K-ToBI (Korean ToBI) Labelling Conventions

(version 3.0, revised in January 2000)

Sun-Ah Jun

Dept. of Linguistics, UCLA

1. Background

K-ToBI 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 more directly on J_ToBI, the Japanese ToBI system, 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 (1993, 1995, 1996, 1998; also see 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 Korean synthesis development project. The second version (Beckman & Jun 1996) was an updated one 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 revised one from the second version by Sun-Ah Jun after Korean ToBI Workshop in Korea, August 1998, and 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, 1996, 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, 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. THLH (T=H if the AP initial segment is aspirated or tense, T=L otherwise), but not by final lengthening. The intonational structure of Seoul Korean is schematically represented in Figure 1.

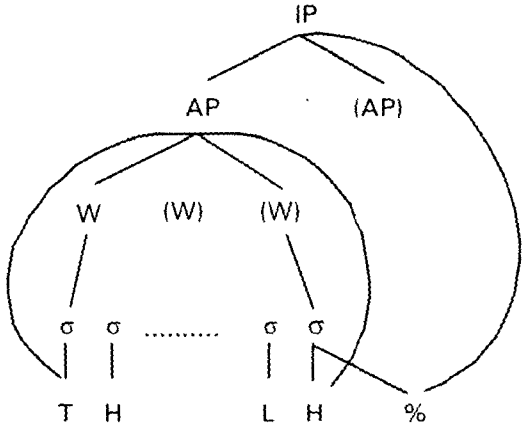


Figure 1. Intonational Structure of Seoul Korean
IP: Intonation Phrase AP: Accentual Phrase
w: phonological word σ: syllable
T= H, when the syllable initial segment is aspirated/tense, otherwise, T= L
%: Intonation phrase boundary tone

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 an IP, and 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 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 1995), and HL% marks both a declarative and a wh-question. More research is needed to identify distinctive pragmatic meaning for each boundary tone. _

2. Structure of K-ToBI

The original English ToBI, Japanese ToBI, and Korean ToBI have four parallel tiers (word, tone, break-index, and miscellaneous), but allow the free proliferation of site-specific extra tiers. Each ToBI users are encouraged to add their own customized tiers to label events of site-specific interest, and keep records of why each particular tier was added and how it is used. By comparing extra tiers across labeling sites, we may find that we all agree on the desirability of some generally used tiers and can make it to be obligatory for Seoul Korean or other dialects.

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 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 a 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 distinctive 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 a 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, 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 can be 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 shorter 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, nor predictable, it is not known yet if all these variations are indeed not distinctive nor predictable. By labelling surface tonal patterns, we will be able to investigate if there is any meaning difference among these patterns.

Second, the earlier version of K-ToBI labels 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 is no initial H in an AP, H- was not labelled. However, when an AP-like phrase is realized in an L tone, 'LHa' was not labelled. Instead L% was labelled at a tone tier and a break index '2m' was labelled at a break index tier to indicate that the boundary juncture does not match the tone pattern. That is, the degree of juncture is the same as that of the usual AP boundary, i.e., '2', but the tonal mark, L%, shows the boundary of an Intonation Phrase. Sometimes this was indeed the case. However, observation of more natural data revealed that an AP boundary is sometimes realized in an L tone due to the tonal interaction of adjacent tones and stylistic variations. At the moment, the detail condition of AP final L tone and its pragmatic meaning are not known.

We hope to get answers to these issues by labelling 'La' for a falling AP boundary at a 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, the second syllable of an AP is H when the AP has more than 3 syllables. Thus, 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 first or second syllable or rarely on the third syllable, without considering the origin of the H tone or the alignment of the peak to syllables. However, quantitative data show that phonetic realization of these H tones differs depending on their origins and locations. F0 is significantly higher for the H tone on the AP initial syllable (i.e., **H**HLH) than the H tone after the AP initial L tone (i.e., L**H**LH). In addition, this raised f0 value in the beginning of the HHLH pattern influences the following syllables, if there is any, by raising f0 values, 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 LHLH or the second H in HHLH is predictable, we did not label these tones in the earlier version. But these are not always predictable, and furthermore, as mentioned earlier, the individual tone itself forming an AP does not seem to be meaningful. What is important in Korean intonational phonology is the boundary marker of an AP and an IP. 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 at a phonetic tone tier will provide us data by which we can determine whether these tones are meaningful in Korean intonational phonology. It will also provide valuable information to researchers working on speech synthesis and recognition.

Fourth, by separating a 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. However, the accentual phrasing of this dialect is the same as that of the Seoul dialect. Thus, a phonological tone tier for Seoul Korean will remain the same for transcribing tones for the Chonnam dialect, while a phonetic tone tier of these two dialects will differ conforming to the surface realization of each dialect.

In the following sections, each of the five tiers is defined, and labels and symbols proper for each tier are introduced. In addition, example sentences illustrate in a text format how to label information in 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. Since what constitutes a "word" in Korean is a matter of some debate, we cannot be very specific here about how frequently to place word labels. 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. At the other end of the scale, another site may need only that there be as many labels as there are spaces in a standard Hangul transcript of the text. 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, and that eventually a consensus will emerge from these needs. In this version, we consider 'word' as a sequence of segments divided by space in a written text.

If the labeling platform is xwaves and xlabel (or any similar labeling platform 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 marked, using some site specific convention for the Hangul or romanized spelling. A romanization convention used at UCLA site is in Appendix A.

3.2 A 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 an IP-final boundary tone, only IP final boundary tone (%) will be labeled at the end of an IP.

To mark the end of an 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 nine different boundary tones, i.e. H%, L%, HL%, LH%, HLH%, LHL%, HLHL%, LHLH%, LHLHL%. To simplify the description of IP boundary tones, 'T' is used below as a variable of the IP boundary tones. Instructions of where to put phonological tone labels are given below. Sentence examples labelled with phonological tones will be shown after the description of a phonetic tone tier in the next section.

LHa marks the end of an IP-medial AP, aligned with the end of 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 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 an AP and final to an IP, only the IP boundary tone is written at the end of the word.

3.3 A phonetic tone tier

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

La this is a less common AP final tone, sometimes seen when the following AP begins with a H tone. 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 the AP final H

tone or before the IP final H boundary tone. Do not label this tone if it is *predictable* from adjacent tone labels. For example, when an AP is continuously falling from an initial H to final La, L+ should not be labelled. Also when an AP initial is L and final is La, L+ should not be labelled. 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, 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 pitch point 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 (or sometimes the third syllable when the AP is long or uttered fast or produced under focus) of an AP. This label should be placed aligned with the f0 peak on 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 H boundary, Ha, and the second row shows AP patterns with a L 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.

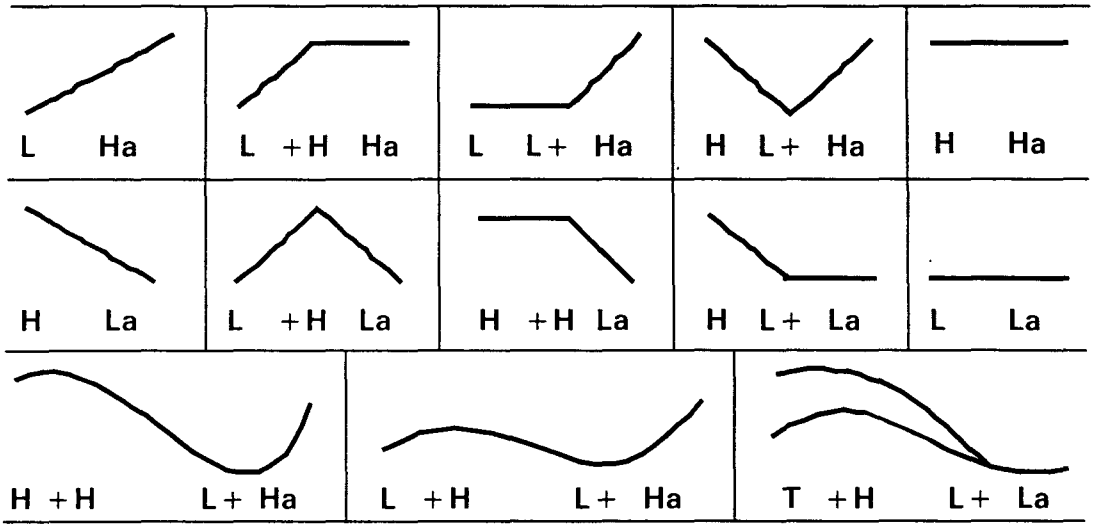


Figure 2. Schematic f0 contours of fourteen tonal patterns of AP.

