**K-ToBI (Korean ToBI) Labelling Conventions**
(version 3.1, November 2000)

Sun-Ah Jun

## 1. Background

K-ToBI (Korean TOnes and Break Indices) is a prosodic transcription convention for standard (Seoul) Korean. It is based on the design principles of the original English ToBI (see Silverman et al., 1992; Beckman & Hirschberg, 1994; Pitrelli et al., 1994), and the Japanese ToBI system (J_ToBI), devised by Jennifer Venditti (see Venditti, 1995; Campbell & Venditti, 1995). Like the other ToBI systems, therefore, K-ToBI assumes intonational phonology with a close relationship to a hierarchical model of prosodic constituents as proposed by Pierrehumbert and her colleagues (e.g., Pierrehumbert, 1980, Beckman & Pierrehumbert, 1986, Pierrehumbert & Beckman, 1988). The intonational analysis and attendant prosodic model of Seoul Korean adopted for K-ToBI are based on Jun (1990, 1993, 1996, 1998; see also Lee (1989) and de Jong (1989) for earlier studies). A first version of K-ToBI was developed at ATR Interpreting Telecommunication Systems in Japan in late 1994 by Mary Beckman and Sun-Ah Jun, as part of a Korean synthesis development project. The second version (Beckman & Jun, 1996) was an updated version modified in November 1996 by the same authors in accordance with the discussion of the Japanese/Korean working group at the Prosody Transcription Workshop held just before ICPhS (International Congress on Phonetic Sciences) in Stockholm, August 1995. The current version is a revision of the second version by Sun-Ah Jun after the Korean ToBI Workshop in Korea, August 1998. This version was presented at the workshop "Intonation: Models and ToBI Labelling", a satellite meeting of ICPhS in San Francisco in August 1999. Before introducing the revised K-ToBI labelling conventions, a brief description of the intonational structure of Seoul Korean proposed in Jun (1993, 1998) is in order.

### 1.1 Intonational structure of Seoul Korean

The intonational structure of the standard dialect (=Seoul) of Korean has two intonationally defined prosodic units: Intonation Phrase (IP) and Accentual Phrase (AP). An AP is smaller than an IP and larger than a phonological word, which is a lexical item plus a case marker or postpositions. An IP is marked by a boundary tone (%) and final lengthening. An AP is marked by a phrasal tone sequence, THLH (where T=H if the AP initial segment is aspirated or tense, and T=L otherwise), but not by final lengthening. The intonational structure of Seoul Korean is schematically represented in Figure 1.

An IP can have one or more APs, which in turn can have one or more phonological words, w. An IP is marked by a boundary tone at the end, but not the beginning, of the IP, which delivers various pragmatic meanings as well as information about the sentence type. The boundary tone is realized in the IP-final syllable, and depending on the shape of f0 contour starting from the onset of the IP-final syllable, at least nine boundary tones have been identified (L%, H%, LH%, HL%, LHL%, HLH%, HLHL%, LHLH%, LHLHL%). For example, H% and LH% differ in the timing of rising; LH% rises later than H%, showing a f0 valley at the beginning of the IP-final syllable. The same is true

with HL% vs. LHL% or HLH% vs. LHLH%. In general, tones ending with H% often have a function of seeking information (i.e. question) and those ending with L% often have a function of making a statement. However, it is often the case that tones and meaning have a many-to-many relationship. That is, more than one boundary tone can be used to mark the same meaning or sentence type, and more than one meaning is realized by the same boundary tone. For example, a wh-question can be marked by L%, H%, LH%, HL%, or HLH% (see Jun & Oh, 1996), and HL% marks both a declarative and a wh-question. More research is needed to identify distinctive pragmatic meanings for each boundary tone.

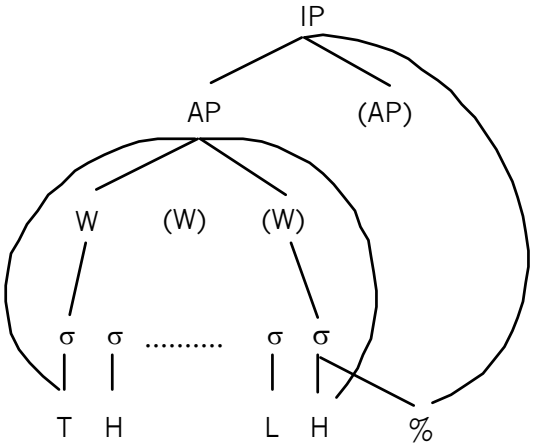



Figure 1. Intonational Structure of Seoul Korean
IP: Intonation Phrase AP: Accentual Phrase
w: phonological word σ: syllable
T= H, when the syllable initial segment is aspirated/tense, otherwise, T= L
%: Intonation phrase boundary tone

**2. Structure of K-ToBI**

The original ToBI system (i.e., English ToBI) has four parallel tiers (word, tone, break-index, and miscellaneous), but allows the free proliferation of site-specific extra tiers. Sites with aligner for English, for example, have generally added a phones tier for phonetic segmentation, and J_ToBI users have agreed to add an obligatory "finality" tier where intonational phrases that sound "final" to a discourse turn are minimally marked as such (until they can develop a more complete discourse model of discourse finality to govern a hierarchy of labels for this tier). In accordance with this general design principle, the current version of K-ToBI expands the tone tier into two tiers, a phonological tone tier and a phonetic tone tier, in order to describe surface tonal patterns which are not predictable from the underlying tones. Therefore, a K-ToBI transcription for an utterance consists minimally of a recording of the speech, an associated record of the fundamental frequency contour, and the transcription-proper symbolic labels for events on the following five parallel tiers:

1. a word tier
2. a phonological tone tier

3. a phonetic tone tier
4. a break-index tier
5. a miscellaneous tier

**2.1 Motivation of revision**

The expansion of the tone tier was devised to label the surface tonal pattern of an accentual phrase (= AP) separately from the underlying tones marking the AP boundary. This was motivated by the following four reasons. First, the ToBI labeling system assumes that tones are labeled only when they are distinctive (Beckman & Ayers, 1994, http://ling.ohio-state.edu/~tobi/). Non-distinctive pitch events that are automatically extractable from the signal should not be labeled. This is true for English ToBI. However, in Korean, distinctive pitch events do not come from an individual phrasal tone but as a set of tones forming an AP. Furthermore, though the most common tone pattern of an AP is LHLH or HHLH when the AP is longer than three syllables, an AP in Seoul Korean can be realized in at least fourteen different tonal patterns, with more variation when the AP has fewer than three syllables (i.e., LH, LHH, LLH, LHLH, HH, HLH, HHLH, LL, HL, LHL, HHL, HLL, LHLL, HHLL). Though these various patterns do not seem to differ in meaning among themselves, and though they do not seems to be predictable, it is not yet known if all these variations are indeed neither distinctive nor predictable. By labelling the surface tonal patterns, we will be able to investigate whether there is any meaning difference among these patterns.

Second, the earlier version of K-ToBI labeled only two types of tones for an AP: 'H-' marking an AP-initial H tone, when realized, and 'LHa' marking the end of an AP. When there was no initial H in an AP, H- was not labelled, conforming to the surface realization. However, in the rare event that an AP-like phrase ended in an L tone, that tone was labelled 'L%' instead of 'La' since a phrase-final L tone was found at an IP-final position most of the time and we did not want to increase the tonal inventory of APs without enough evidence. Then, in order to indicate that the AP-like boundary juncture did not match the tone pattern, a break index '2m' was placed on the break index tier: the degree of juncture was the same as that of the usual AP boundary, i.e., '2', but the tonal mark, L%, showed the boundary of an Intonation Phrase. Sometimes this was indeed the case. However, observation of more natural data revealed that there are AP boundaries which are sometimes realized with an L tone due to the tonal interaction of adjacent tones and stylistic variations. At the moment, the detailed conditions on an AP-final L tone and its pragmatic meaning are not known. We hope to get answers to these issues by labelling a falling AP boundary as 'La' on the phonetic tone tier.

By allowing 'La' to mark an AP boundary, this revised version now has a different definition of the break index '2m'. Before, it was used for a mismatch between tone and break index covering two cases: "2-like break but not AP-like tone" and "AP-like tone but not 2-like break". In the current version, a break index '2m' refers only to the former: "2-like break but not AP-like tone". "AP-like tone but not 2-like break" will be labelled in two ways depending on the degree of perceived juncture: either 1m (1-like break with AP-like tone) or 3m (3-like break with AP-like tone).

Third, the AP-initial tone in Seoul Korean is in general either L or H depending on the initial segment of an AP: H when the segment is aspirated or tense, but L otherwise. Regardless of this tonal difference on the first syllable of an AP, the second syllable of an

AP is H when the AP has more than 3 syllables. As a result, an AP can have H on the first syllable or on the second syllable or both. In the earlier version of K-ToBI, we labeled 'H-' at the first occurrence of a high-pitched syllable, either the first or second syllable or, rarely, the third syllable, without considering the origin of the H tone or the alignment of the peak to syllables. However, quantitative data show that the phonetic realization of these H tones differs depending on their origins and locations. F0 is significantly higher for the H tone on the first syllable of an AP (i.e., **H**HLH) than the H tone after the AP initial L tone (i.e., L**H**LH). In addition, this extra-high f0 value in the beginning of the HHLH pattern influences the following syllables, if there are any, by raising the f0 values of these syllables, compared to those in the LHLH pattern, up to the penultimate syllable of an AP (see Lee (1999) for more detail). Assuming that the initial L in **L**HLH or the second H in H**H**LH is predictable, we did not label these tones in the earlier version. But it turns out that these are not always predictable, and furthermore, as mentioned earlier, the individual tones forming an AP do not seem to be meaningful. That is, none of the surface tonal variations which deviate from the underlying tonal sequence seem to have a different meaning. What is meaningful in Korean intonational phonology is the phrasing, marked by the boundary tone of an AP or an IP. For example, wh-questions and yes/no-questions are distinguished only by intonational phrasing (Jun & Oh, 1996) and syntactically ambiguous sentences are disambiguated by differences in AP boundary locations (Schafer & Jun, submitted). Therefore, in this revised version, we will label the AP and IP boundaries at a phonological tone tier, and the individual AP tones at a phonetic tone tier aligned with the corresponding surface f0 event. Labelling surface tonal events on a phonetic tone tier will provide us data by which we can determine what the pragmatic meaning of these tones is, if there is any, and get information about the timing and magnitude of the f0 realization of these tones. This will provide valuable information to researchers working on speech synthesis and recognition.

