FIN-ToBI Tones and Break Indices for Finnish

Jarmo Välikangas Department of Phonetics, University of Helsinki

3.1.2002

Abstract

ToBI (Tones and Break Indices) is a general notation system for spoken language. The basic idea behind it is to describe prosodic aspects such as intonational patterns and standardise the description of prosody. At adequate levels, ToBi system is adaptable for describing Finnish boundary tones, phrase accents and breaks.

1 Introduction

ToBI (Tones and Break Indices) system was originally developed [3] for the description of intonational features of English [10] and it has been succesfully localised for many languages, like Dutch, Greek, Spanish, Italian, Japanese, Cantonese, Pan-Mandarin, French, Serbo-Croatian and Swedish.

ToBI system associates speech signal (fundamental frequency contour) with textual information. The basic system uses 4 text tiers.

- 1. An ortographic tier. As the name says, this tier represents the ortographic form of words. In English, ortographic form is 'a straightforward transcription of all words in the utterance' [2].
- 2. A break-index tier. The break-index is a 5-step scale which describes how the breaks are interpreted. See Table 1.
- 3. A tone tier. Pitch events (as fundamental frequency contour) can be marked tones like intonational boundaries (phrasal tones) or accented syllables (pitch accents). Markings for phrasal tones are relative L (low) and H (high) and they are assigned at intonation or intermediate phrases. Intonational boundaries are marked with boundary tone '%'. For example, low final intonational phrase boundary tone is marked 'L%'. Intermediate phrase boundaries (accents) are marked with '-' when they are connected with intonational boundary and '+' when

Type	Description
0	clitilication
1	phrase-medial word boundaries
2	a strong disjuncture marked by a pause
3	intermediate intonation phrase boundary
4	full intonation phrase boundary

Table 1: ToBI break indices

they are not connected with intonational boudary. Tone target syllables are marked with '*'. Intermediate phrase boundary markings are usually combined like 'L+H*' (rising peak accent). All the unaccented syllables are left without markings.

4. A miscellaneous tier. Miscellaneous tiers can be used for any other text like comments or extra-linguistic information, for example.

2 Finnish prosodic and intonational features

2.1 Prosodic features

Prosodic features of speech may be described in terms of energetic, tonal and temporal phenomena. These are perceived as loudness (volume), melody (pitch) and duration (including pauses).

2.2 Intonational Features

2.2.1 Stress

Finnish stress as an intonational feature is twofold comprising lexical stress and sentence accent. As Karlsson has described, Finnish lexical stress has a fixed place on the first syllable of word, indicating "there is a word junction before this syllable" [8]. This has been also noticed by Tuomainen et al when they concluded that word stress is the primary cue to word boundaries in Finnish. They also noticed that the change in fundamental frequency contour is the primary cue for the perception of word boundaries [12]. Within the words, lexical stress describes relationship or prominence between syllables and there are two kind of syllables, namely strong and weak (see Table 2). On the segmental level, strong syllables tend to have more energy, sound segments are longer and the vowel quality is more distinct.

Although word stress is a built-in feature of Finnish speech, different linguistic functions cause sentence-level features to dominate and stress should be judged in the context of the whole utterance (Table 2). Iivonen et al have categorised sentence accent as three types which are (1) **main stress**

	me	nem	me	ju	nal	la	muu	laan
word stress	x			х			х	
sentence accent				х			х	

Table 2: <We travel to Muula by train.> (<Menemme junalla Muulaan.>) Word stress and sentence accent, strong syllables are marked 'x'.

(neutral, contrastive or emphatic), (2) secondary stress and (3) nonstressed [4]. Stress as an intonational feature can manifest itself by all three prosodic features and describe the relative prominence in an utterance.

2.2.2 Phrasing

The most natural idea of phrasing is a breath group. Articulation and phonation follow breathing which provides phrasing with natural start and end, consecutively. However, spontaneous speech can have different units when speaker decides to use prosodic repertoire for communicative aims by using pauses for deviding speech flow into smaller phrases. See Figure 2.

