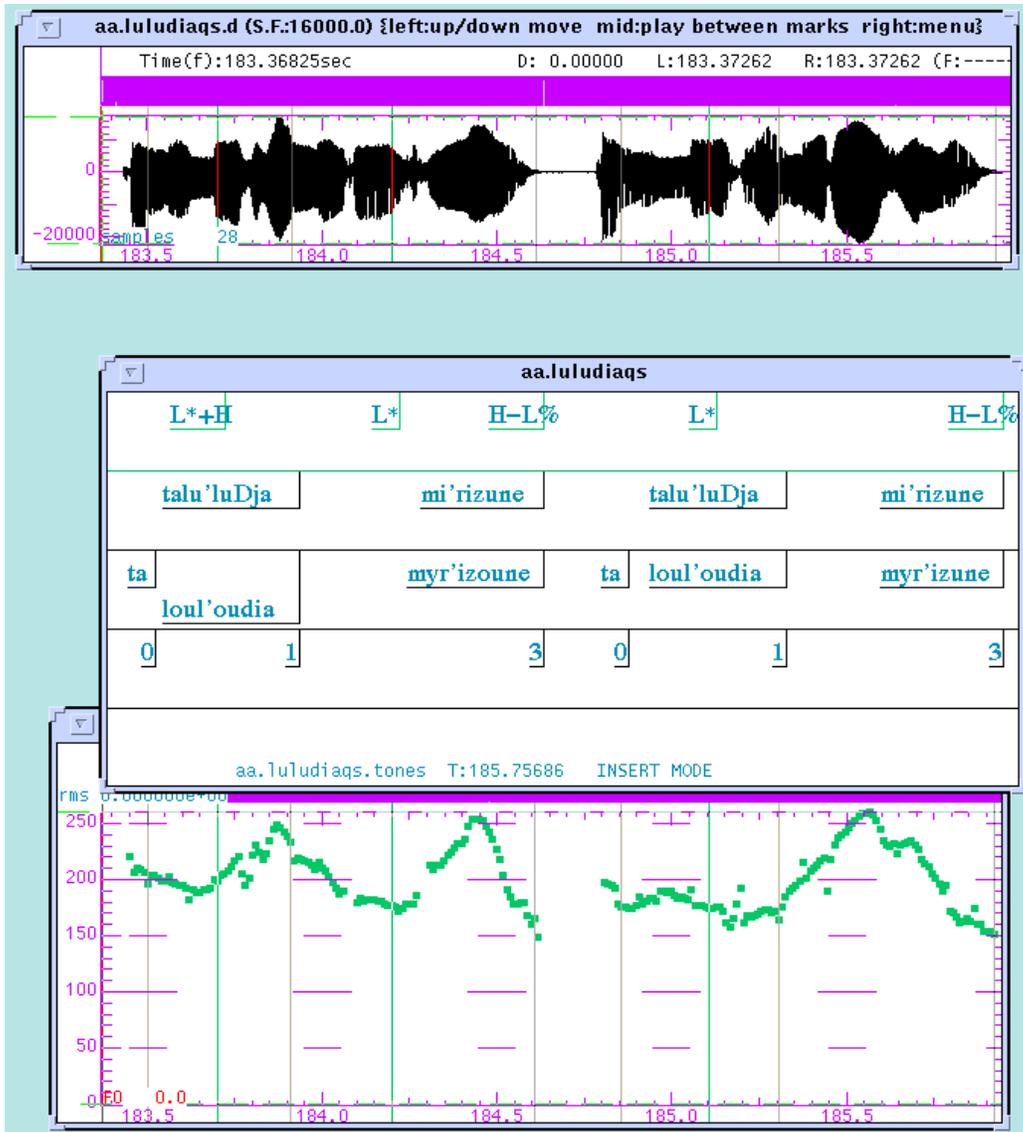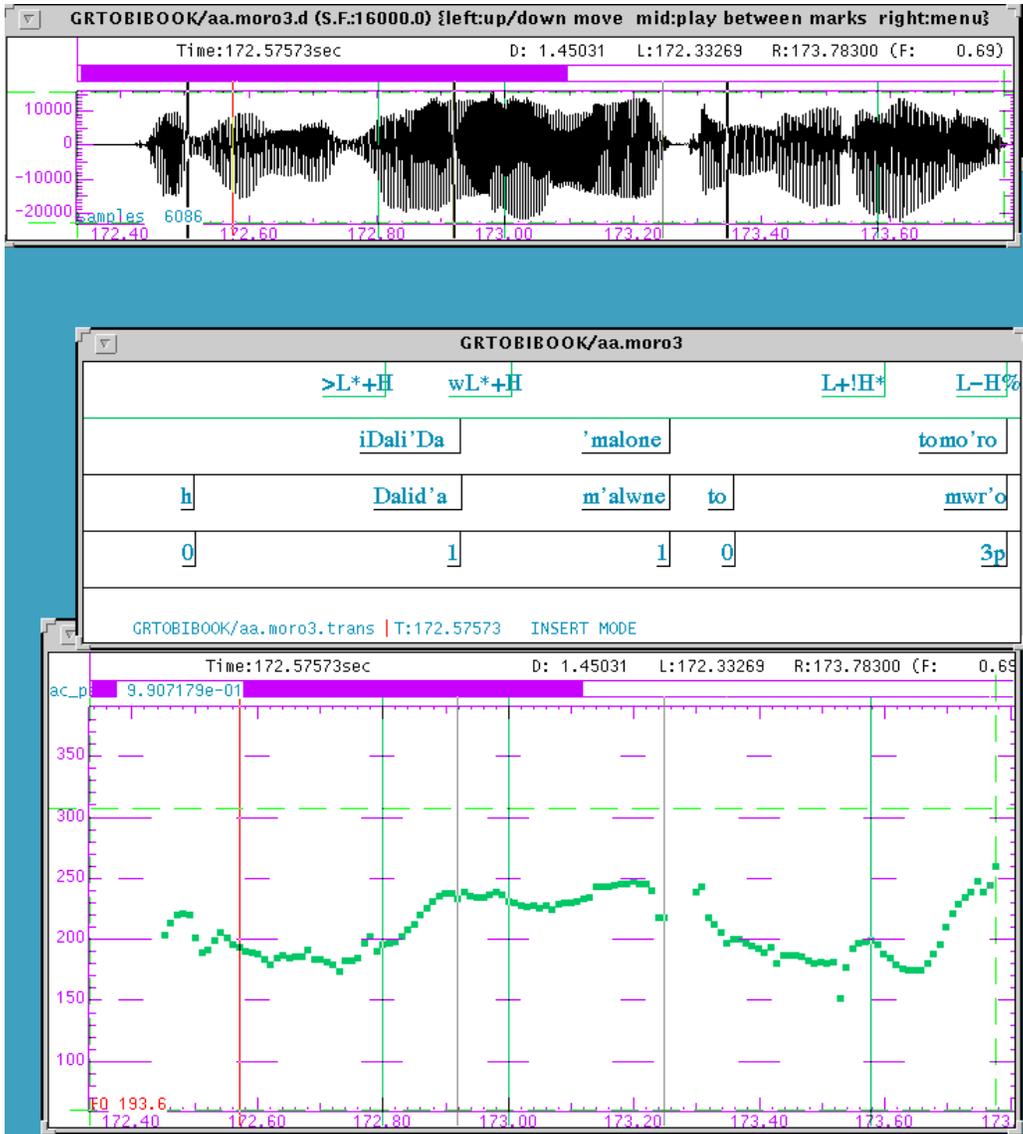