Fourth, by separating the tone tier into phonological and phonetic tone tiers, we can easily accommodate tonal transcriptions of other dialects. For example, unlike Seoul Korean, the tonal pattern of an AP in the Chonnam dialect (Southwestern dialect of Korean) is LHL or HHL (Jun, 1989, 1993, 1996, 1998), with the alternation of the AP-initial tone being caused by the same principles as in Seoul Korean. Though the tonal patterns differ between the two dialects, the accentual phrasing is the same for these dialects. Thus, the boundaries marked in a phonological tone tier for Seoul Korean will remain the same for the Chonnam dialect, while a phonetic tone tier of these two dialects will differ conforming to the surface realization of each dialect. I assume this will be true for other dialects of Korean which do not have a lexical pitch accent.

In the following sections, each of the five tiers is defined, and the proper labels and symbols for each tier are introduced. In addition, example sentences illustrate in a text format how to label information on each tier, and pitch tracks of all sentences are shown in Appendix B.

## 3. Tiers

### 3.1 The word tier

The word tier in K-ToBI corresponds to the "orthographic tier" in English ToBI. In this tier, words may be labeled using either Hangul orthography or some conventional Romanization, depending on what is more convenient for the users' labeling platform or

on what is most appropriate for exporting to relevant applications. In the current K-ToBI, words are transcribed following the Romanization convention originally used at KAIST, Korea, and adopted by ATR, Japan. A table showing the mapping between Korean characters, IPA symbols, and Roman letters is given in Appendix A.

What constitutes a "word" in Korean is controversial, and we anticipate that different sites may find that the intended applications pose specific needs as to how finely an utterance should be broken up into words. For example, the intended applications at one site might require that a word label be placed for each morpheme string that has its own separate entry in some on-line dictionary. Another site may want to label a word as often as there are spaces in a standard Hangul transcript of the text. In this version, we consider 'word' as a sequence of characters separated by a space in a written Hangul text. That is, a word will be labelled at the end of each Hangul item separated by space.

If the labeling platform is xwaves and xlabel (or any similar labeling platform such as *PitchWorks* that works in terms of time flags), the word label should be placed at the end of the final segment in the word, as determined by the labeler from the waveform or spectrogram record. That is, each word should be marked at its right edge. Filled pauses and the like should also be labeled using some site-specific convention for the Hangul or Romanized spelling.

**3.2 The phonological tone tier**

A phonological tone tier will be used to mark the boundary tone of an Intonation Phrase (IP) and the boundary tone of an IP-medial Accentual Phrase (AP). Since an AP boundary tone in an IP-final position is overridden by the IP-final boundary tone, only the IP-final boundary tone (%) will be labeled at the end of an IP.

To mark the end of an IP-medial AP, we will use 'LHa' as a short term for LHLHa or HHLHa. This implies that the most common AP-final tone in Seoul Korean is a rising tone (LH). To mark the end of an IP, we will use one of the nine different boundary tones, i.e. H%, L%, HL%, LH%, HLH%, LHL%, HLHL%, LHLH%, LHLHL%. Instructions on where to put phonological tone labels are given below. To simplify the description of IP boundary tones, 'T' is used below as a variable of the IP boundary tones. The meaning of each boundary tone and sentence examples labelled with phonological tones are given in the next section.

LHa marks the end of an IP-medial AP, aligned with the end of the AP-final segment determined from the waveform. The LHa tone should be placed at or just before the corresponding break index marker regardless of the actual location of the peak.

T% marks the end of an IP, aligned with the end of the IP-final segment determined from the waveform. 'T' can be H, L, HL, LH, HLH, LHL, HLHL, LHLH or LHLHL. A T% tone at a phonological tone tier should be placed at or just before the corresponding break index marker regardless of the actual location of the peak. When a word is final to both an AP and an IP, only the IP boundary tone is written at the end of the word.

### 3.3 The phonetic tone tier

A phonetic tone tier will be used to mark the surface realization of AP tones and IP tones. For AP tones, we will have three initial tones (i.e. L, H, and +H) and three final tones (i.e. La, Ha, and L+). Among the initial tones, L and H are for the tone on the first syllable of an AP, and +H is for the tone on the second syllable (and sometimes the third when the AP is long and focused) of an AP. Among the final tones, La and Ha are for the tone on the final syllable of an AP, and L+ for the penult of an AP. Therefore, the '+' sign in Korean ToBI refers to a syllable boundary and implies a grouping of tones; +H is part of the AP-initial tone realized on the second syllable of an AP, and L+ is part of the AP-final tone realized on the penult of an AP. This is different from the '+' in English bitonal pitch accents such as L+H* or L*+H, where the starred tone is associated with a stressed syllable with the unstarred tone being realized either before (i.e., a leading L tone in L+H*), or after the starred tone (i.e., a trailing H tone in L*+H).

When an AP has three syllables, the tone on the second syllable can be either L (ex. LLH) or H (ex. LHH). In this case, we will consider the medial L as a part of the *final* AP tone and the medial H as a part of the *initial* AP tone because we believe that both are derived from the underlying LHLH pattern. That is, LLH is parsed as L-LH with the undershoot of the first H of L**H**LH, and LHH is parsed as LH-H with the undershoot of the second L of LH**L**H. Therefore, LLH will be labelled as L, L+, and Ha, and LHH will be labelled as L, +H, and Ha, on each of the three syllables. The realizations and locations of three AP-final tones and three AP-initial tones are described below.

AP-final tones:

Ha — This is the most common AP-final tone of an IP-medial AP. It can be either the end of a rising tone or a high flat tone. This label is placed aligned with an actual f0 peak on (or near if the peak is delayed or early) the AP final syllable.

La — This final tone is less common. It is sometimes seen when the following AP begins with a H tone or when it is predictable for the following AP to end with L%. This label is placed aligned with an actual f0 valley on the AP-final syllable.

L+ — This tone is not for the final syllable of an AP, but to label the low-toned penultimate syllable of an AP, either before Ha (the AP-final H) or H% (the IP-final H). Do not label this tone if it is *predictable* from adjacent tone labels, such as when an AP is continuously falling from an initial H to a final La, or when an AP-initial tone is L and the final tone is La. When not predictable, this label is placed aligned with an actual f0 valley on the penult of an AP. When there is no valley but only a low plateau after an initial H or before a final H, place this label at the beginning of the low plateau when preceded by an initial H, or at the end of the plateau when followed by a final H.

AP-initial tones:

| | |
|---|---|
| L | This tone marks an L tone on the first syllable of an AP. This label should be placed aligned with the f0 valley on the first syllable of an AP. |
| H | This tone marks a H tone on the first syllable of an AP. This label should be placed aligned with the f0 peak on the first syllable of an AP (but avoid the first few pitch points at the beginning of a vowel which is most likely due to the segmental perturbation). |
| +H | This tone marks the H tone on the second syllable of an AP (or sometimes the third syllable when the AP is long, uttered quickly, or produced under focus). This label should be placed aligned with the f0 peak around the second syllable. When the peak continues over the following syllable, place this label aligned with the latest f0 peak of the phrase-initial peak. |

Schematic f0 contours of fourteen types of AP realizations and corresponding phonetic tone labels are shown in Figure 2. The first row shows AP patterns with a high boundary, Ha, and the second row shows AP patterns with a low boundary, La. The third row shows contours of a long AP where all four underlying tones are realized with either a Ha or La boundary. 'T' in the last contour is either H or L.

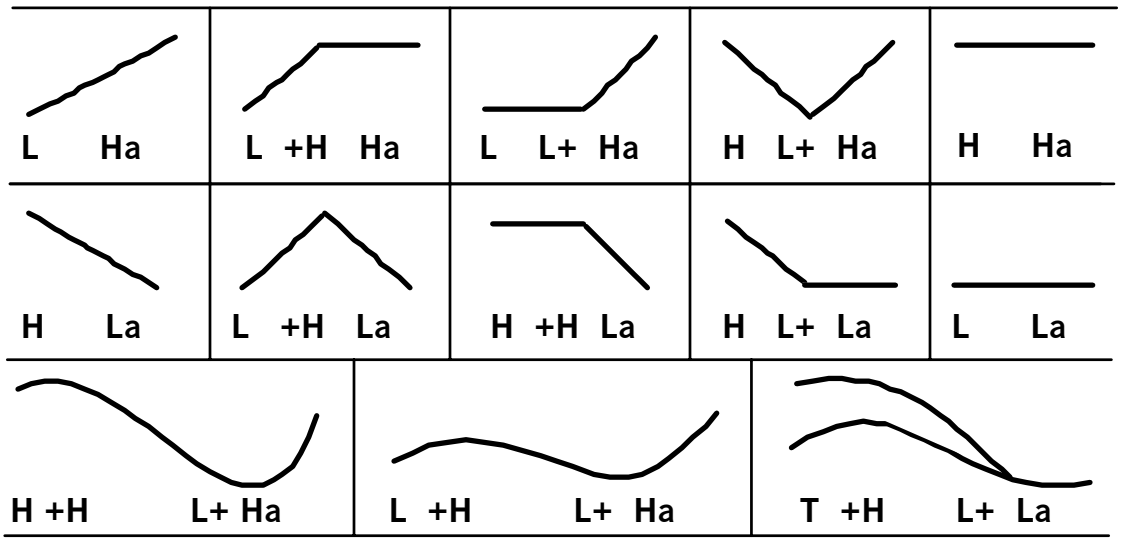



Figure 2. Schematic f0 contours of fourteen surface tonal patterns of APs.