2.2.3 Tune

The subglottal pressure of speech production gets lower towards the end of speech utterance. This is perceived as a lowering declination i.e. lowering melody line. In terms of prosody, lowering melody line is typically connected with standart Finnish declaration sentence. Finnish clause final features are supposed to be the lowest point of fundamental frequency contour and after the last word stress the vowel quality is often characterised as laryngalised (creaky) voice [5, 4]. No specific final fundamental contour for questions has been addressed in spoken, spontaneous Finnish, but speakers are able to use fundamental frequency for contrast and emphasis which may sometimes be understood as questions [1, 4]. The rising fundamental frequency contour in the sentence final position, "dot-intonation", is sometimes used to express continuation, too [4]. Hirvonen and Iivonen have claimed that the interrogative function of Finnish is executed as initial global high fundamental frequency [5, 6]. In the context of style research, livonen connects the initial high fundamental frequency also with information focusing. In news-style read speech, new topics (new newsitems) are highlighted by the use of higher initial fundamental frequency [7].

3 Tobi EXAMPLES

3.1 Tones

- Interphrasal accents Välimaa-Blum has introduced pierrehumbertian pitch accent analysis in Finnish and the results are relevant starting point for ToBI notation. Välimaa-Blum has found two complex accents 'L+H*' and 'L*+H' and two boundary tones (H% and L%) [11]. Complex accent 'L+H*' and boundary tones are illustrated in Figure 1 below.
- L^*+H Prominent low syllable and ascending fundamental frequency contour denotes late accentuation, displacement of normal accent from the first syllable. No example here.
- L+H* Välimaa-Blum describes this kind of complex accent as 'neutral contour [where] the accents form a gradually declining pattern of L+H* accents'. See Figure 1.



Figure 1: <The general grade for the food good, points for environment zero.> (<Yleisarvosana ruoasta hyvä, ympäristö pisteitä nolla>) Complex accents 'L+H*' and fundamental frequency contour of Finnish declarative sentence.

Intonational boundary tones

- %H As noted before, Iivonen et al has claimed that initial high fundamental frequency contour can be interpretated as a question or as an introduction to the new thematic material [7].
- %L Low initial boundary tone seems to be (no examples here) strongly connected with coordinating and subordinating conjunctions like <mutta> (but) and <sillä> (because). No examples here.

- H% Rising boundary tone can be used in interphrasal position (see Figure 2), but is very seldom used in the sentence final position, too. No examples here, see section 2.2.3 for references.
- L% The lowest point of fundamental frequency contour is considered as typical end of Finnish declaration sentence. See Figures 2, 3 and section 2.2.3.



Figure 2: <Eila went to pick berries, Noora went to the mill, Leevi went fishing and the others went swimming.> (<Eila meni marjaan, Noora meni myllylle, Leevi meni ongelle ja muut menivät <u>uimaan</u>.>) Figure 2 describes how last words of intermediate phrases have boundary tones H% and final intonational boundary tone L%. <u>Underlined</u> words refer to the words with boundary tones.

Ladd has interestingly suggested that in some cases, complex phrase accent seems to be a result of natural declination [9]. This implies different interpretation of accents in Finnish, too. There's not always 'L' accent but a normal lowering declaration which is followed by simple accent 'H*'. This would bring down complex accents to simple ones and reserve possible 'L' or 'L*' occurrences for marked cases with accent. As noted before, Välimaa-Blum has interpreted 'L*+H' as late accent where primary stress is displaced from the first syllable of the word [11]. Since there are no such phonologically distinctive features in Finnish, simple 'H*' can also cover complex 'L+H*' cases. Despite the absence of accent 'L', phonetically prominent features can be communicated and mapped with other tiers.

3.2 Break Indices

0 The weakest form of break can be interpretated as clitilisated morphing of two syllables. For example, $\langle \text{onko se...} \rangle$ ($\langle \text{is it...} \rangle$) can be reduced to $\langle \text{onkse...} \rangle$, omitting $\langle \text{o} \rangle$ at the word boundary. Clitilisation

File: /work/docs/esitykset/tobi/mv_0001.wav Page: 1 of 1 Printed: Thu Jan 03 18:54:48



Figure 3: <The general grade for the food good, points for environment zero.> (<Yleisarvosana ruoasta hyvä, ympäristö pisteitä nolla>). An example of ToBI annotation with tonal, break and ortographic tiers.

doesn't comment differencies between standart written forms and estimated transcription form, nor abbreviations, for example. A typical instance is free variation of allophones, like in Figure 1 where written form <ruoan> is changed to phonetic form [ru:an].