For the IP boundary tones, five tones ending in L boundary tone (i.e., L%/HL%/ LHL%/HLHL%/LHLHL%) are placed at the f0 minimum toward the end of an IP final syllable, and four tones ending in H tone (i.e., H%/LH%/HLH%/LHLH%) are placed at the f0 maximum toward the end of an IP final syllable. For complex boundary tones ending in H% but has a H tone before the final H tone (ex., HLH%, LHLH%), '>' referring to an 'early peak' should be placed at the f0 peak corresponding to the first H tone in the boundary. For complex boundary tones ending in L% but has one or more H tones before the final L tone (ex. HL%, LHL%, LHLHL%), '>' should be placed aligned with each peak in f0 contour. The following provides surface realization rules of each boundary tone, and its location relative to words and f0 contours.

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.

H%: a rising boundary tone that begins to rise before the IP final syllable, and reached 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 very commonly used for questions, continuation rises and explanations. It is also used to signal 'being annoyed' or unpleasant (ex. [kI\ethAnik`A kI\´nE]! 'I have already told you so. (Why do you keep asking me?)' or [p´\´s`´!] '(Did you) throw it out? (I can't believe that!)').

HL%: HL%: a falling boundary one that rises to a peak before the last syllable, and then falls during the last syllable. -- essentially a combination of H% followed by L%. However, in some cases such as in news, this boundary tone is realized earlier than the final syllable showing the H tone on the penult and L on the final syllable. This is easily noticeable because H tone does not show up on the penultimate syllable otherwise. This tone should be placed at the end of the phrase aligned with the minimum f0 value, and the location of H is marked by '>' aligned with the f0 peak whether it be on the penultimate or final syllable. This tone is very common in declaratives and wh-questions.

LHL%: a rising-falling boundary tone that rises during the IP final syllable. Unlike HL%, the IP final syllable starts a minimum f0 value, and rise begins considerably later -- essentially a combination tone consisting of LH% followed by L%. This tone should be placed at the end of the phrase aligned with the minimum f0 value, and the location of H is marked by '>' above the f0 peak. This tone is also common in declaratives like HL%, but implies being more persuasive, insisting, and confirmative. It is also used to show

being annoyed or unpleasant. (ex. [hAtSima]! 'Don't do it (I told you before)')

HLH%: a fall-rise 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 small rise. This tone should be placed at the end of the phrase aligned with the maximum f0 value. The location of the first H is marked by '>' above the f0 peak. The tone is not as common as other types, and some speakers use this type more often than others. This type is used when a speaker is confident and expecting listeners' agreement.

LHLH% a rising-falling-rising boundary tone -- a combination of LHL% and H%. The timing of rise is like LHL%, followed by H%. This tone should be placed at the end of the phrase aligned with the maximum f0 value. The location of the first H is marked by '>' above the f0 peak. This tone is less common than others.

HLHL% a falling-rising-falling boundary tone. The timing of rise is like HLH%, followed by L%. This tone should be placed at the end of the phrase aligned with the minimum f0 value. The location of the two Hs are marked by '>' above the f0 peak. This tone is more common than LHLH%, and has a meaning of intensifying HL%, confirming and insisting one's opinion. Sometimes, like LHL%, it delivers nagging or persuading meanings.

LHLHL% a rising-falling-rising-falling boundary tone. The timing of rise is like LHLH%, followed by L%. This tone should be placed at the end of the phrase aligned with the minimum f0 value. The location of the two Hs are marked by '>' above the f0 peak. The meaning of this tone seems to be similar to LHL%, but with more 'intense meaning of being annoyed'. It is not clear at the moment if this tone is categorically different from LHL%.

Schematic f0 contours of eight types of IP boundary tone realizations are shown in Figure 3. The first row shows an IP boundary 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.

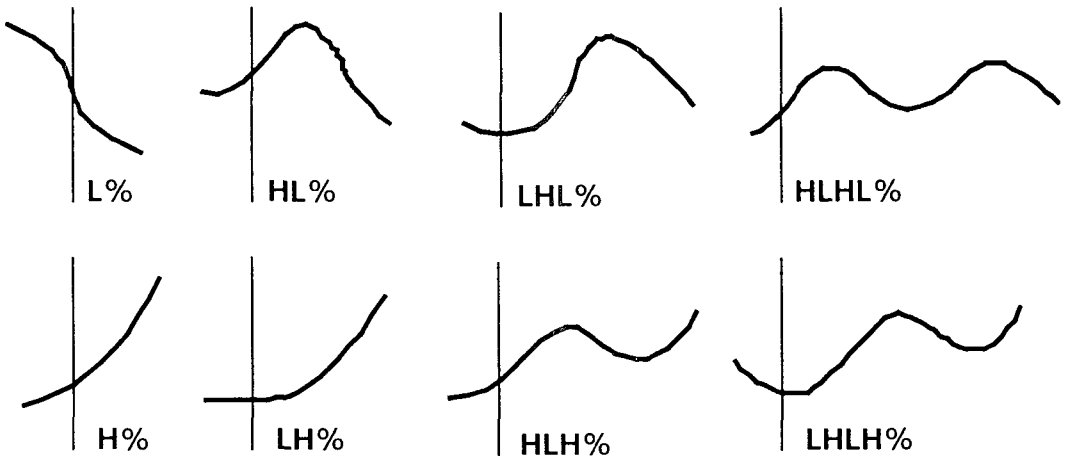


Figure 3. Schematic f0 contours of eight boundary tones of IP.

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

X	underspecified tonal event of non-AP final boundary tone. (Tone is there, but the tonal value have yet to be assigned)
a	underspecified AP final tone
%	underspecified IP final tone
X?	uncertain of the type of a tone, which is not an AP final nor IP final boundary tone. (a 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 and a phonetic tone are shown below. Romanization of Korean font is shown in Appendix A, and f0 tracks of each example are shown in Appendix B.

Examples of AP realization and IP boundary tones:

Ex.1. <<4boundary-H%>> gIrASEjo 'Is that so?'
phonological tone tier H%
phonetic tone tier L+H L+ H%

Ex.2. <<4boundary-LH%>> gIrASEjo 'Is that so?'
phonological tone tier LH%
phonetic tone tier L+H LH%

Ex.3. <<4boundary-HL%>> gIrASEjo 'Is that so?'
phonological tone tier HL%
phonetic tone tier L+H L+ HL%

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

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%
'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%
'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%
'Hyungmin's family-TOP Younga-ACC hate'
-> 'Hyungmin's family hates Younga'

Ex.12. <<t1p1s2>>-early doQgi bujEU du hjEQtA zuQesE ...
phonological tone tier LHa LHa LHa L%
ponetic tone tier L Ha L L+Ha L Ha H +H L+ L%
'motivation providing-POSS two types among ...'
-> 'Among 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 L+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 maNnaTTa.
phonological tone tier: LHa LHa LHa L%
phonetic tone tier: L Ha H +H 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 a rating for 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 pronouns which are different from a real "word", not pronounced by itself (e.g. /su/ 'way', /te 'place', /k't 'thing').

1 for phrase-internal "word" boundaries not marked by such cliticization phenomena, and can be pronounced by itself. "Word" is defined as a sequence of sounds separated by space in orthography, and is 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 associated typically with the tonal pattern 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 associated typically 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 completely redundant to 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 without any tonal markings, or with 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 with the tonal markings of an AP.

In an xwaves xlabel type system, the break index label should be associated 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 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 at a break index tier.

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

Example sentences:

Ex.12. <<t1p1s2>>-early	doQgi	bujEU	du	hjEQtA	zuQesE	...
break index	2	2	2-	1	3-	
	'motivation providing-POSS two types among ...'					
	-> 'Among two types which provide motivation,'					

Ex. 13 <<t1p2s8-1m>>	sEQzaQhago iNnIN gEsi	saraiNnIN gEsida			
break index tier	1m	2	1	3	
	'to grow-prog. is	to live-prog.'			
	-> 'Growing means that it is alive'				

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

Ex. 15. <<t1p1s2>>-late iRbaNzEgiN kEsIn waNzEnhwa.
1 3 3
'general-rel thing-TOP completeness'
-> '(Among two types which provide motivation,) the general thing is completeness'

Ex 16.. <<break-L8c3>> azumEninga ENze maNdIrEjo?
break index 2 1 3
'madam-NOM when make-Q'
-> 'When is Madam making (it)?'

Ex. 17. <<t1p2s6>> zIG. saNhonIN saraiSImjE aMsEgIN zugEiNnIn gEsida
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 rock is dead'

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

Ex. 19. <<t1p2s5>>-early
glrEna. gatIN hjENmigyEQe sanho zogagIR noko bomyEN
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.