For the IP boundary tones, the whole tone is placed toward the end of the IP-final syllable aligned with the f0 maximum for H-ending boundary tones (i.e., H%/LH%/HLH%/LHLH%) and the f0 minimum for L-ending tones (i.e., L%/HL%/ LHL%/HLHL%/LHLHL%). For complex boundary tones which include H before the last tone (e.g., **H**L%, **H**LH%, L**H**LH%, L**H**L**H**L%), the label '>' should be placed at the f0 peak corresponding to each non-final H tone. Here, '>' can mean an 'early peak' as in English ToBI (i.e. some examples of HL%; see next paragraph), but most of the time it

simply indicates the location of H so that it provides information about pitch range. At the moment, it is not clear if complex boundary tones with more than 3 tones (i.e., LHLH%, HLHL%, LHLHL%) have a distinct meaning of their own other than intensifying the meaning of the less complex tones with 2 or 3 tones (e.g., HLHL% intensifies the meaning of HL%). More K-ToBI labelled data would be needed to clarify this issue. Until then, we will label all boundary tones on the phonetic tone tier.

Currently, the type of an IP boundary tone is determined by the f0 shapes realized on the IP-final syllable. Though this is accurate most of the time, we found in news broadcasting that the H tone of HL% is sometimes realized on the penultimate syllable of an IP, possibly to keep the same rhythm across phrases. This style is also found in movies or dramas which describe the times of Old or Middle Korean, especially in the dialogues of high-class people. In addition, Park (2000) found examples where the H of HL% is realized earlier than the penult of an IP. This happened when an object was postposed after a verb whose boundary tone in the original sentence was HL%. This is one of the three possible ways of 'afterthought' realization in Korean: 1) both the verb-final syllable and the postposed object-final syllable carry the HL% tone, 2) the verb and the object form one IP, and the object-final syllable carries the HL% tone, and 3) the verb and the object form one IP, but the HL% tone is split so that the H tone is realized on the verb-final syllable and the L tone is realized on the object-final syllable. The third possibility is when the part of a boundary tone is realized before the IP-final syllable. In this case, the label '>' should be placed at the f0 peak of the verb-final syllable. So far, this type of split boundary tone has been found only for HL%. More data are needed to see if this is possible for other boundary tones.

The following shows surface realization rules of each boundary tone, and its location relative to words and f0 contours.

<u>IP-final boundary tones:</u>

L% : A level ending, or a gently falling boundary tone spread over much of the IP-final AP from the f0 peak at the beginning of the AP. This tone should be placed at the end of the phrase, aligned with the minimum f0 value. This tone is the most common in stating facts, and in declaratives in reading.

H%: A rising boundary tone that begins to rise before the IP-final syllable, and reaches its peak during the final syllable. Therefore, the rise is earlier than that in LH%. This tone should be placed at the end of the phrase, aligned with the maximum f0 value. This tone is the most common in seeking information as in yes/no-questions.

LH%: A rising boundary tone that is more localized than H%, rising sharply from a valley well within the final syllable. That is, by comparison to H%, this is a sharper later rise, starting after the onset of the final syllable. This tone should be placed at the end of the phrase, aligned with the maximum f0 value. This is commonly used for questions, continuation rises, and explanatory endings. It is also used to signal annoyance, irritation or disbelief (e.g.,

<gIrEtaniKa gIrEne!> ‘I have already told you so. (Why do you keep asking me?)’ or <bEryESE!> ‘(Did you) throw it out? (I can’t believe that!)’).

HL%: A falling boundary tone that rises to a peak *before* the last syllable, and then falls during the last syllable. Though it seems to be a combination of H% and L%, the H part of this boundary tone is not as high as a simple H% and the L is not as low as a simple L%. This tone should be placed at the end of the phrase, aligned with the minimum f0 value, and the location of H should be marked by ‘>’ aligned with the f0 peak. This tone is most common in declaratives and wh-questions. It is also commonly used in news broadcasting.

LHL%: A rising-falling boundary tone that, unlike HL%, rises *within* the IP-final syllable — essentially a combination tone consisting of LH% followed by L%, but the f0 peak is not as high as that of LH%. This tone should be placed at the end of the phrase, aligned with the minimum f0 value, and the location of H should be marked by ‘>’ above the f0 peak. It sometimes intensifies the meaning of HL%, but like LH%, it also delivers the meanings of being persuasive, insisting, and confirmative. It is also used to show annoyance or irritation. (e.g., <hazima>! ‘Don’t do it (I told you before)’)

HLH%: A falling-rising boundary tone — a combination of HL% and H%. That is, the timing of the rise is the same as HL% but followed by a shallow dip and then another rise. This tone should be placed at the end of the phrase, aligned with the maximum f0 value. The location of the first H should be marked by ‘>’ above the f0 peak. The tone is not as common as the other types mentioned so far, and some speakers use this type more often than others. This tone is used when a speaker is confident and expecting listeners’ agreement.

LHLH% A rising-falling-rising boundary tone. The timing of the rise is like LH%. This tone should be placed at the end of the phrase, aligned with the maximum f0 value. The location of the first H should be marked by ‘>’ above the f0 peak. This tone is less common than others, and has a meaning of intensifying some of the LH%’s meanings, i.e., annoyance, irritation or disbelief.

HLHL% A falling-rising-falling boundary tone. The timing of the rise is like HL%. This tone should be placed at the end of the phrase, aligned with the minimum f0 value. The location of the two Hs should be marked by ‘>’ above the f0 peak. This tone is more common than LHLH%, but not as common as single, bi- or tritonal boundary tones. It sometimes intensifies the meaning of HL%, confirming and insisting on one’s opinion, and sometimes, like LHL%, it delivers nagging or persuading meanings.

LHLHL% A rising-falling-rising-falling boundary tone. The timing of the rise is like LH% followed by LHL%. This tone should be placed at the end of the phrase, aligned with the minimum f0 value. The location of the two Hs should be marked by '>' above the f0 peak. This tone is rare and its meaning is similar to that of LHL%, but has a more intense meaning of being annoyed.

Schematic f0 contours of eight types of IP boundary tone realizations are shown in Figure 3. The first row shows IP boundaries ending with L% and the second row shows those ending with H%. The vertical line shown in each contour marks the beginning of the IP-final syllable. The f0 scale is not normalized.

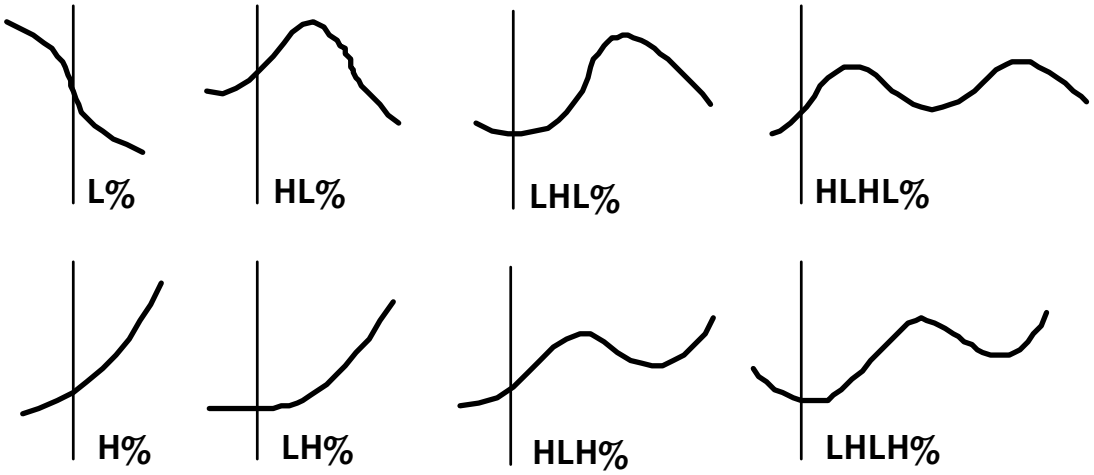



Figure 3. Schematic f0 contours of eight boundary tones of IP.

Finally, for a case of uncertain or underspecified tonal events, for both AP and IP, use the following labels on the phonetic tone tier. Underspecified tone labels should be used when a labeler knows there is a tone, but has not assigned a label yet.

| | |
|---|---|
| X | Underspecified tonal event of a non-AP-final tone. (Tone is there, but the tonal value has yet to be assigned) |
| a | Underspecified AP-final tone |
| % | Underspecified IP-final tone |
| X? | Uncertain of the type of a tone which is neither an AP-final nor IP-final boundary tone. (The labeler is not sure of the tone type.) |
| Xa? | Uncertain of the type of an AP-final boundary tone. |
| X%? | Uncertain of the type of an IP-final boundary tone. |

Example sentences labelled with a phonological tone tier and a phonetic tone tier are shown below. File names are in "<< >>" and example sentences are given in a Romanization of the Korean alphabet (see Appendix A). F0 tracks of each example with corresponding labels are shown in Appendix B. "-early", "-middle", or "-late" indicates a region of the sound file.

Examples of tone labelling on both the phonological tone tier and the phonetic tone tier:

Ex.1. <<boundary-H%>> gIrASEyo
phonological tone tier H%
phonetic tone tier +H L+ H%
-> 'Is that so?'

Ex.2. <<boundary-LH%>> gIrASEyo
phonological tone tier LH%
phonetic tone tier +H LH%
-> 'Is that so?'

Ex.3. <<boundary-HL%>> gIrASEyo
phonological tone tier HL%
phonetic tone tier L+H L+ > HL%
-> 'Is that so?'

Ex.4. <<boundary-LHL%>> gIrASEyo
phonological tone tier LHL%
phonetic tone tier L+H > LHL%
-> 'Is that so?'

Ex.5. <<J3A2-HLH% >> onIR zEnyEge nuga mEgEyo
phonological tone tier LHa HLH%
phonetic tone tier L L+Ha L+H L+ > HLH%
'Today night who eat?'
-> 'Who is eating tonight?'

Ex.6. <<IPboundary-HL%>> baraMgwa hANnimi
phonological tone tier LHa HL%
phonetic tone tier L Ha H L+ >HL%
'The North Wind and the Sun-NOM'
-> 'The North Wind and the Sun .....'