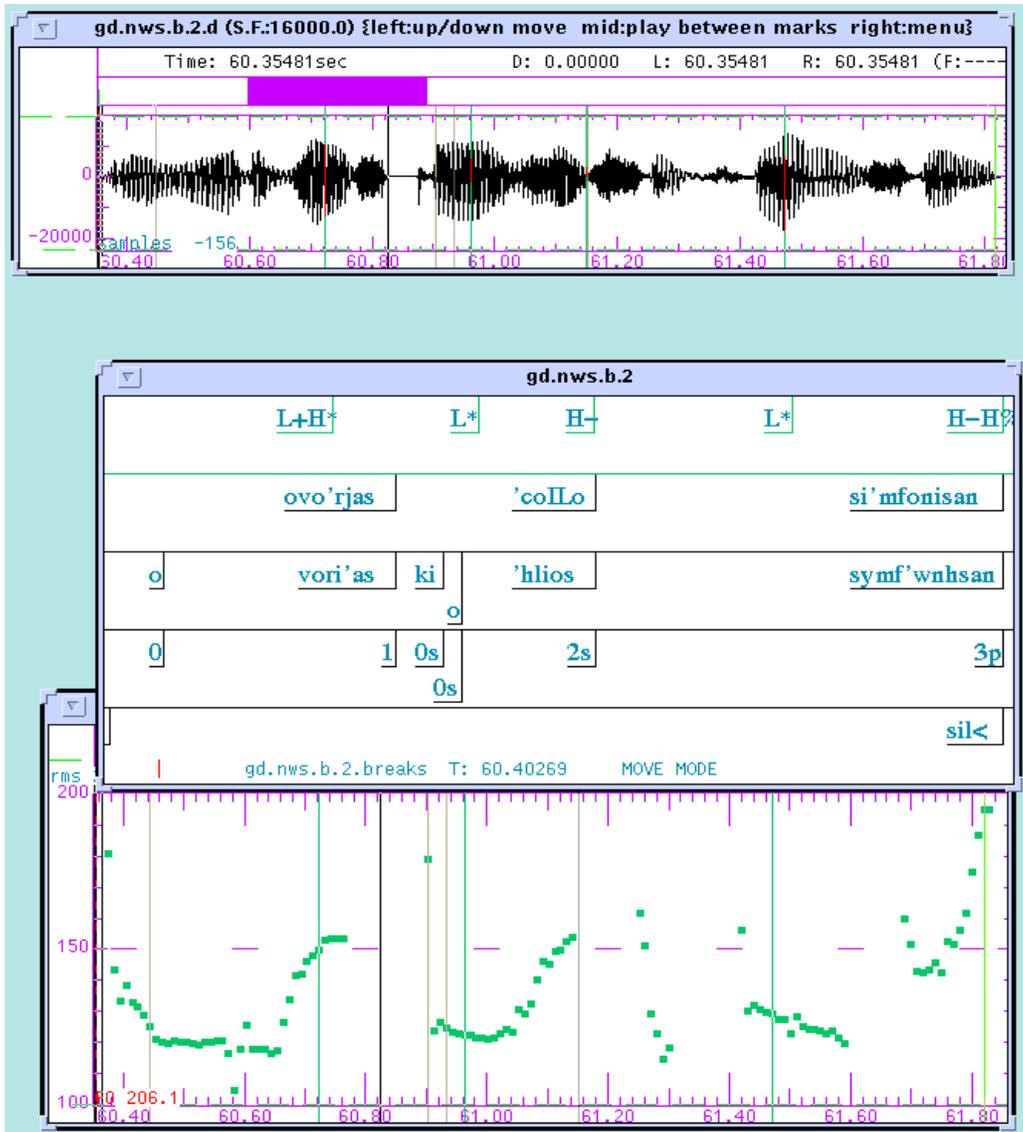
Figure 4.5