2 2 2 0 2 0 0 3

'coral-nom. growing change-prog.-rel. thing-ACC to see'

-> 'We can see that coral is changing while it's growing'

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

Tak zikigoiNniN sarami nuguNgohani zERmIN coQgaK ANSoni pakiNsImnida

3- 1 2 2 2m 2p 1 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 misc 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 at both their temporal beginnings and ends, using label pairs below, so that the interval is delimited by the < .... > pair.

laughter<	beginning of interval of 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'.

Ex. 17. <<t1p2s6>>

	zIG.		saNhonIN			saraiSImyE			aMsEgIN			zugEiNnIN			gEsida
PL tone:	L%			LHa			L%			LHa					L%
PT tone:	H	L%	H+H		Ha	H-H		L%	L		Ha	L+H			L%
BI:		3		2			3			2-			0		3
misc:												<Vdev>			

'i.e.. coral-TOP alive and rock-TOP dead-progressive rel.marker to be'
-> 'That is. a coral is alive and a rock is dead'

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

	Tak	zikigoiNniN	sarami	nuguNgohani		zERmIN	coQgaK	ANSoni	pakiNsImnida				
PL:	H%		LHa		LHa		LHa	LHa		LHa		HLH%	
PT:	L H%	+H		L+ Ha	L		L+Ha	L	Ha	L	Ha	L+H	La H L+ > HLH%
BI:	3-		1	2		2		2m		2p		1	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		
PL:	LHa		LHa		LHa		LHa		H%		
PT:	L	Ha	L	Ha	L	L+Ha	L+H	L+ La	L+H	L+ H%	
BI:	2		2		2		2		3-		
misc:					<disfl>						

'Thesedays that church-NOM. eh.. Year 2000-NOM millennium-NOM'
-> 'Thesedays. that kind of church. eh. Year 2000. millennium....'

Ex.23. <<millennium>>-middle

	ize		nAnyENbutE		um		sizagi		dwegu
PL:	LHa			LH%			LHa		HL%
PT:	H ?a	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 ...'

Ex. 24 <<millennium>>-final

	usEN		manIN		gyohwe(do)		ceiNziga		dweNda		gIreyo.		ne
PL:	LHa			LHa			LHa						HL%
PT:	L Ha		L	Ha	L		La	H+H				L+	HL%
BI:	2-			2-			2		1		1	3	
misc:													<other spkr>

'First of all many church (too) change-NOM become they say'

-> 'They say, first of all many churches will change too'

4. Online Data Files and Future Versions

All examples (sound file, f0 track, and labels) shown in this manual can be accessed in the SUN workstation in Phonetics Lab, Linguistics department, UCLA. This directory includes more examples, some labeled and some not, for labellers to practice transcribing the K-ToBI system. As more speech data become available, this labeling guidelines may be refined further. To get speech files and label files, contact jun@humnet.ucla.edu. Texts of this and earlier K-ToBI manual are in Phonetics Lab Web site: http://www.humnet.ucla.edu/humnet/linguistics/faciliti/facilities/k_tobi.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 and visit http://ling.ohio-state.edu/~tobi/]

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

Beckman, Mary & Jun, Sun-Ah (1996) "K-ToBI (Korean ToBI) Labelling Convention" Version 2. Manuscript. Ohio State University and UCLA. Manuscript is available in [http://www.humnet.ucla.edu/humnet/linguistics/faciliti/facilities/k_tobi.html.]

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 7th International Conference on Korean Linguistics*, pp. 87-104. University 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 (1995) "Asymmetrical prosodic effects on the laryngeal gesture in Korean," In Bruce Connell and Amalia Arvaniti, eds., *Phonology and Phonetic Evidence: Papers in Laboratory Phonology IV*, pp. 235-253. Cambridge University Press.

Jun, Sun-Ah & Oh, Mira (1996) "A prosodic analysis of three types of Wh-phrases in Korean", *Language and Speech* 39(1):37-61.

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

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.

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.

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. Manuscript 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

Hangul [IPA]	Roman letters	
	Onset	Coda
§≤ [p]	b	B
§ß [t]	d	D
§° [k]	g	G
§∏ [tS]	z	D
§Ω [p^{h}]	p	B
§° [t^{h}]	t	D
§^{a} [k^{h}]	k	G
§∫ [tS^{h}]	c	D
§≥ [p']	P	B
§® [t']	T	D
§¢ [k']	K	G
§π [tS']	C	D
§µ [s]	s	D
§∂ [s']	S	D
§æ [h]	h	---
§© [\]	r	R
§± [m]	m	M
§§ [n]	n	N
§∑ [N]	---	Q

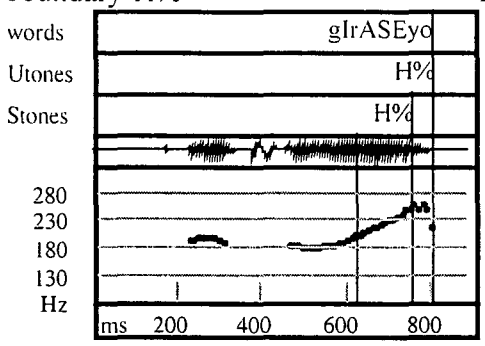
2. Vowels

Hangul [IPA]	Roman letters
æΔ [A]	a
æÓ [´]	E
ø¿ [o]	o
øÏ [u]	u
¿∏ [I]	I
¿Ã [i]	i
ø° [E]	e
æ÷ [E]	A
¿« [¨i]	U
æfl [ja]	ya
ø© [j´]	yE
ø‰ [jo]	yo
¿Ø [ju]	yu
øπ [jE]	ye
æÍ [jE]	yA
øÕ [wa]	wa
øˆ [w´]	wE
ø‹ [wE]	we
ø÷ [wE]	wA
¿ß [wi]	wi

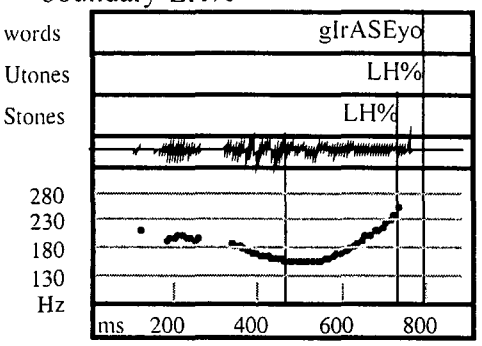
Appendix B

Pitch tracks and labels are drawn using, PitchWorks (Scicon). A word tier is labelled as 'words', a phonological tone tier as 'Utones' and a phonetic tone tier as 'Stones', a break index as 'break', and a miscellaneous tier as 'misc'. In #1-4 below, the vertical line drawn in the waveform and pitch track but not through the tiers marks the beginning of the last syllable, '-yo' [jo]. The number given in each graph matches that in the main text.