Ex.7. <<IPboundary-LH%>> dubENCA,
phonological tone tier LH%
phonetic tone tier L +H LH%
-> 'Second,'

Ex.8. <<2syllAP-LHa>> nanIN yEQarIR miwEhAyo
phonological tone tier LHa LHa L%
phonetic tone tier L Ha L L+Ha L+H L+L%

‘I-TOP Younga-ACC hate’
-> ‘I hate Younga’

Ex.9. <<5syllAP-LHLHa>> yEQmaNinenIN yEQarIR miwEhAyo
phonological tone tier LHa LHa L%
phonetic tone tier L +H L+Ha L L+ Ha L +H L+ L%
‘Youngman’s family-TOP Younga-ACC hate’
-> ‘Youngman’s family hates Younga’

Ex.10. <<6syllAP-LHLHa>> yEQi EmEninIN yEQarIR miwEhAyo
phonological tone tier LHa LHa L%
phonetic tone tier L+H L+ Ha L L+ Ha L +H L+ L%
‘Youngi’s mom-TOP Younga-ACC hate’
-> ‘Youngi’s mom hates Younga’

Ex.11. <<5syllAP-HHLHa>> hyEQmininenIN yEQarIR miwEhAyo
phonological tone tier LHa LHa L%
phonetic tone tier H +H L+ Ha L Ha L +H L+ L%
‘Hyungmin’s family-TOP Younga-ACC hate’
-> ‘Hyungmin’s family hates Younga’

Ex.12. <<t1p1s2>>-early doQgi buyEU du hyEQtA zuQesE ...
phonological tone tier LHa LHa LHa L%
phonetic tone tier L Ha L L+Ha L Ha H +H L+ L%
‘motivation providing-POSS two types among ...’
-> ‘Among the two types which provide motivation,’

Ex. 13. <<t1p2s8-1m>> sEQzaQhago iNnIN gEsi saraiNnIN gEsida
phonological tone tier LHa LHa L%
phonetic tone tier H+H Ha L L+Ha H+H L+ L%
‘to grow-prog. rel.cl. marker-thing-NOM to live-prog.’
-> ‘Being growing means that it is alive’

Ex. 14. <<gazEQgyosa>> nanIN siRryEGiNnIN zibaNU gazEQgyosarIR maNnaTa.
phonological tone tier: LHa LHa LHa LHa L%
phonetic tone tier: L Ha H +H Ha L L+Ha L+H L+Ha L L%
‘I-TOP powerful family’s tutor-ACC. met’
-> ‘I met the tutor of a powerful family’

**3.4 The break index tier**

Break indices represent the degree of juncture perceived between each pair of words and between the final word and the silence at the end of the utterance. They are to be marked after all words that have been transcribed in the word tier. All junctures — including those after fragments and filled pauses — must be assigned an explicit break index value; there is no default juncture type.

Values for the break index are chosen from the following set:

0 For cases of clear phonetic marks of "clitic" groups; e.g. application of vowel coalescence rules. Also for cases of 'incomplete nouns', monosyllabic nouns which are, though separated by spaces, not used by themselves but need a modifier (e.g. <su> 'way', <de> 'place', <gED> 'thing').

1 For phrase-internal "word" boundaries which are not marked by such cliticization phenomena and can be pronounced by itself.

2 For cases of a minimal phrasal disjuncture, with no strong subjective sense of pause — that is, a sense of phrase edge of the type that is typically associated with the tonal pattern at the right edge of the Accentual Phrase.

3 For cases of a strong phrasal disjuncture, with a strong subjective sense of pause (whether it be an objective visible pause or only the "virtual pause" cued by final lengthening) — that is, a sense of phrase break of the type that is typically associated with the tonal pattern at the right edge of an Intonation Phrase.

Note that while the Accentual Phrase and Intonation Phrase are defined in the prosodic model by tonal markings, the break index value indicates the labeler's subjective sense of disjuncture and not simply the juncture that typifies the apparent tones. Thus, the break index tier markings are not made completely redundant by the tone tier markings for break index levels 2 and 3. In cases of mismatch, the break index number should follow the perceived juncture rather than the tones, and it should be flagged with the diacritic "m", as in:

1m A disjuncture that typically would correspond to a phrase medial word boundary, but is marked by the tonal pattern of an AP.

2m A medium strength disjuncture that typically would be marked by the tonal pattern of the AP, but has no tonal markings, or has the tonal markings of an IP edge.

3m A highest strength disjuncture that typically would be marked by the tonal pattern of the IP, but has the tonal markings of an AP.

In an xwaves/xlabel-type system or any system which allows time-aligned labels, the break index label should be aligned with a point in time at the end of each word, as indicated in the word tier. It should be located exactly at, or slightly to the right of, this word marker, so that break indices can be unambiguously associated with other tiers. Transcriber uncertainty about break-index strength is to be indicated with a minus ("-") diacritic affixed directly to the right of the higher break index — e.g. "1-" to indicate uncertainty between "0" and "1"; "2-" to indicate uncertainty between "1" and "2"; and

so on. Note that since the "m" diacritic suggests certainty about the break index analysis in the face of conflicting tonal evidence, the "-" diacritic should not be used together with "m".

For a case of uncertain or underspecified break index labels, use the following labels on the break index tier.

| | |
|---|---|
| x | Underspecified break index |
| #- | Break uncertain between # and #-1 level (ex. 2-: not sure of 2 or 1) |
| #p | Pause or disfluency after this level of juncture; 1p for abrupt cutoffs after or in the middle of a word; 2p for prolongation of an AP-final syllable, but not meant to be IP final. |

Example sentences with break indices:

Ex.12. <<t1p1s2>>-early doQgi buyEU du hyEQtA zuQesE ...
Break index 2 2 2- 1 3-
'motivation providing-POSS two types among ...'
-> 'Among the two types which provide motivation,'

Ex. 13 <<t1p2s8-1m>> sEQzaQhago iNnIN gEsi saraiNnIN gEsida
Break index 1m 2 1 3
'to grow-prog. rel.cl.-thing-NOM to live-prog.'
-> 'Being growing means that it is alive'

Ex. 14. <<gazEQgyosa>> nanIN siRryEGiNnIN zibaNU gazEQgyosarIR maNnaTa.
Break index 2 1m 2 2- 3
'I-TOP powerful family's tutor-ACC. met'
-> 'I met the tutor of a powerful family'

Ex. 15. <<t1p1s2>>-late iRbaNzEgiN gEsIn waNzENhwa,
1 3 3
'general-rel thing-TOP completeness'
->(Among the two types which provide motivation,) what's in common is completeness'

Ex 16. <<break-L8C3>> azumEniga ENze maNdIrEyo?
Break index 2 1 3
'madam-NOM when make-Q'
-> 'When is Madam making (it)?'

Ex. 17. <<t1p2s6>> zIG, saNhonIN saraiSImyE aMsEgIN zugEiNnIN gEsida
Break index 3 2 3 2- 1 3
'That is, coral-TOP alive and rock-TOP dead-progressive rel.marker to be'
-> 'That is, coral is alive and a rock is dead'

Ex. 18. <<t1p2s10>> igEsIN uridIR maIMU segyeedo hAdaQdweNda.
Break index 3- 2 2 2 3
'This our mind world too apply to'
-> 'This also applies to our mind'

Ex. 19. <<t1p2s5>>-early
gIrEna, gatIN hyENmigyEQe saNho zogagIR noko bomyEN
Break index 3- 2 3 1 2- 1 3
'but, same microscope-LOC coral piece-ACC to put and see if'
-> 'But, if you see a piece of coral under the same microscope, ...'

Ex. 20. <<t1p2s5>>-late
saNhoga sEQzaQhamyENsE byENhwahago iDTanIN gEsIR aR Su iDTa.
Break index 2 2 2 0 2 0 0 3
'coral-nom. growing change-prog.-rel. thing-ACC to see'
-> '(We) can see that the coral is changing while growing'

Ex. 21. <<coQgaG-HLH%>>
TaG zikigoiNniN sarami nuguNgohani zERmIN coQgaG ANSoni pakiNsImnida
Break index 3- 1 2 2 2 2p 2- 3
'firmly guard-PROG man who-is young bachelor Anthony Parkinson-be'
-> 'The man who is guarding firmly is the young bachelor, Anthony Parkinson'

**3.5 The miscellaneous tier**

The miscellaneous tier will be used for any comments or markings (e.g., silence, audible breaths, laughter, disfluencies, and so on) desired by particular transcription groups. The only conventions K-ToBI specifies for this tier are that events that cover some clearly specifiable interval (such as breaths, silence or laughter) be labeled by the < .... > pair, aligned with both their temporal beginnings and ends. Event labels are written only before '>'.

| | |
|---|---|
| < | beginning of an interval (laughter) |
| laughter> | end of a period of laughter |

Examples showing all tiers are shown below. PL refers to a phonological tone tier and PT refers to a phonetic tone tier. Break index is abbreviated as 'BI', and miscellaneous tier as 'misc'.

Ex. 17. <<t1p2s6>> zIG, saNhonIN saraiSImyE aMsEgIN zugEiNnIN gEsida
PL: L% LHa L% LHa L%
PT: H L% H+H Ha H+H L+L% L Ha +H L+ L%
BI: 3 2 3 2- 1 3
misc: <Vdev>
'i.e., coral-TOP alive and rock-TOP dead-progressive rel.marker to be'

-> ‘That is, coral is alive and a rock is dead’

Ex. 21. <<coQgaG-HLH%>>

TaG zikigoiNniN sarami nuguNgohani zERmIN coQgaG ANSoni pakiNsIMnida

PL: H% LHa LHa LHa LHa HLH%

PT: L H% +H L+Ha L L+Ha L Ha L Ha L+H L+ HLH%

BI: 3- 1 2 2 2 2p 2- 3

misc: <Vdev> <sil>

‘firmly guard-PROG man who-is young bachelor Anthony Parkinson-be’

-> ‘The man who is guarding firmly is the young bachelor, Anthony Parkinson’

Ex. 22. <<millennium>>-early

yozIM gIrEN gyohwega i- icENnyENi miRreniEmi ize

PL: LHa LHa LHa LHa H% LHa

PT: L Ha L Ha L L+Ha L+H L+ La L+H L+ H% H La

BI: 2 2 2 2 3- 1m

misc: <disfl>

‘These days that church-NOM, eh., Year 2000-NOM millennium-NOM now’

-> ‘These days, that kind of church, eh, Year 2000, millennium now...’