Figure 4.6

Figure 4.7

44
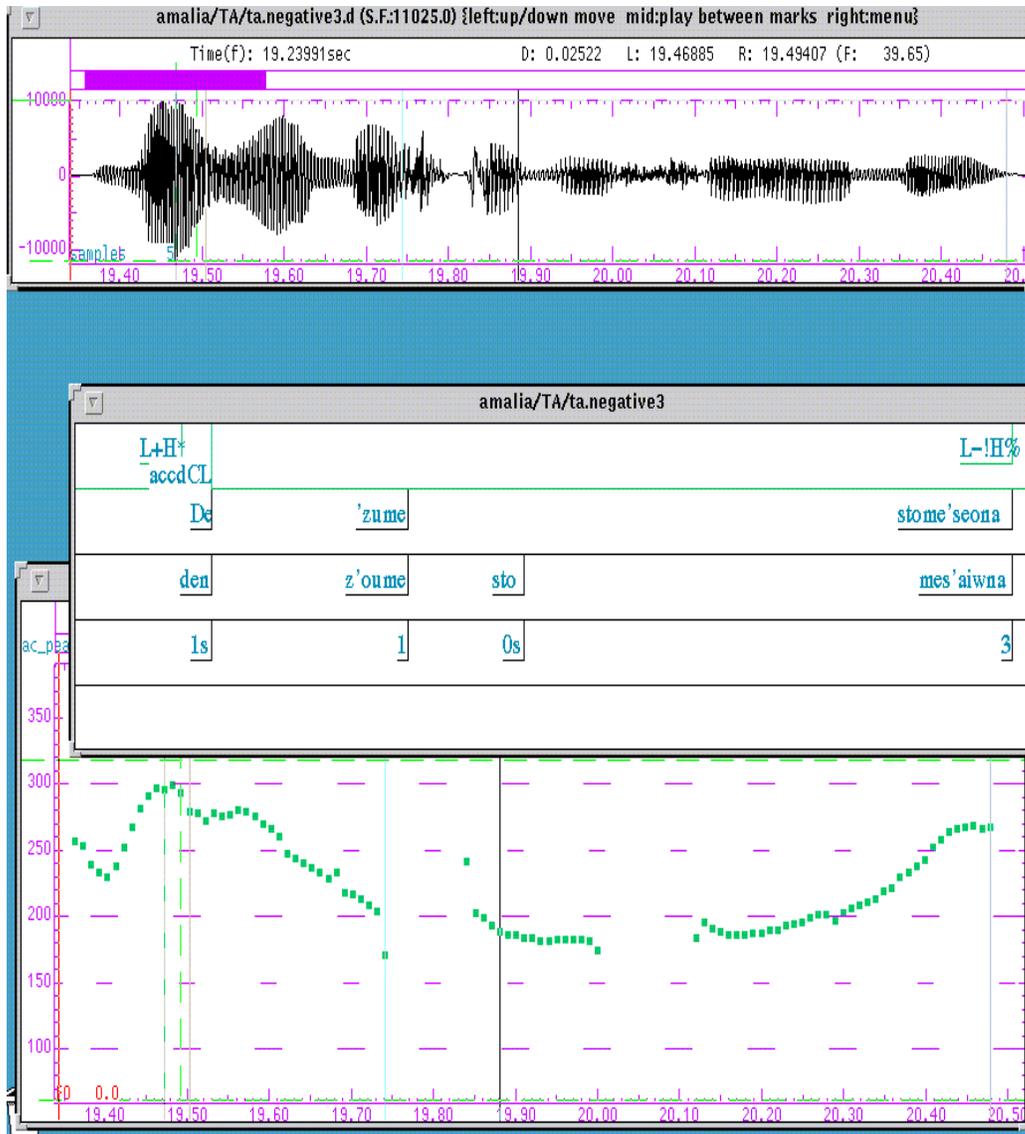