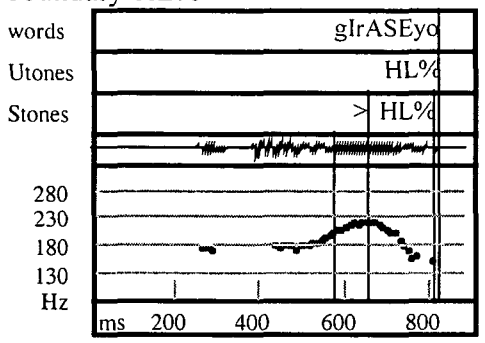
1. <<boundary-H%>>



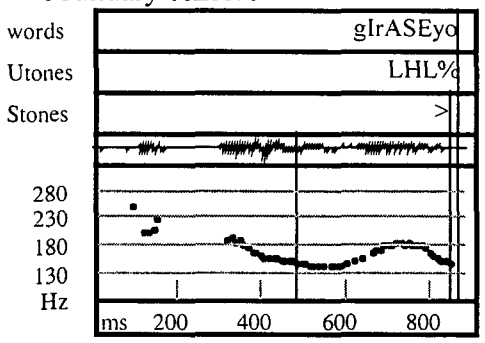
2. <<boundary-LH%>>



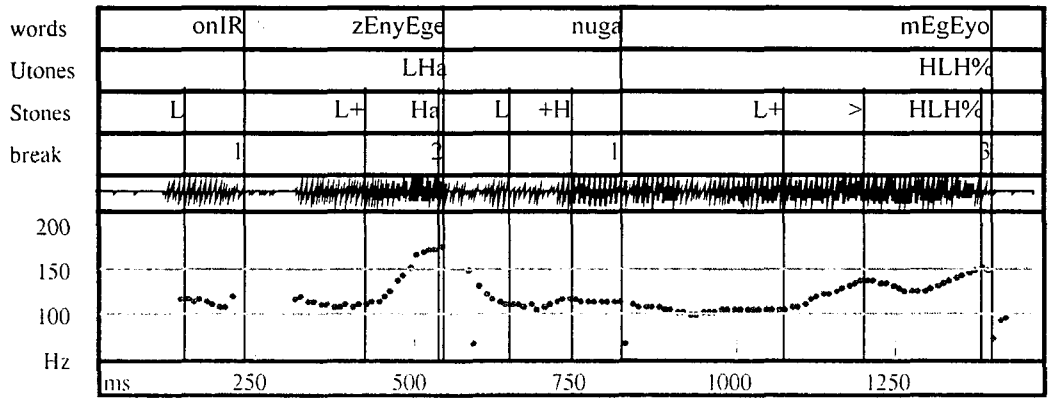
3. <<boundary-HL%>>



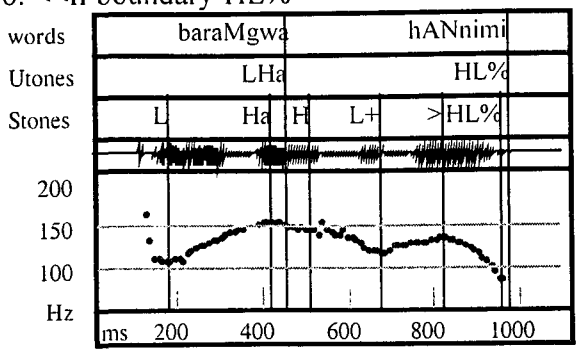
4. <<boundary-HLH%>>



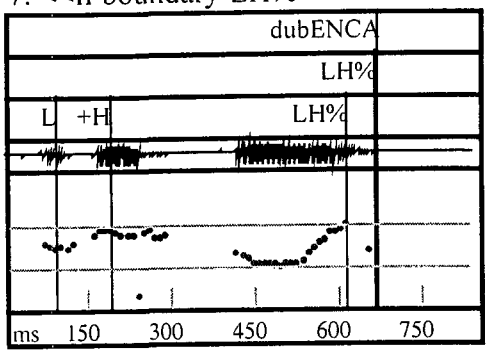
5. <<J3A2-HLH%>>



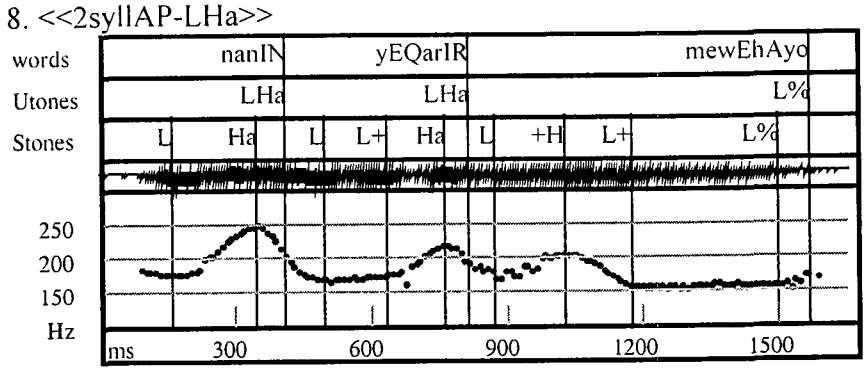
6. <<IPboundary-HL%>>



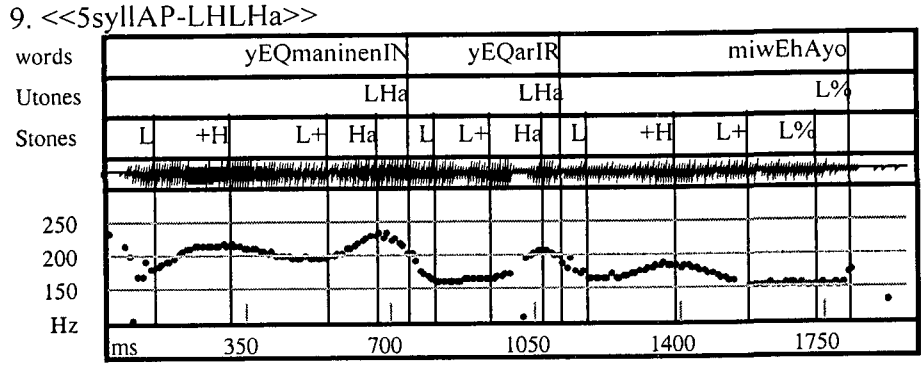
7. <<IPboundary-LH%>>



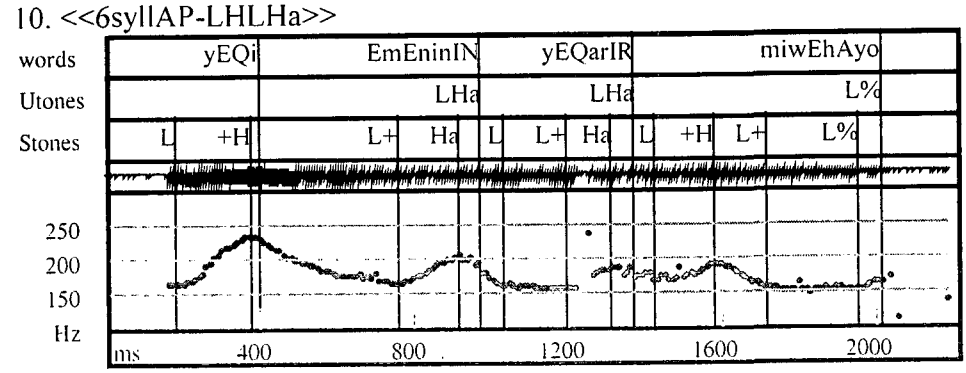
8. <<2syllAP-LHa>>