Ex.23. <<millennium>>-middle

ize nAnyENbutE (ne) sizagi dwegu

PL: LHa LH% LHa HL%

PT: H La L+H LH% H Ha L > HL%

BI: 1m 3 2- 3

misc: <other spkr>

‘now next year-from (yes) beginning-NOM become’

-> ‘Now, (it will) start from next year (Yes)...’

Ex. 24 <<millennium>>-late

usEN manIN gyohwe(do) ceiNziga dweNda gIreyo. (ne)

PL: LHa LHa LHa HL% L%

PT: L Ha L Ha L La H+H L+ HL% H L%

BI: 2- 2- 2 1 1 3 3

misc: <other spkr>

‘First of all many church (too) change-NOM become they say (yes)’

-> ‘They say, first of all, many churches will change, too (Yes)’

**4. Online Data Files and Future Versions**

All examples (sound file, f0 track, and labels) shown in this manual can be accessed from the Sun workstation in the Phonetics Lab of the UCLA Department of Linguistics. This directory includes more examples, some labeled and some not, for labelers to practice transcribing in the K-ToBI system. As more speech data become available, these

labeling guidelines may be further refined. To get speech files and label files mentioned in this paper, contact jun@humnet.ucla.edu. This and earlier versions of K-ToBI manual are available on the author's web site (http://www.linguistics.ucla.edu/people/jun/sunah.htm), and also on UCLA Phonetics Lab web site (http://www.linguistics.ucla.edu/ faciliti/uclaplab.html).

**References**


Beckman, Mary & Gayle Ayers (1994) "Guidelines for ToBI Labelling". Unpublished ms. Ohio State University. Version 3. March 1997. Downloadable ms [http://ling.ohio-state.edu/Phonetics/etobi_homepage.html]. For information on obtaining by ftp, send e-mail to tobi@ling.ohio-state.edu or visit http://ling.ohio-state.edu/~tobi/]

Beckman, Mary & Hirschberg, Julia (1994) "The ToBI Annotation Conventions". Ms. Ohio State University.

Beckman, Mary & Jun, Sun-Ah (1996) "K-ToBI (Korean ToBI) Labelling Convention" Version 2. Ms. Ohio State University and UCLA. Ms is available in [http://www.linguistics.ucla.edu/people/jun/sunah.htm.]

Beckman, Mary & Pierrehumbert, Janet (1986) "Intonational Structure in Japanese and English", *Phonology Yearbook* 3:255-309.

Campbell, Nick & Venditti, Jennifer (1995) "J-ToBI: An Intonational Labeling System for Japanese," Paper presented at the Autumn meeting of the Acoustical Society of Japan.

de Jong, Kenneth (1989) "Initial Tones and Prominence in Seoul Korean," a paper presented at the 117th meeting of the Acoustical Society of America, Syracuse, N.Y.; A paper published in the *Ohio State University Working Papers in Linguistics*, No. 43, pp. 1-14 (1994).

Jun, Sun-Ah (1989) "The Accentual Pattern and Prosody of Chonnam Dialect of Korean," in S. Kuno et al. (eds.) *Harvard Studies in Korean Linguistics* III. pp. 89-100. Harvard Univ. Cambridge, Mass.

Jun, Sun-Ah (1990) "The Prosodic Structure of Korean - in terms of voicing," in E-J. Baek (ed.) *Proceedings of the Seventh International Conference on Korean Linguistics*. Vol. 7: 87-104, Univ. of Toronto Press.

Jun, Sun-Ah (1993) *The Phonetics and Phonology of Korean Prosody*. Ph.D. Dissertation, the Ohio State University. [Published in 1996 by Garland, New York]

Jun, Sun-Ah (1996) "Influence of Microprosody on Macroprosody: A Case of Phrase Initial Strengthening", *UCLA Working Papers in Phonetics* 92: 97-116

Jun, Sun-Ah (1998) "The Accentual Phrase in the Korean Prosodic Hierarchy", *Phonology*. 15.2: 189-226

Jun, Sun-Ah & Oh, Mira (1996) "A Prosodic Analysis of Three Types of Wh-phrases in Korean", *Language and Speech* 39(1):37-61.

Lee, Hyuck-Joon (1999) Tonal Realization and Implementation of the Accentual Phrase in Seoul Korean. MA thesis, UCLA.

Lee, Sook-hyang (1989) "Intonational Domains of the Seoul Dialect of Korean," a paper presented at the 117th meeting of the Acoustical Society of America, Syracuse, N.Y.; An abstract in *Journal of the Acoustical Society of America*, vol. 85, suppl. 1, p. S99.

Park, Mee-Jeong (2000) "Where Prosody Meets Grammar: Taxonomy of Korean Prosodic Boundary Tones". Ms. UCLA.

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

Pierrehumbert, Janet & Beckman, Mary (1988) *Japanese Tone Structure*, MIT Press.

Pitrelli, John; Beckman, Mary; & Hirschberg, Julia (1994) "Evaluation of Prosodic Transcription Labeling Reliability in the ToBI Framework," *Proceedings of the 1992 International Conference on Spoken Language Processing*, vol. 1, pp. 123-126.

Schafer, Amy & Jun, Sun-Ah (submitted) "Effects of Accentual Phrasing on Adjective Interpretation in Korean", in M. Nakayama (ed.), *East Asian Language Processing,* Stanford, CSLI. [Proceedings of the International East Asian Psycholinguistics Workshop], August, 1999, Columbus.

Silverman, Kim; Beckman, Mary; Pitrelli, John; Ostendorf, Mari; Wightman, Colin; Price, Patti; Pierrehumbert, Janet; & Hirschberg, Julia (1992) "ToBI: A Standard for Labeling English Prosody," *Proceedings of the 1992 International Conference on Spoken Language Processing*, vol. 2, pp. 867-870.

Venditti, Jennifer (1995) Japanese ToBI Labeling Guidelines. Ms with examples, Ohio State University. [For information on obtaining by ftp, send e-mail to venditti@ling.ohio-state.edu.]

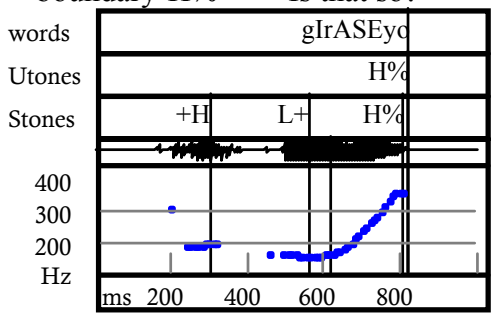
**Appendix A: Romanization Convention**

| 1. Consonants | | | | 2. Vowels | | |
|---|---|---|---|---|---|---|
| Hangul | [IPA] | Roman | letters | Hangul | [IPA] | Roman letters |
| | | Onset | Coda | 아 | [ɑ] | a |
| ㅂ | [p] | b | B | 어 | [ɘ] | E |
| ㄷ | [t] | d | D | 오 | [o] | o |
| ㄱ | [k] | g | G | 우 | [u] | u |
| ㅈ | [ʧ] | z | D | 으 | [ɨ] | I |
| ㅍ | [$p^h$] | p | B | 이 | [i] | i |
| ㅌ | [$t^h$] | t | D | 에 | [e/ɛ] | e |
| ㅋ | [$k^h$] | k | G | 애 | [ɛ] | A |
| ㅊ | [$ʧ^h$] | c | D | 의 | [ɯi] | U |
| ㅃ | [p'] | P | B | 야 | [ja] | ya |
| ㄸ | [t'] | T | D | 여 | [jɘ] | yE |
| ㄲ | [k'] | K | G | 요 | [jo] | yo |
| ㅉ | [ʧ′] | C | D | 유 | [ju] | yu |
| ㅅ | [s] | S | D | 예 | [je/jɛ] | ye |
| ㅆ | [s'] | S | D | 얘 | [jɛ] | yA |
| ㅎ | [h] | H | - | 와 | [wa] | wa |
| ㄹ | [l] | R | R | 워 | [wɘ] | wE |
| ㅁ | [m] | M | M | 외 | [we/wɛ] | we |
| ㄴ | [n] | N | N | 왜 | [wɛ] | wA |
| ㅇ | [ŋ] | - | Q | 위 | [wi] | wi |

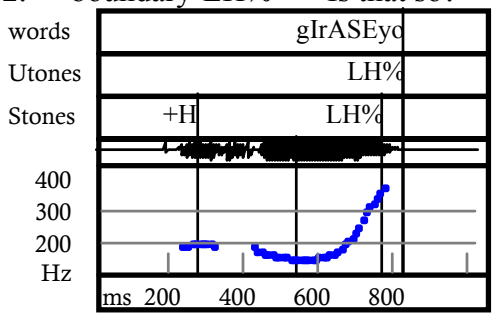
## Appendix B

Pitch tracks and labels are made using *PitchWorks* (Scicon). The word tier is labelled as 'words', the phonological tone tier as 'Utones' and the phonetic tone tier as 'Stones', the break index as 'break', and the miscellaneous tier as 'misc'. In #1-4 below, a vertical line marking the beginning of the last syllable, '-yo' [jo], is drawn before the line marking a boundary tone or '>'. The figure numbers match the numbers of the examples in the main text.

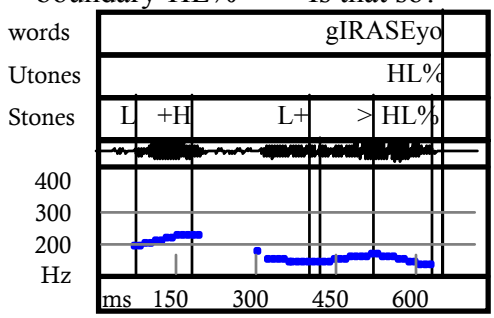
1. <<boundary-H%>> 'Is that so?'