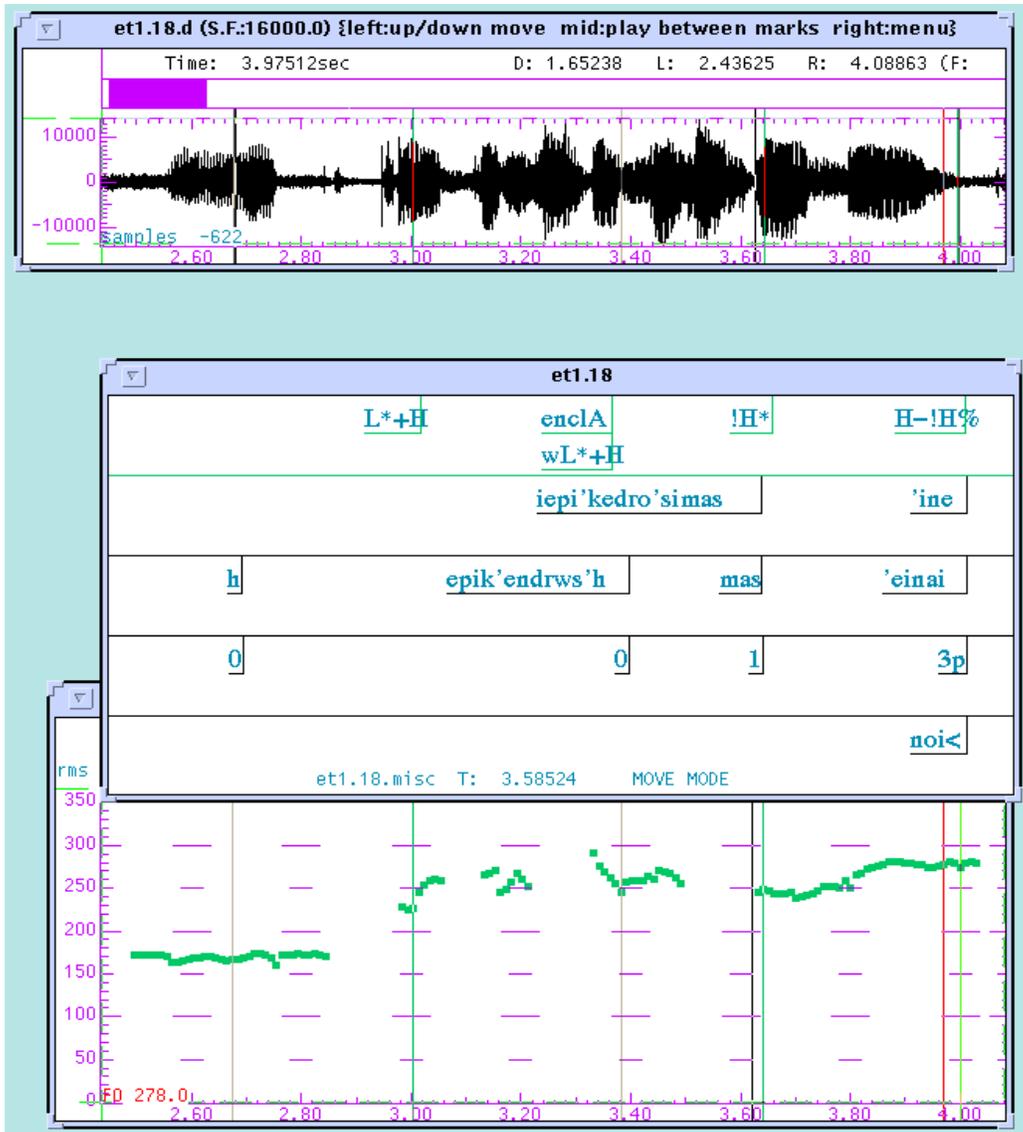