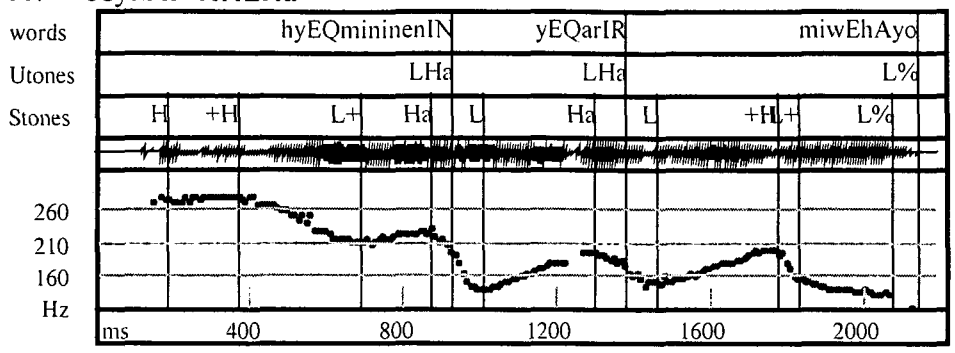
9. <<5syllAP-LHLHa>>



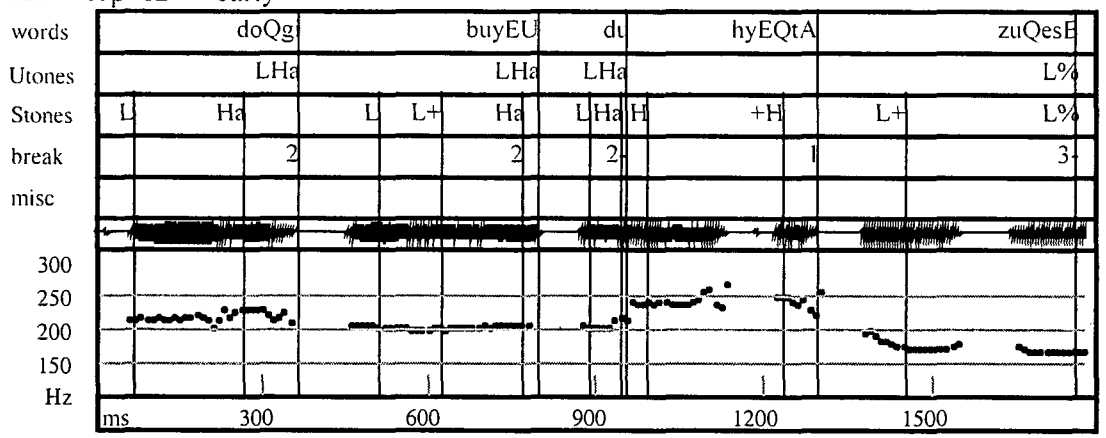
10. <<6syllAP-LHLHa>>



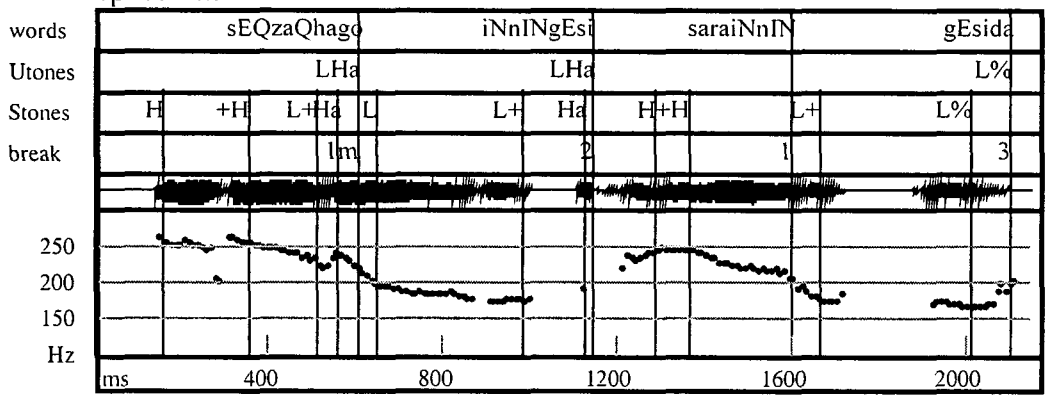
11. <<5syllAP-HHLHa>>



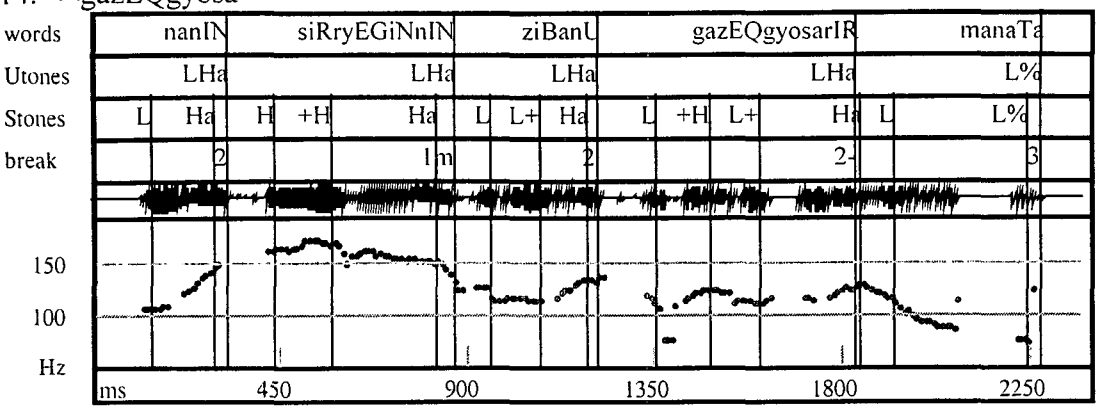
12. <<t1p1s2>>-early



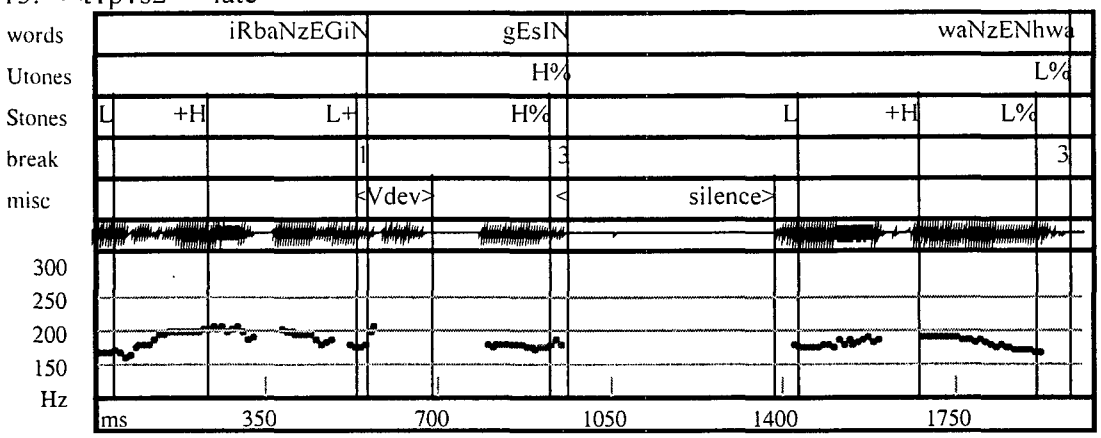
13. <<t1p2s8-1m>>



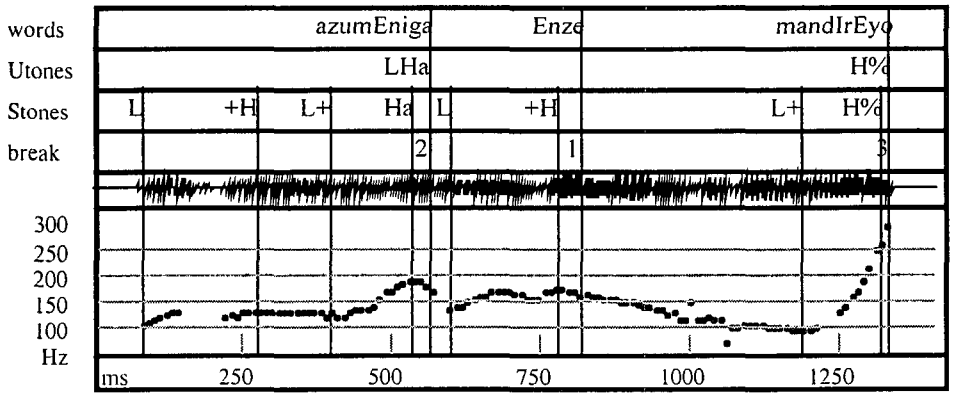
14. <<gazEQgyosa>>



15. <<t1p1s2>>-late

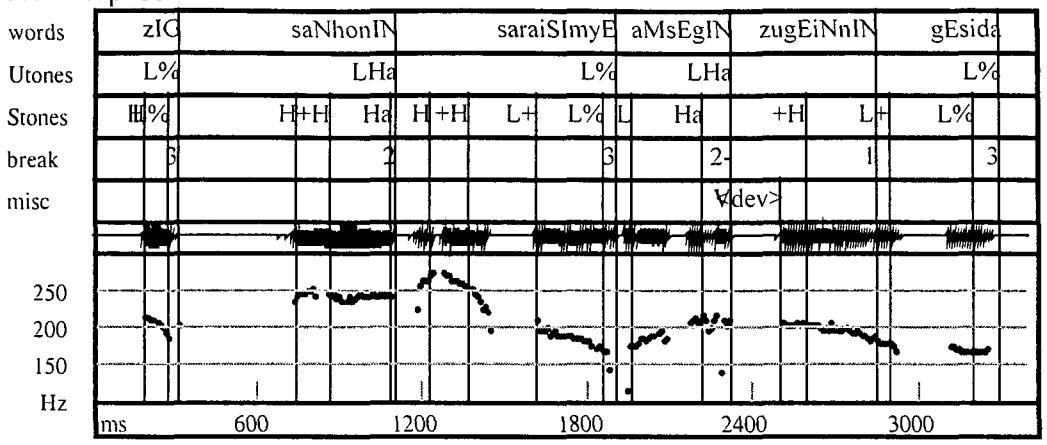


16. <<break-L8c3>>

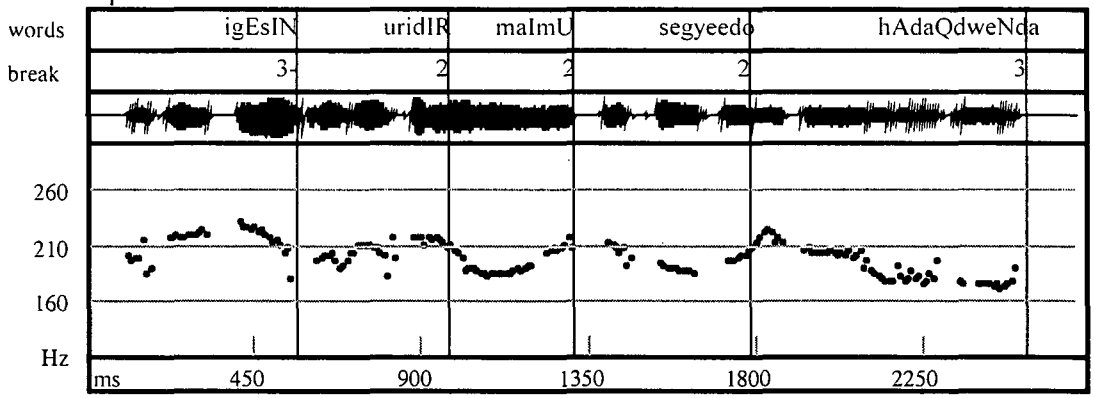


Appendix B (continue)