- This stage of break indices describes normal word boundaries. In Finnish, juncture may need more detailed description, like <tämä nero> -<tämän ero> (<this genius> - <difference of this>) as well as sandhi and initial dubbling, too.
- 2 'A strong disjuncture marked by a pause or virtual pause', as the definition claims, is open to various interpretations. In spontaneous speech the distinction between strong and weak disjunctures, whether rhytmic or tonal, is unclear.
- 3 Within full intonational phrase, intermediate intonation phrase boundary is the single phrase tone from the last pitch accent to the boundary. See Figure 3 where boundary tone H% precedes break 3.
- 4 Full break is strongly connected with tonal boundaries and it is rather unambiguous as such.

The combination of boundary tones, tone accents and breaks are illustrated in Figure 3. The first tier is a tone tier, where tonal events ('H*', 'L%') are marked at the prominent places of the fundamental frequency contour. The next tier lists break indices (1, 3, 4) and connects them with the word boundaries and ortographic tier.

4 Discussion

A suggestion for basic tonal paradigm of Finnish intonation is

$$\left\{\begin{array}{c}\%L\\\%H\end{array}\right\} \left\{\begin{array}{c}\mathrm{H}^{*}\\L\%\end{array}\right\} + \left\{\begin{array}{c}H\%\\L\%\end{array}\right\}$$

where optional initial boundary tone '%H' or '%L' is followed by one or more intermediate phrase accents 'H*' and final intonational boundary tone 'H%' or 'L%'.

The least unambiguous break indices are 1 and 4 which describe interphrasal word boundaries and intonational phrase breaks. More complicated break indices, such as 2, are unclear and they also need more research in the context of filled pauses and discontinuations.

From the phonetic point of view, the basic problem of ToBI is the relativity of notation. The mark 'H*', which stands for phrase accents, does not really tell much about the nature of phonetic substance. There may be a twist in the fundamental frequency contour, to some extent, but 'H*' doesn't expose where the slope starts and where it ends. In the more global perspective of fundamental frequency contour, 'H*' accent looks different in the initial position of the utterance than in the final position. In addition to this, in the realms of phonetics, accents may involve other prosodic repertoire than fundamental frequency contour. Prominent features of speech may be carried out by means of energy and temporal arrangements, too.

Behind the scenes, there seems to be many points of interest which ToBI brings out. A strictly phonetic view requests more detailed inspection, but on the other hand, discourse analysis would happily connect final boundary tone with relevant place for turn-taking, for example. Furthermore, tones, accents and breaks are alluring gateway to other linguistic functions like information focus or speech acts.

References

- Lehessaari A.-L. Alkoholin vaikutus puheen prosodiikkaan. PhD thesis, University of Helsinki, 1996.
- [2] S. Baumann et al. Tobi a phonological system for the transcription of german intonation. In Proc. Prosody 2000: Speech Recognition and Synthesis Workshop, 2000.
- [3] Mary E. Beckman et al. Guidelines for tobi labelling. Technical report, The Ohio State University Research Foundation, 1993, version 3, March 1997.

- [4] Iivonen A. et al. Puheen intonaatio. Helsinki: Gaudeamus, 1987.
- [5] P. Hirvonen. Finnish and English Communicative Intonation. PhD thesis, University of Turku, 1970.
- [6] A. Iivonen. Kommunikatiivisten aktien ja puhetyylien prosodiasta. In Tuomainen J. Ojala S., editor, 21. Fonetiikan päivät, pages 1–18, 2001.
- [7] A. Iivonen et al. Prosodic characteristics in english, finnish and german radio and tv newscasts. In Klippi A. Iivonen A., editor, *Studies in Logopedics and Phonetics*, number 6 in Publications of the Department of Phonetics, pages 47–68. Department of Phonetics, University of Helsinki, 1996.
- [8] F. Karlsson. Yleinen kielitiede. Helsinki: Yliopistopaino, 1994.
- [9] R. Ladd. Prosody : theory and experiment : studies presented to Gösta Bruce, volume 14 of Text, speech and information technology, chapter 2, pages 37–51. Kluwer Academic Publishers, 2000.
- [10] J. Pierrehumbert et al. Plans and Intentions in Communication and Discourse. SDF Benchmark Series in Computational Linguistics. MIT Press, 1990.
- [11] Välimaa-Blum R. A pitch accent analysis of intonation in finnish. Ural-Altaische Jahrbücher N.F., 12:82–9, 1993.
- [12] J. Tuomainen et al. Fundamental frequency is an important acoustic cue to word boundaries in spoken finnish. In Ohala J.J. et al., editor, *Proceedings of the 14th International Congress of Phonetic Sciences*, pages 921–923, August 1-7 1999.