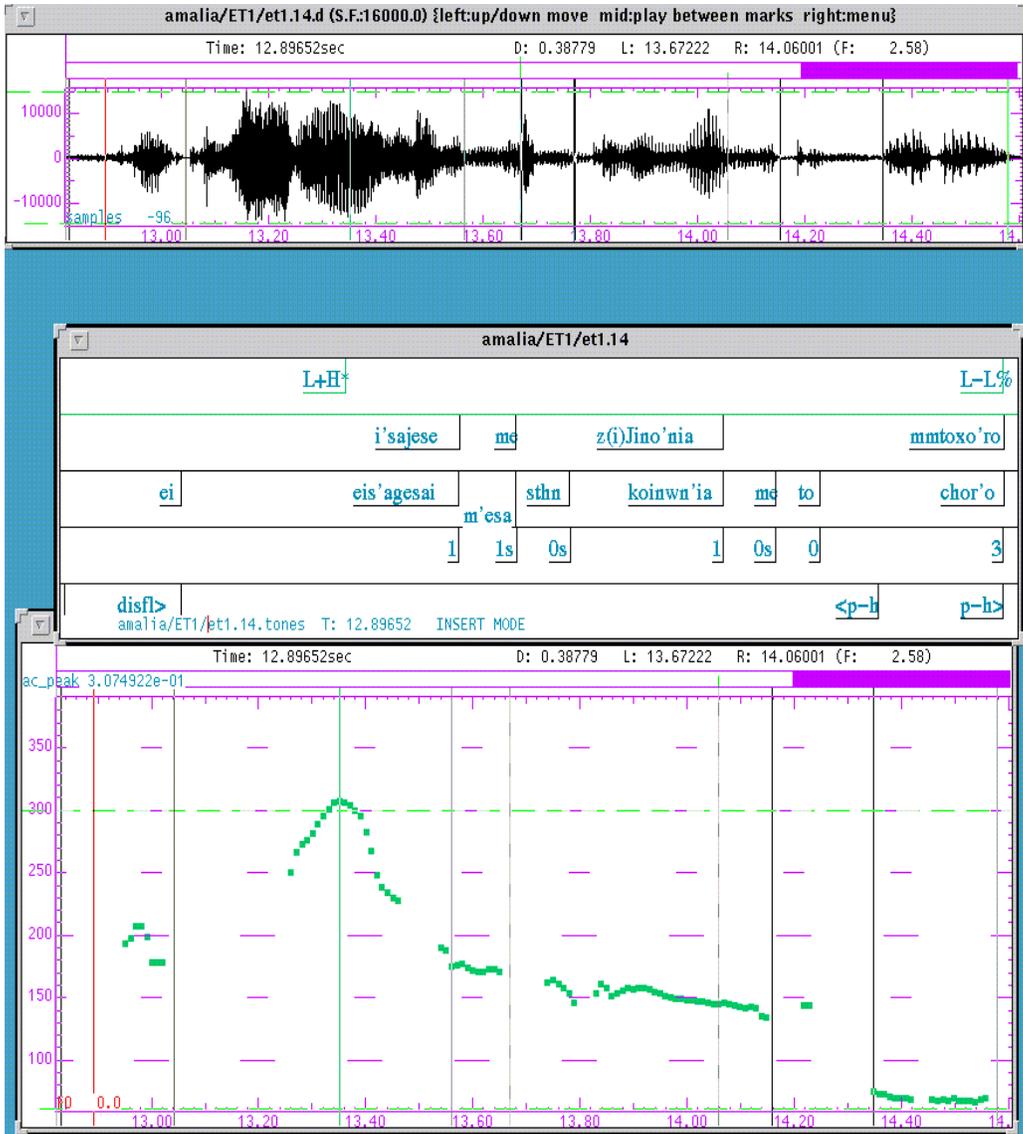
**Figure 4.8**