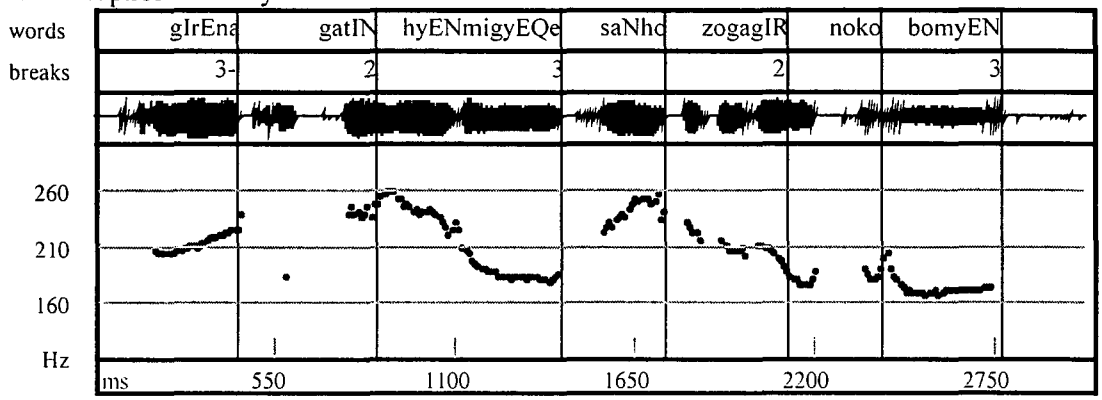
17. <<t1p2s6>>



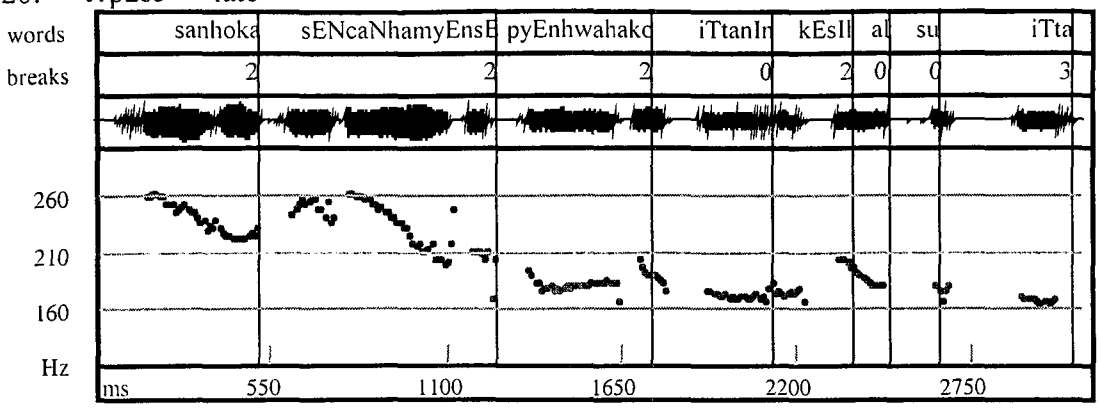
18. <<t1p2s10>>



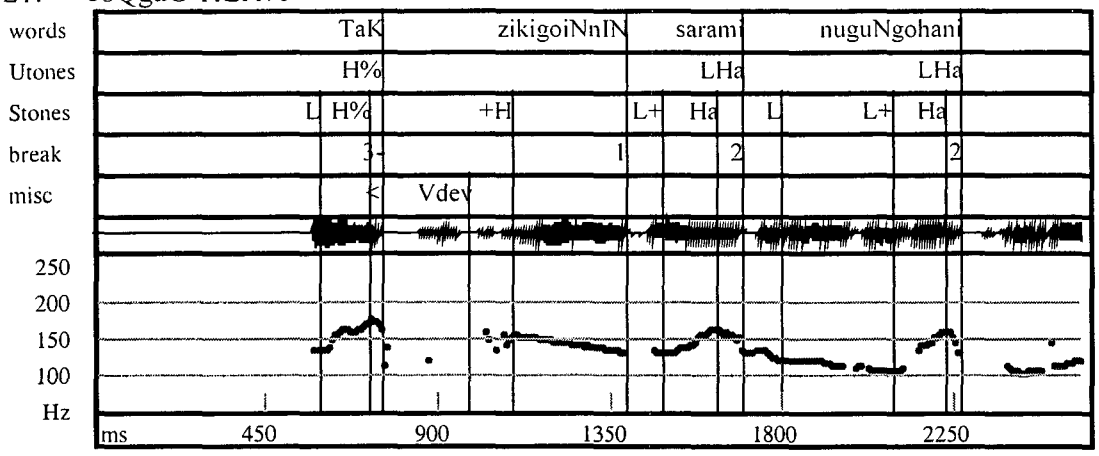
19. <<t1p2s5>>-early

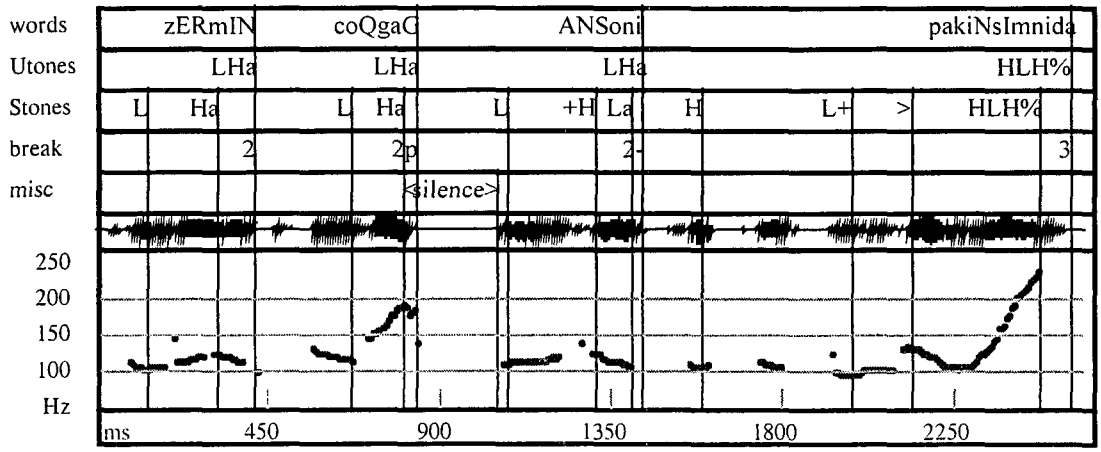


20. <<t1p2s5>>-late

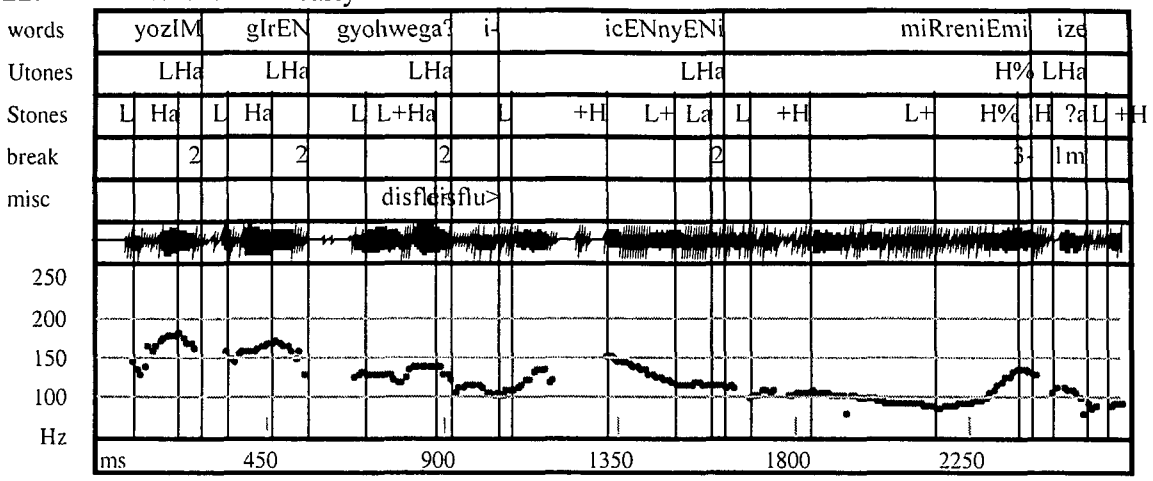