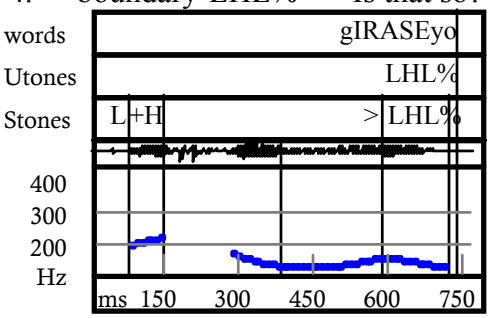

2. <<boundary-LH%>> 'Is that so?'



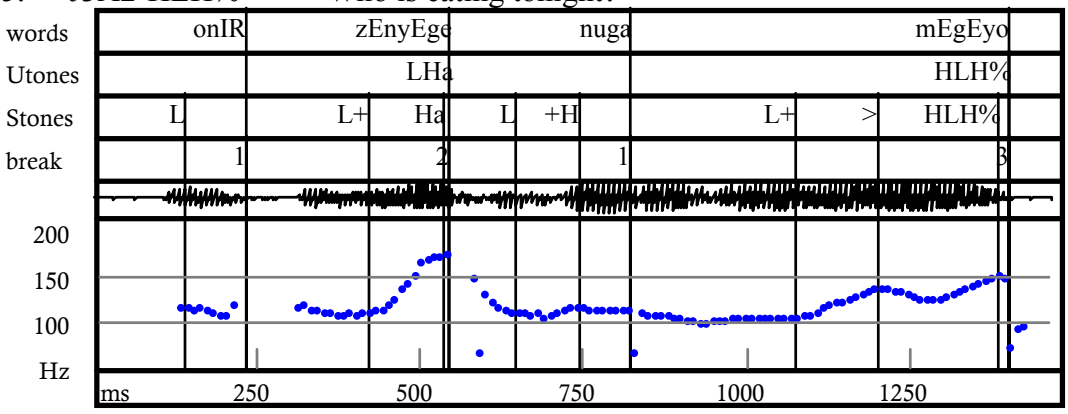

3. <<boundary-HL%>> 'Is that so?'



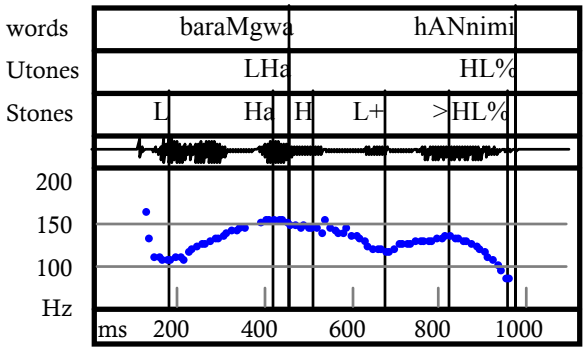

4. <<boundary-LHL%>> 'Is that so?'



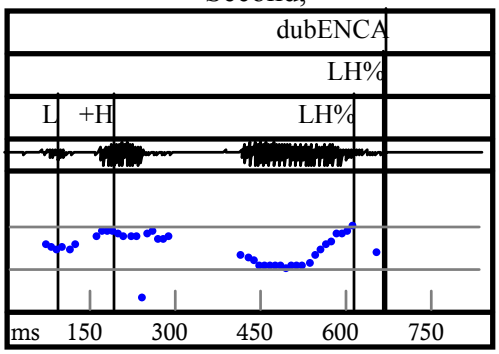

5. <<J3A2-HLH%>> 'Who is eating tonight?'



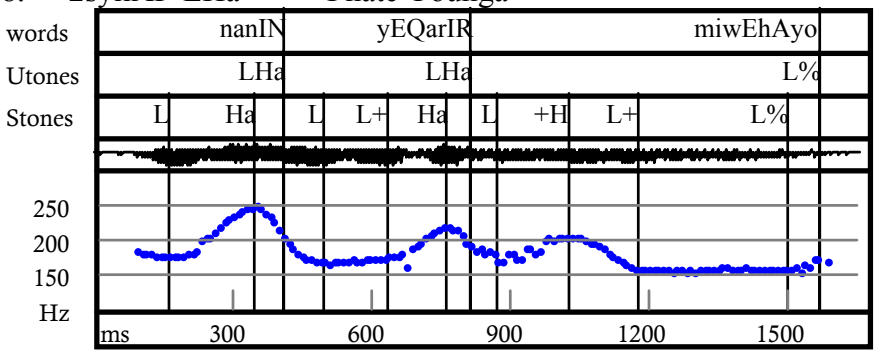
6. <<IPboundary-HL%>>
'The North Wind & the Sun-NOM..'



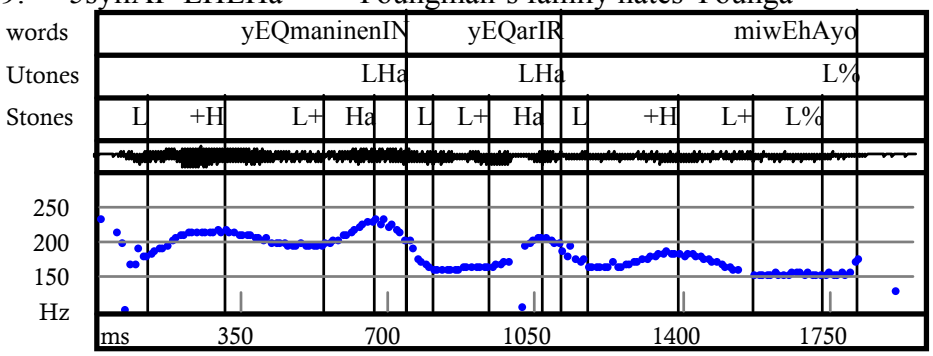

7. <<IPboundary-LH%>>
'Second,'



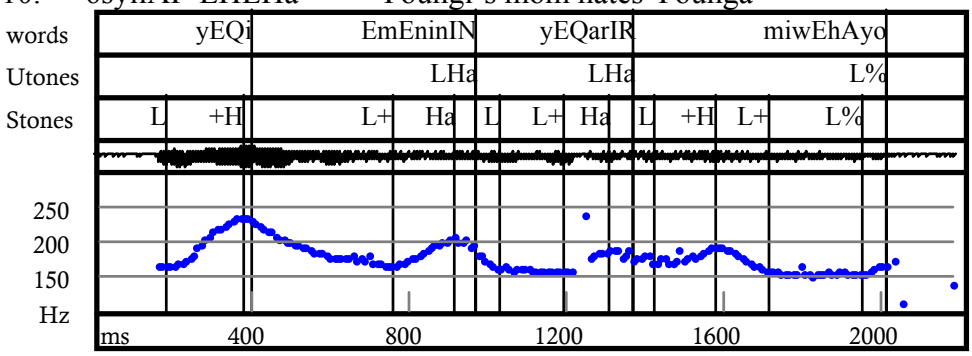

8. <<2syllAP-LHa>> 'I hate Younga'



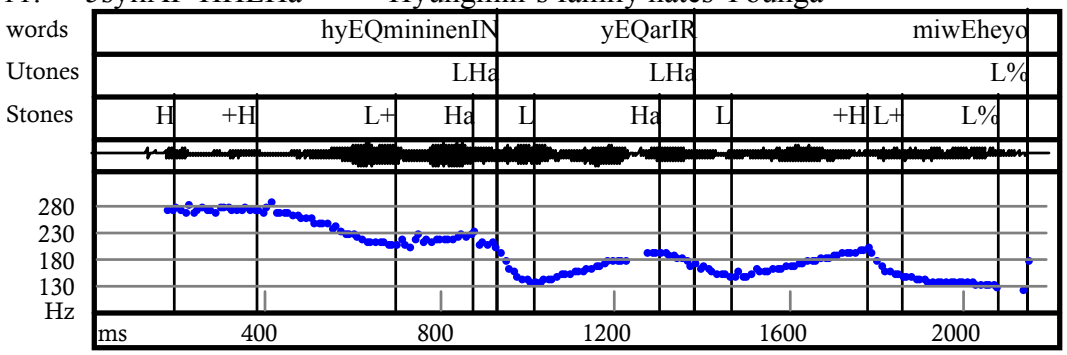

9. <<5syllAP-LHLHa>> 'Youngman's family hates Younga'



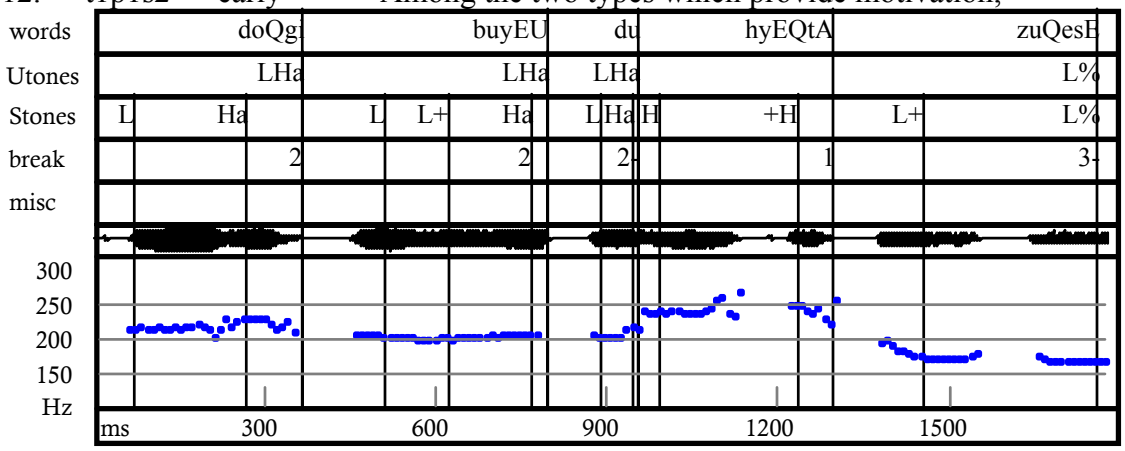

10. <<6syllAP-LHLHa>> 'Youngi's mom hates Younga'