Figure 4.9
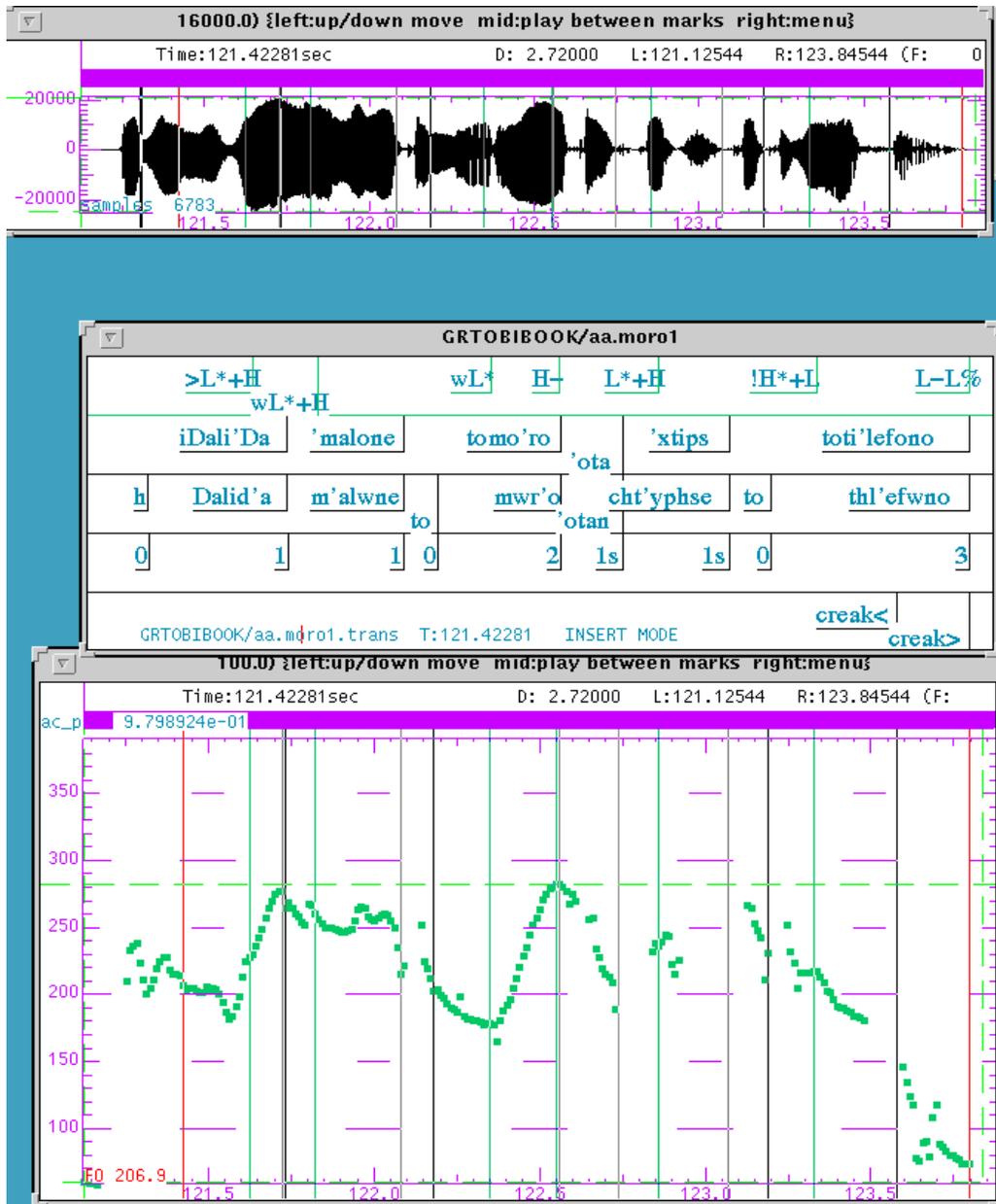
Figure 4.10

Figure 4.1: This example (the flowers smell 'Do the flowers really smell?') illustrates the different alignment of the H tone in L*+H (the accent on [l̩ˈludia]) and L+H* (the accent on [miˈrizune]). In particular it shows how the H tone is aligned with the first postaccentual vowel in the former accent, but in the middle of the accented vowel in the latter. Further this figure shows a !H-!H% phrasal configuration, realised as a plateau in the middle of the speaker's range.

Figure 4.2: This example ('S/he is talking to Charalambos') shows a typical H* nuclear accent (downstepped, in this case). Note the lack of a dip at the beginning of this accent; cf. the L+H* in Figure 4.1.

Figure 4.3: These two examples (both 'they smell') illustrate the difference between H*+L (on the left) and H* (on the right) on a one word utterance. As can be seen the H*+L is falling throughout the accented syllable [ri], while the H* accent involves a shallow rise on the accented syllable.

Figure 4.4: This illustration shows the same question twice (the flowers smell), with focus on the word [miˈrizune] 'they smell' on the left, and on the word [luˈluðia] 'flowers' on the right ('Do the flowers SMELL?' and 'Is it the FLOWERS that smell?' respectively). In both cases, the L* nucleus is realised as a low plateau with an additional dip in F0 on the accented syllable itself. Note also the different alignment of the H- in the two contours: as described in section 4.3.4, the H- aligns with the unstressed penultimate syllable [zu] in the question on the left (where the nucleus is on

the final word), but with the stressed syllable [ri] in the question on the right (where the nucleus is early).