21. <<coQgaG-HLH%>>

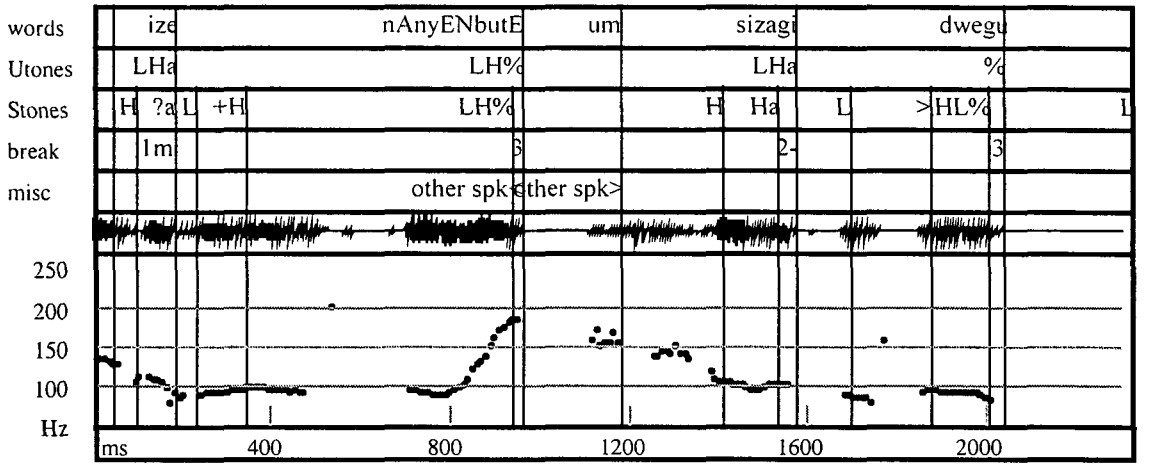




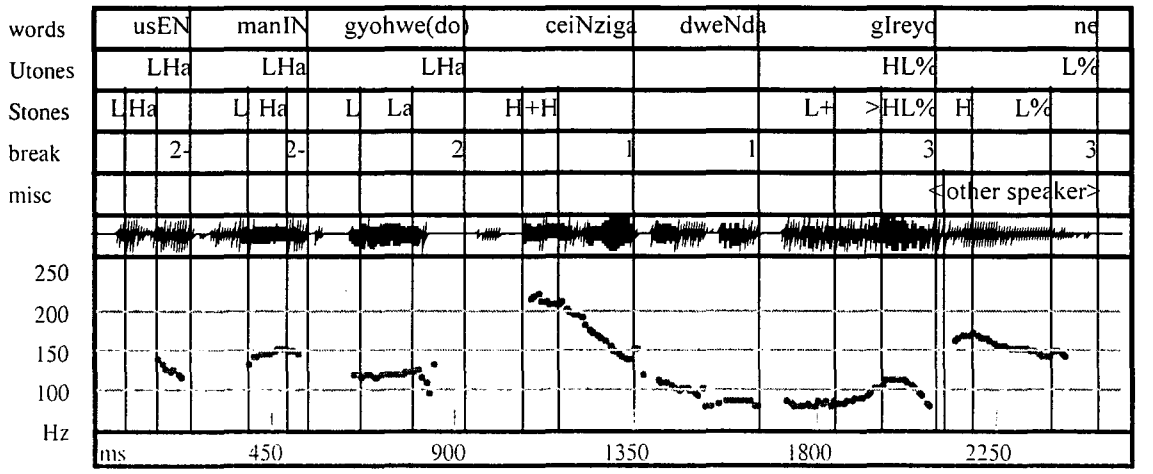
22. <<millennium>>-early



23. <<millennium>>-middle

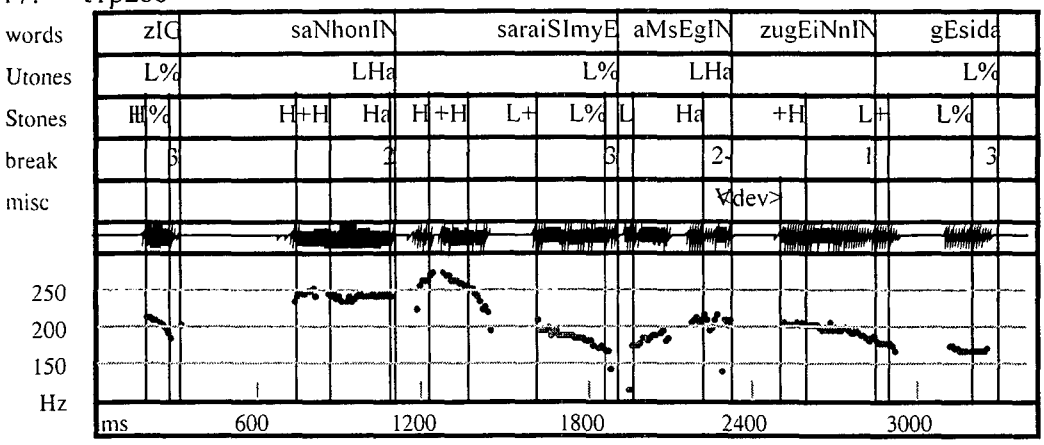


24. <<millennium>>-final

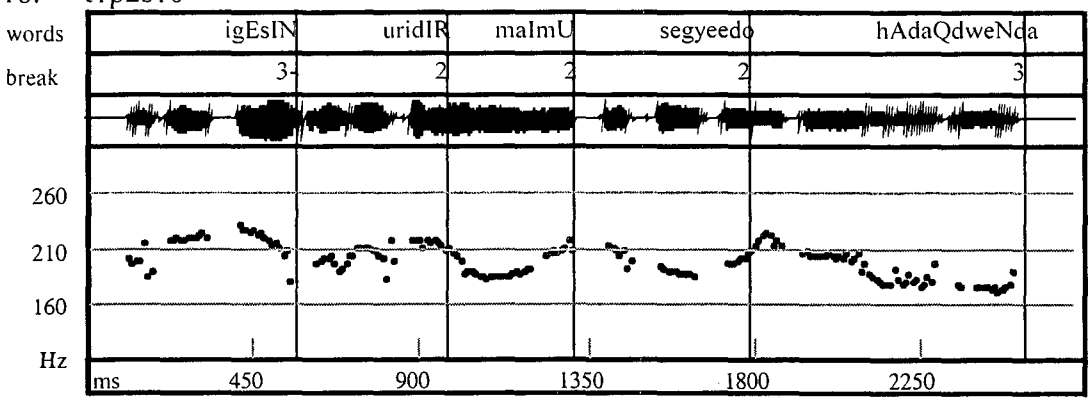


Appendix B (continue)