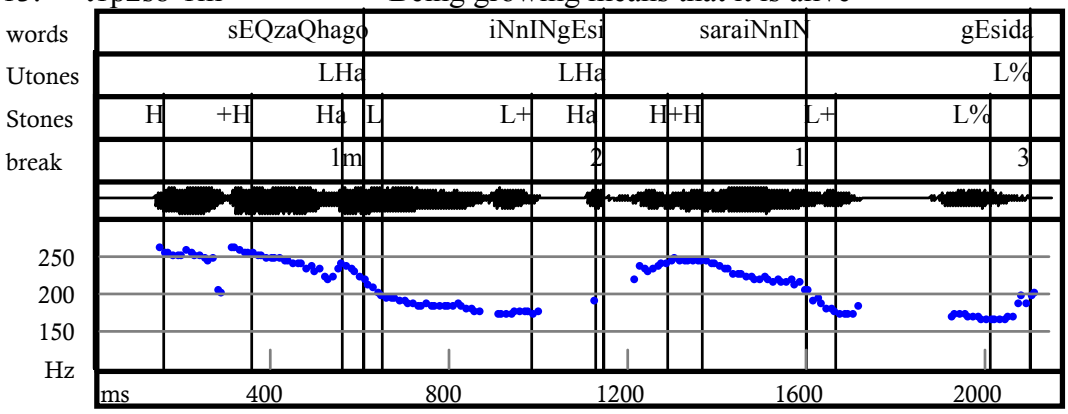
11. <<5syllAP-HHLHa>>  ‘Hyungmin’s family hates Younga’



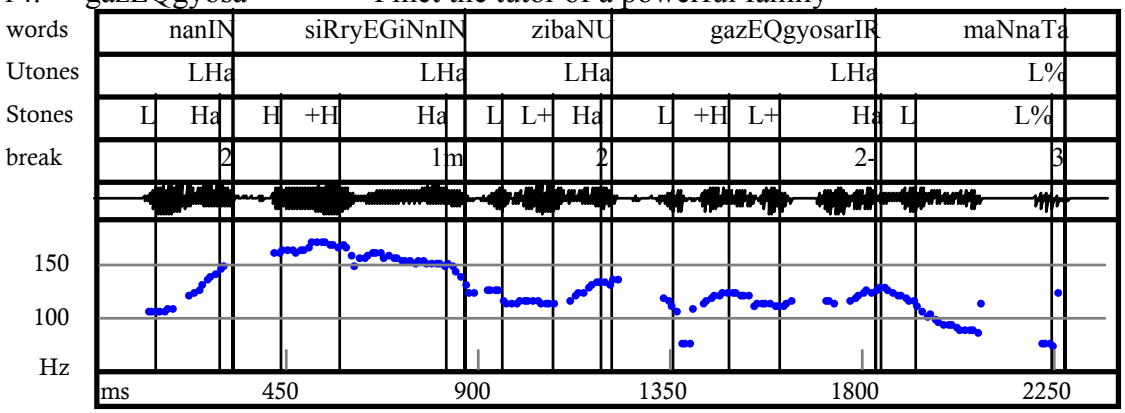

12. <<t1p1s2>>-early  ‘Among the two types which provide motivation,’



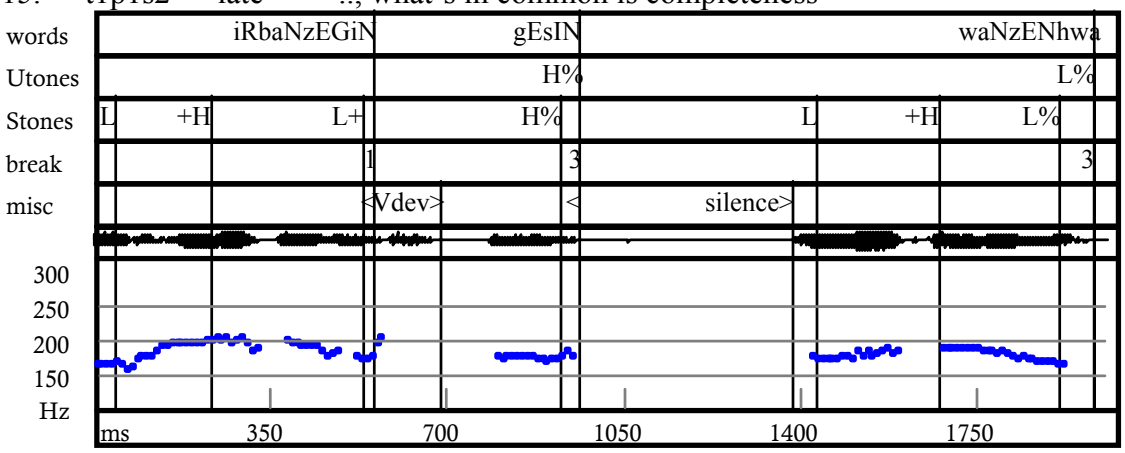

13. <<t1p2s8-1m>>  ‘Being growing means that it is alive’



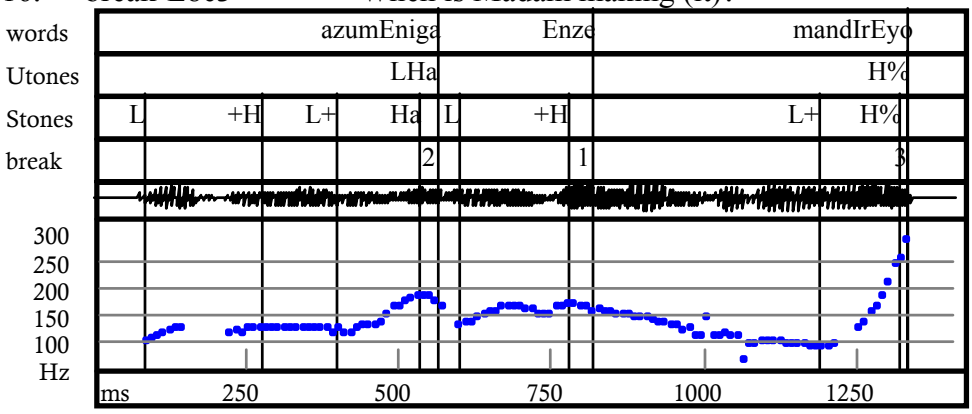
14. <<gazEQgyosa>> ‘I met the tutor of a powerful family’



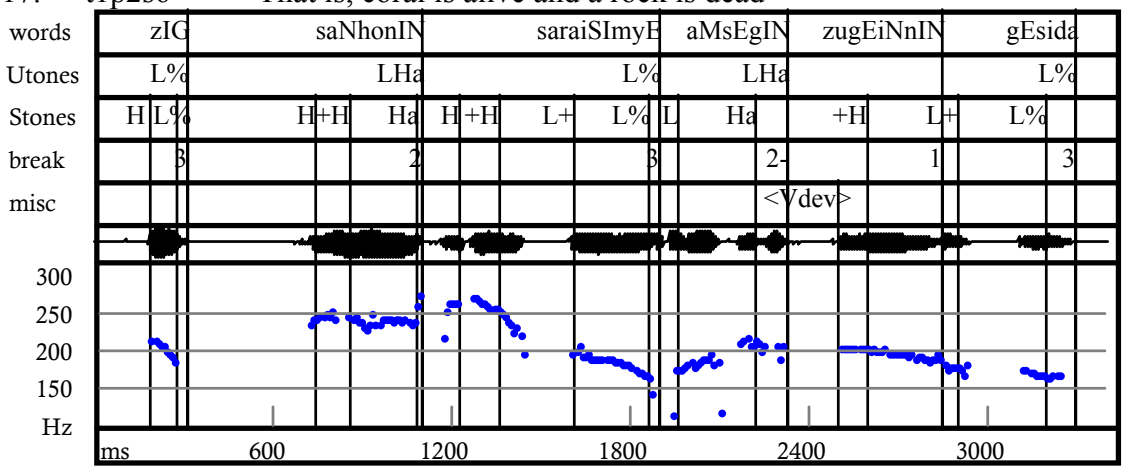

15. <<t1p1s2>>-late ‘.., what’s in common is completeness’



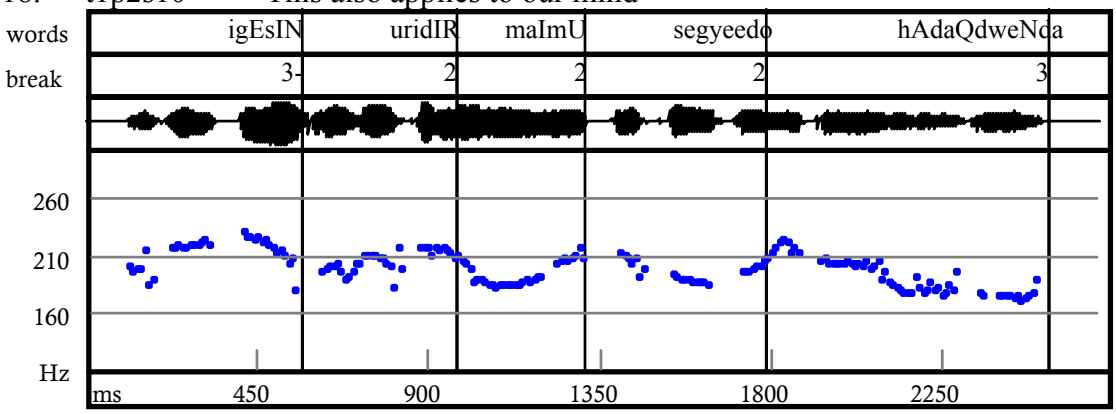

16. <<break-L8c3>> ‘When is Madam making (it)?’