Figure 4.5: This example ('Dalida was scolding the baby [when she fainted]!') exemplifies the L-H% phrasal configuration, which is preceded in this case by a L+!H* accent on [moˈro]. Note also the early aligned (>L*+H) and undershot (wL*+H) realisations of the two L*+H accents on the clashing [iðaliˈða] and [ˈmalone] respectively.

Figure 4.6: This example ('The north wind and the sun agreed…') illustrates the much higher scaling of H-H% relative to H-. Further the example shows the diphthongisation of /o/ and /i/—pronounced as a rising diphthong—in /ke o ˈiʎos/ 'and the sun', a phrase realised as one PrWd, [ˈcoi.ʎos], as evidenced by the alignment of the L* with the whole syllable [coi]. Finally, this example illustrates consonant degemination across ip boundaries (BI 2s): /ke o ˈiʎos_ simˈfonisan/ > [ˈcoiʎosiˈmfonisan].

Figure 4.7: This example ('We do not live in the Middle Ages!') illustrates the typical pattern of a negative declarative expressing reservation. Note that the negative particle /ðen/, which is considered a phonological clitic, carries the nuclear (and only) pitch accent of the utterance, and thus forms a separate PrWd from the de-accented verb [ˈzume] 'we live'; yet, sandhi (/n/-deletion before the fricative [z]) does take place as well, although it is said to take place only within PrWd boundaries (Nespor and Vogel 1986). The rest of the utterance is deaccented, with the L- spreading until the last syllable ([na] of [meˈseona] 'Middle Ages'). Finally, compare the scaling of the !H%

49

(relative to that of the L+H* peak) to the scaling of the H- and H% tones in Figures 4.5 and 4.6 (relative to the accentual H tones in those examples).

Figure 4.8: This example ('our focus is…') illustrates the stylised H-!H% configuration on the word ['ine] 'is'. Note also the presence of two accents on the word [epiˌcedroˈsi] 'focus', which here is followed by the enclitic [mas] 'ours', and thus carries enclitic stress on its last syllable [si].