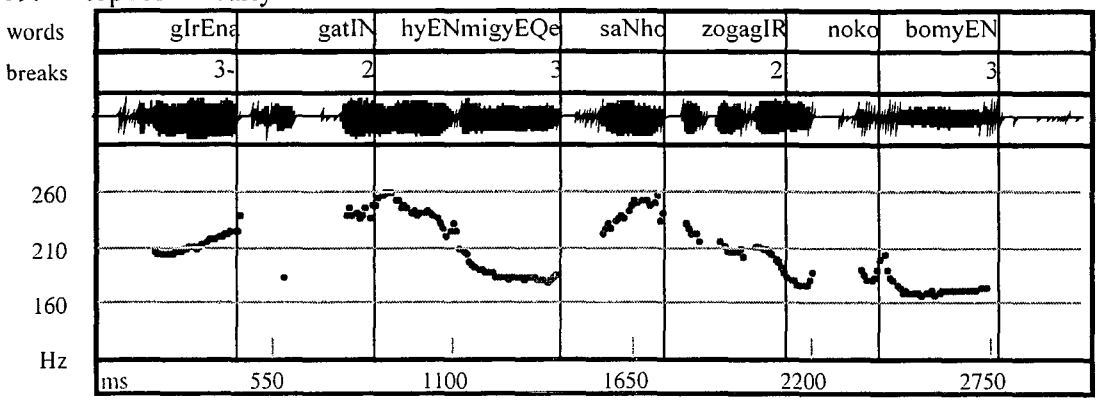
17. <<t1p2s6>>



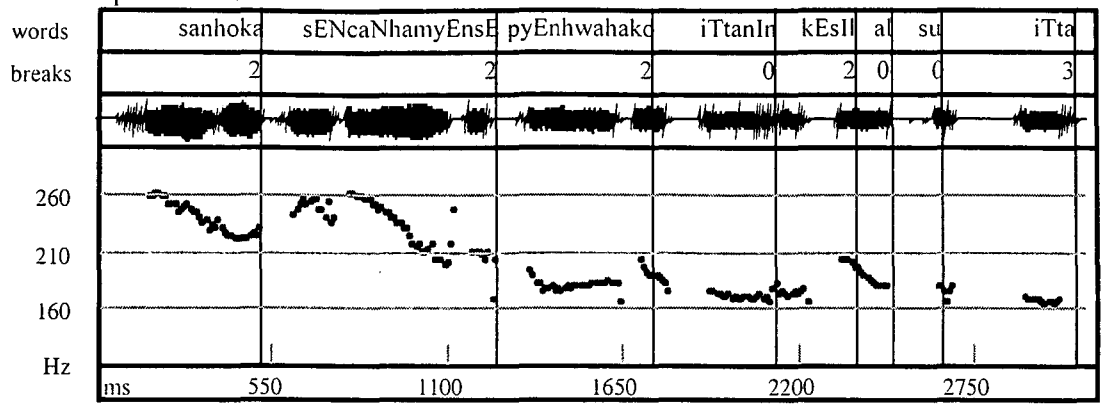
18. <<t1p2s10>>



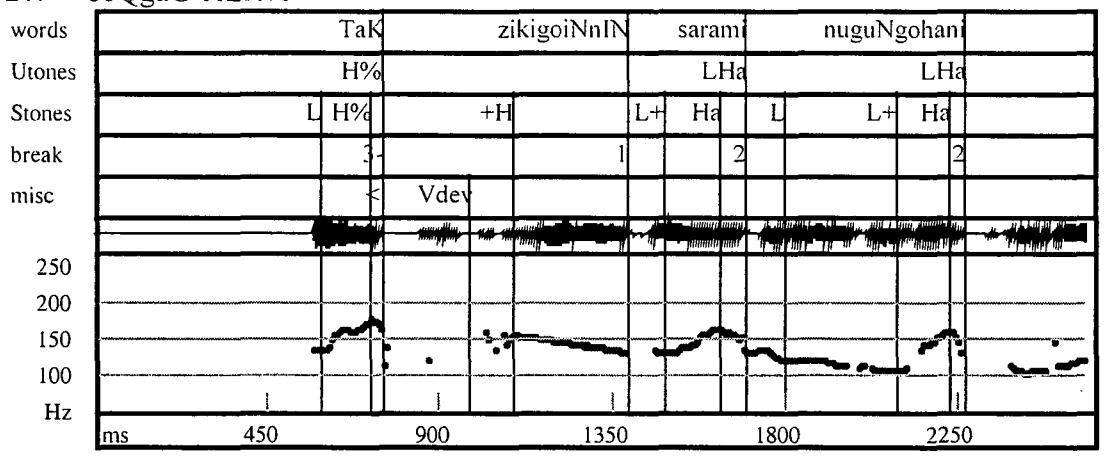
19. <<t1p2s5>>-early

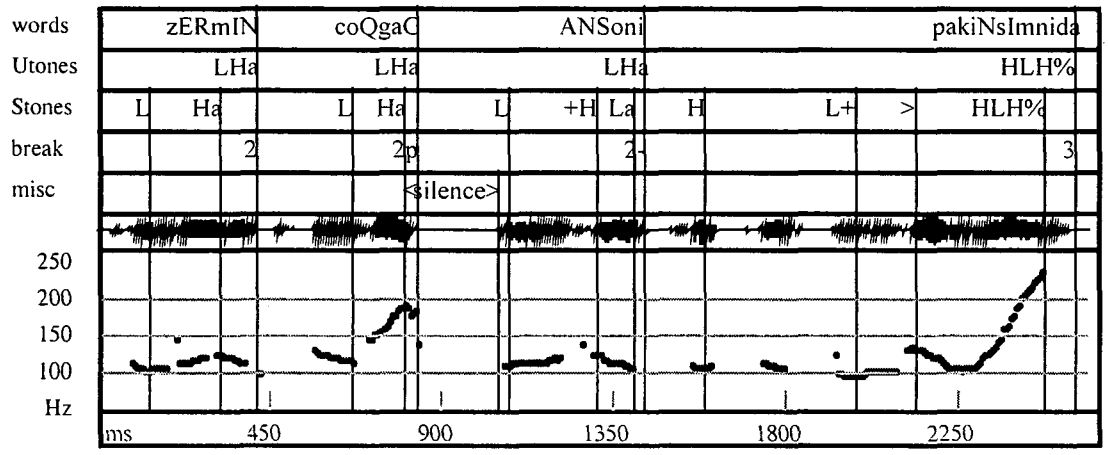


20. <<t1p2s5>>-late

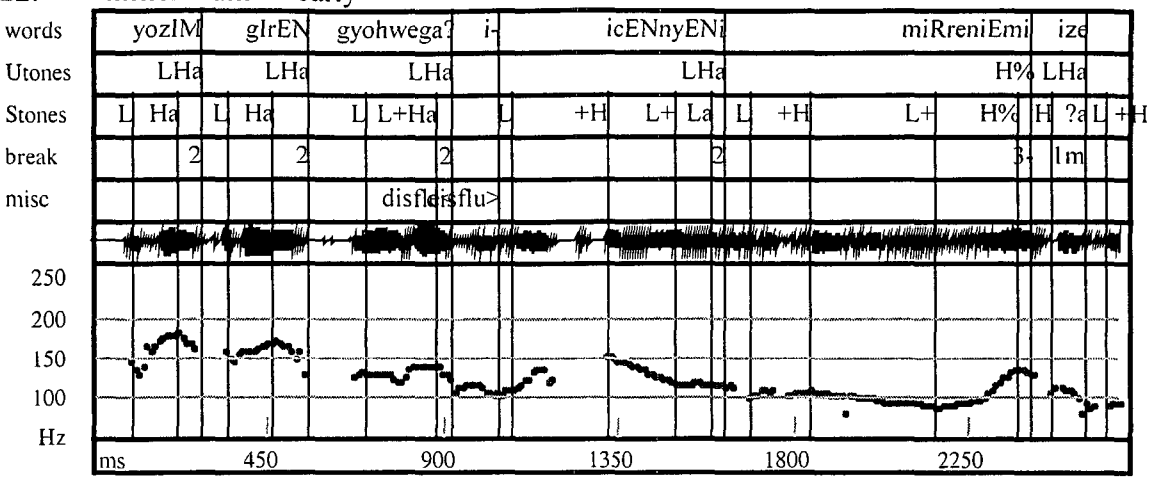


21. <<coQgaG-HLH%>>

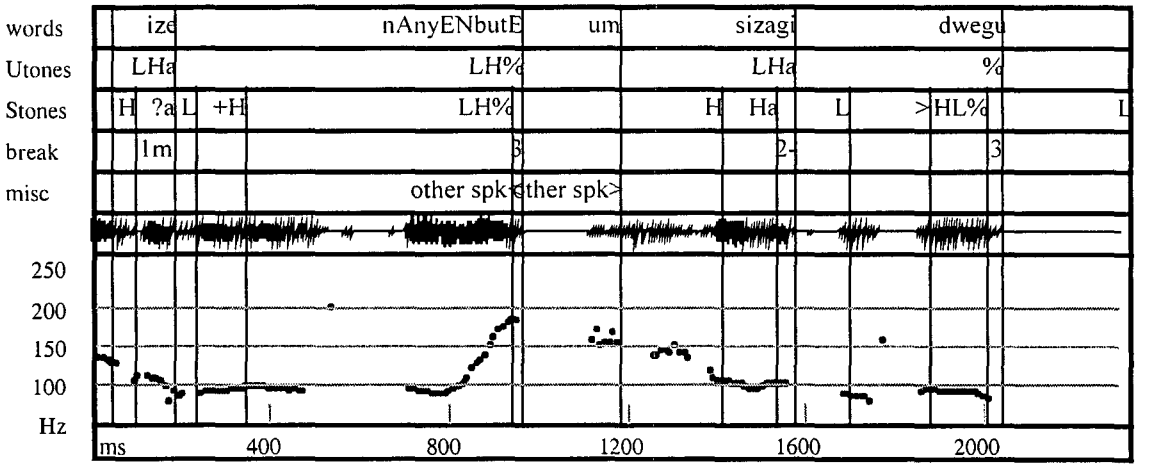




22. <<millennium>>-early



23. <<millennium>>-middle



24. <<millennium>>-final