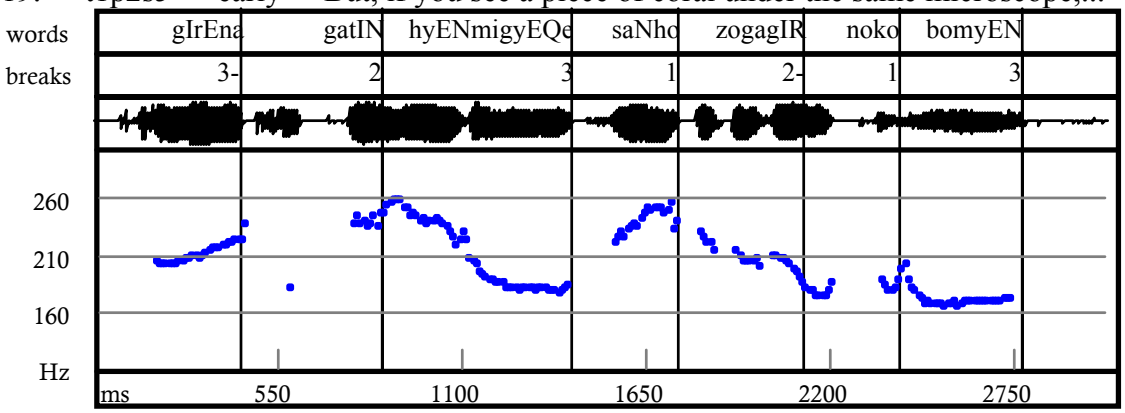
17. <<t1p2s6>>  ‘That is, coral is alive and a rock is dead’



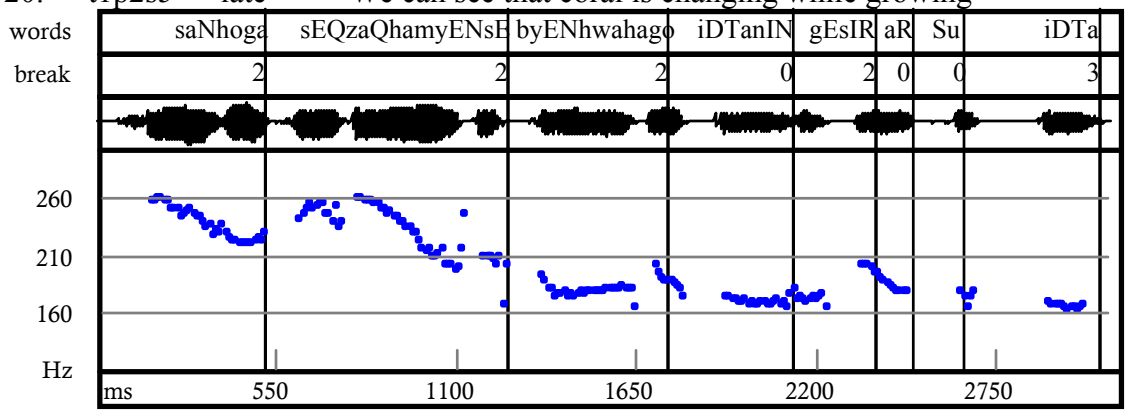

18. <<t1p2s10>>  ‘This also applies to our mind’




19. <<t1p2s5>>-early  ‘But, if you see a piece of coral under the same microscope,...’

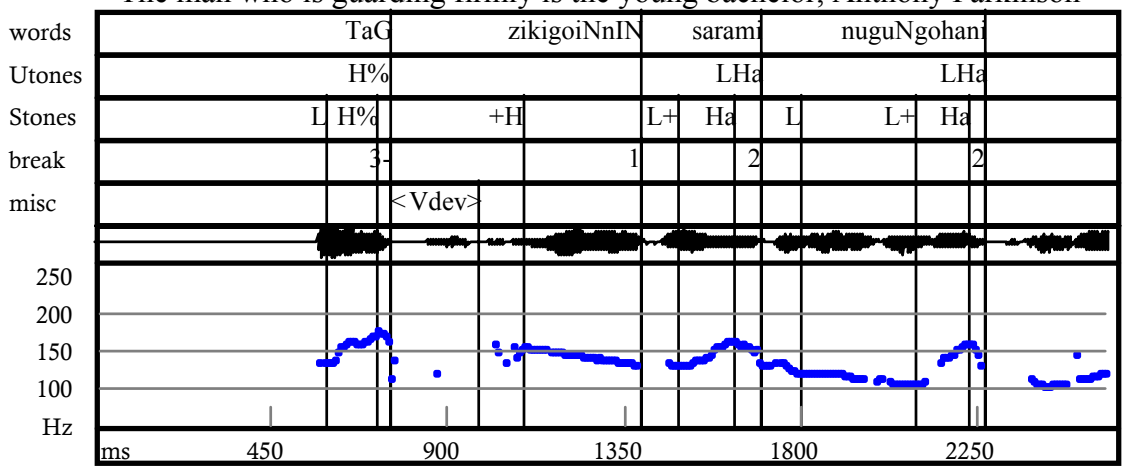


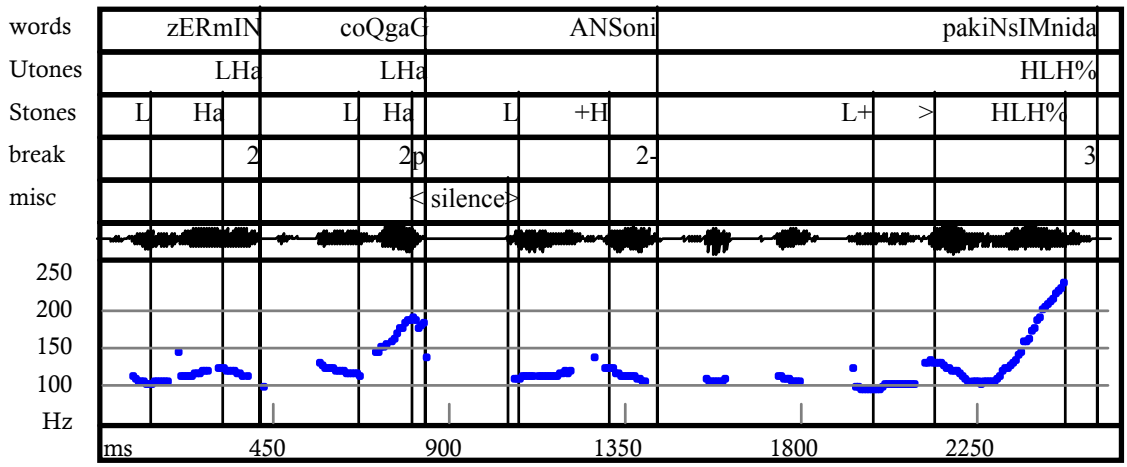
20. <<t1p2s5>>-late ‘We can see that coral is changing while growing’




21. <<coQgaG-HLH%>>

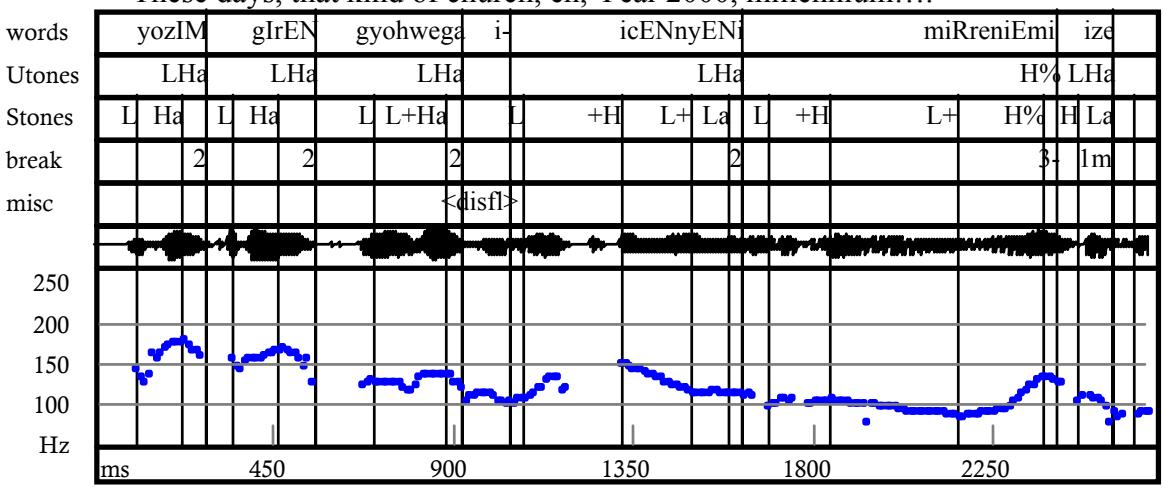
‘The man who is guarding firmly is the young bachelor, Anthony Parkinson’



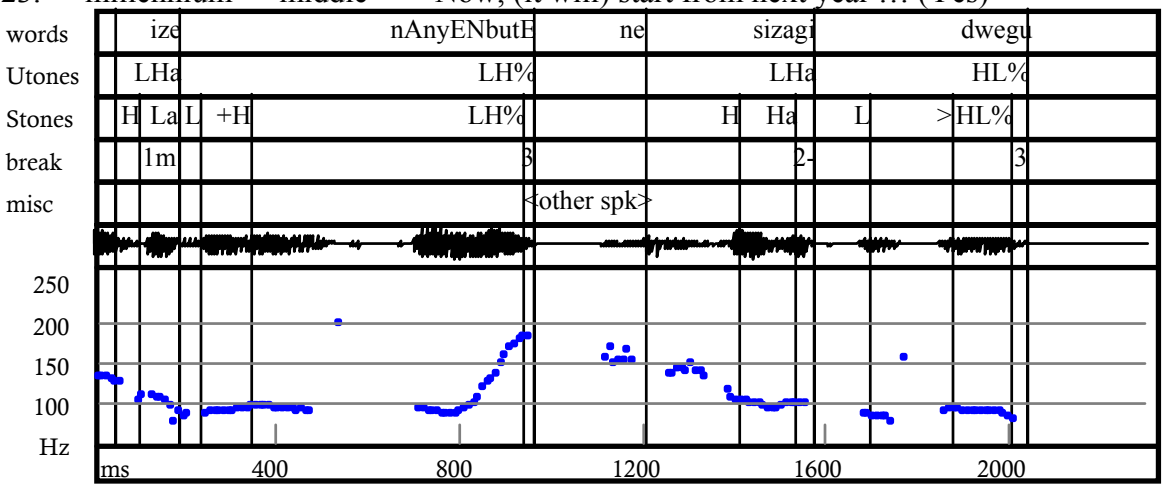



22. <<millennium>>-early

'These days, that kind of church, eh, Year 2000, millennium….'




23. <<millennium>>-middle 'Now, (it will) start from next year … (Yes)'




24. <<millennium>>-late 'They say, first of all, many churches will change, too (Yes)'