Figure 4.9: This example ('You BECOME-PART of society through dance') illustrates de-accenting after early focus. Note also, the several instances of sandhi and fast speech rules. The phonological representation of this utterance is /iˈsaɣese ˈmesa stin kinoˈnia me to xoˈro/ and the expected realisation, according to phonological descriptions, is [iˈsaɟese ˈmesa stiɲɟinoˈnia metoxoˈro]; the actual realisation is [iˈsaɟese ˈmeziɟinoˈnia m̥toxoˈro]. Note finally the presence of pitch-halving at the end of the utterance and the use of the pitch-halving label in the Miscellaneous Tier.

Figure 4.10: This example ('Dalida was scolding the baby when the phone rang') shows two different realisations of L*+H under tonal crowding, >L*+H which is realised earlier than it canonically would (the H tone is aligned with the accented vowel, instead of the first postaccentual vowel), and wL*+H, in which the L* tone is undershot, while the H shows the typical late alignment of H in L*+H accents. In this utterance there is also an undershot L* (wL*) on [moˈro], realised as a rise from low pitch throughout the accented syllable (cf. the canonical L*s in Figure 4.6).

50

ABSTRACT

This paper provides an analysis of the prosodic and intonational structure of Greek within the autosegmental/metrical framework of intonational phonology, and presents Greek ToBI (GRToBI), a system for the annotation of Greek spoken corpora based on this analysis. Both the analysis and the annotation system have largely been developed on the basis of a corpus of spoken Greek. The analysis posits five pitch accents (H*, L*, H*+L, L*+H, L+H*), and two levels of phrasing, the intermediate phrase (ip) and the intonational phrase (IP), which are tonally demarcated by three types of phrase accent (H-, L-, !H-) and three types of boundary tone (H%, L%, !H %) respectively. The annotation system is based on the original ToBI for American English, but has been adapted to take into account facets of Greek prosody (such as extensive sandhi) that merit particular attention. Thus GRToBI has five tiers: the Tone Tier, the Words Tier, the Break Index Tier, the Miscellaneous Tier, and the Prosodic Words Tier (a phonetic transcription of prosodic words, in which stress and sandhi are also annotated for further research on prosodic constituency and its relation to sandhi and stress).

Key words or phrases: intonation, prosody, stress, Greek, ToBI, corpus annotation, spoken corpus, sandhi, intonational phonology, downstep

Headwords for index: intonation, stress, prosody, Greek, pitch accent, phrase accent, boundary tone, metrical structure, secondary stress, sandhi, juncture (put: "see sandhi" (or v.v.), downstep, prosodic structure, phrasing levels, intonational phrase, intermediate phrase

APPENDIX ?: SUMMARY OF GRToBI labels

H*	Pitch accent:
L*	Pitch accent
H*+L	Pitch accent
L+H*	Pitch accent
L*+H	Pitch accent
>	Pitch accent diacritic: marked before L*+H, it indicates that the accent aligns later than normal (usual in tonal crowding)
<	Pitch accent diacritic: marked before L*+H, it indicates that the accent aligns earlier than normal (usual in tonal crowding)
w	Pitch accent diacritic: marked before L* or L*+H it indicates that the L tone is higher than normal (usual in tonal crowding)
!	Pitch accent diacritic: marked before a H tone it indicates downstep, the lower than expected value of the H.
H-	High phrase accent
L-	Low phrase accent
!H-	Downstepped high phrase accent
H%	High boundary tone
L%	Low boundary tone
!H%	Downstepped high boundary tone
0	*Break index* showing strongest cohesion: typical of boundaries internal to prosodic words (e.g. boundaries between clitics and hosts); often, but not always, accompanied by sandhi (see diacritic s).
1	*Break index* marking a prosodic word boundary: marked on the right edge of a prosodic word followed by another prosodic word within the same intermediate phrase.
2	*Break index* marking an intermediate phrase boundary: marked on the right edge of an intermediate phrase followed by another intermediate phrase within the same intonational phrase.
3	*Break index* marking an intonational phrase boundary: marked on the right edge of an intonational phrase.
enclA	*Enclitic accent diacritic*: marked on syllables with enclitic accent.
s	*Sandhi diacritic*: indicates the presence of sandhi at any prosodic boundary
m	*Mismatch diacritic*: indicates the presence of mismatch between a break index and stress evidence for it.
p	*Pause diacritic*: indicates the presence of a pause.
?	*Uncertainty diacritic*: indicates that the labeller is unsure of XXXX