

The Manifestation of Linguistic Information in Prosodic Features of Finnish

*Hansjörg Mixdorff**, *Martti Vainio***, *Stefan Werner**** and *Juhani Järvikivi****

*Faculty of Computer Science, Berlin University of Applied Sciences, Germany
mixdorff@tfh-berlin.de

**Department of Phonetics, University of Helsinki, Finland
martti.vainio@helsinki.fi

***Department of Linguistics, University of Joensuu, Finland
stefan.werner@joensuu.fi; juhani.jarvikivi@joensuu.fi

Abstract

The current paper presents a preliminary study of Finnish prosody using a quantitative model. In a small corpus of segmentally identical utterances the focus, the sentence mode and the place of the phrase boundary were systematically varied. Following the rationale of MFGI, the Mixdorff-Fujisaki Model of German intonation, the phonetic properties of the contrast-lending intonemes were examined. These properties differ strongly from those of German: Although the place of narrow focus is marked by higher accent command amplitudes, tone switches on pre-focal and post-focal accents are never completely deleted. Yes/no questions with final rises are perceived as less natural by Finnish listeners than those marked by a question particle. Non-terminal utterances are not marked by clear rises, but rather by a fall that does not reach the lower edge of the speaker's *F0* range as in declaratives.

1. Introduction

In Text-to-Speech systems the application of quantitative models for the prediction of prosodic parameters is commonplace. In contrast, studies on the phonological systems of prosody are still very often based on impressionistic approaches. This is also the case for the bulk of works on Finnish prosody. One reason might be that commonly employed quantitative models have been originally developed for intensely studied languages, such as English and German, and are therefore based on assumptions concerning the particular language for which they were created.

The present study applies the methodology developed by Mixdorff for German to the analysis of Finnish. The Mixdorff-Fujisaki model of German Intonation (MFGI) utilizes the Fujisaki formula [2] for parametrizing a given *F0* contour. Resulting parameters are then related to intonational events, so-called 'intonemes', which define intonational contrasts in a language. Since the Fujisaki model is per-se production-oriented and physiologically motivated, it is applicable to all languages, though the functions of model components may vary depending on the particular language. The intonational description in MFGI distinguishes between three major types of intonemes: Declarative-final falls

(Information Intonemes $I\downarrow$), question-final rises (Contact Intonemes $C\uparrow$) and non-terminal rises (Non-terminal Intonemes $N\uparrow$). The phonetic properties of these intonemes are quantified using the Fujisaki model by aligning accent command onsets and/or offsets with accented syllables.

2. Finnish Prosody

Finnish has a fairly free word order, rich morphology with suffixation, and enclisis, as well as a relatively large number of grammatical cases (15). According to a typology suggested by Donegan and Stampe [3], this is a set of features which is typical for an intonationally falling language. Due to suffixation and enclisis, the lexical morphemes are at the beginnings of the words leading to a state in which the lexical stress is invariably on the first syllable of the word. Finnish is also a quantity language with phonetically long vowels and geminate consonants; the vowels are always in the same syllable, whereas the long consonants always have a syllable boundary within them. The long sounds are on the average about twice as long as the short ones and they can occur in stressed and unstressed syllables alike.

The basic intonation shape in Finnish is a falling shape with an accent on basically all content words. Finite verbs are usually less prominent than nominals and are sometimes altogether unaccented. According to Välimaa-Blum [4], Finnish can be analyzed as having two pitch-accents ($L+H^*$, L^*+H) and two boundary tones ($L\%$, $H\%$).

Finnish questions are typically marked only by lexical means (by interrogative particles) which has lead most researchers (see Iivonen [5], for instance) to the conclusion that there is no interrogative intonation in Finnish. There is some evidence that, in general, questions start with higher fundamental frequency, but the final rise in questions has been considered an alien feature in Finnish. The final rise can, however, be found in more colloquial speech in echo-questions. Continuation is typically signalled by level intonation and finality by a sharp fall into the bottom of the speakers fundamental frequency range, which usually causes a creaky or whispery voice during the last unstressed syllables of an utterance.

3. Speech Material and Method of Analysis

In order to examine the prosodic properties of Finnish, a small corpus was created. The sentence "Menemme laivalla Lemille" - "We go by boat to Lemi" was read by a single native speaker of Finnish seven times with several different connotations. The conditions examined are listed in Table 1 stating the sentences underlying the utterances, with narrowly focused items set in bold type. The right column contains comments concerning the cases.

As mentioned before, questions in Finnish are usually either marked by interrogative pronouns or the particles 'ko' or 'kö'. For this reason, Cases 4-7 represent relatively unidiomatic renditions of intonationally marked echo-questions. In the language of young people, however, intonationally marked questions can be observed, presumably under the influence of English which is the first foreign language in Finland. Furthermore, thanks to the flexible Finnish syntax, narrowly focused items are preferably placed at the end of an utterance.

In a listening test all utterances were checked for consistency with the intended message. Twelve native speakers of Finnish were asked to determine the most prominent word in each utterance and rate the naturalness of the samples on a scale from 1 to 5. Of the single-phrase utterances, on the average Cases 1-3 and 7-10 received naturalness ratings close to or above 4.0. In contrast, the intonationally marked questions (Cases 4-6), were rated rather poorly and reach only between 2.3 (narrow foci) and 2.7 (broad focus). Case 10 with the narrowly focused item 'laivalla' in the default final position (4.24) is slightly preferred over Case 8 with 'laivalla' in medial position (4.19). Except for Cases 12 and 16 (2.5), the two-phrase-utterances were also rated close to and above 4.0. The poor ratings for the former cases can be explained by a semantic

inconsistency between the focused item and the context supplied in the second phrase, e.g. "We go to **Lemi** by boat, because the bridge is broken." This result nicely shows that the listeners were conscious of the prosodic marking and judged whether or not it matched the context implicitly provided.

Utterances that failed to convey the intended focus were excluded from the following prosodic analysis. The *FO* contours of all examples were extracted at intervals of 10 ms and parametrized using an automatic procedure for determining the Fujisaki parameters [6]. Parameter configurations yielded were checked and if necessary corrected. The constants Fb , α and β were set to 55 Hz, 2/s and 15/s. Syllabic boundaries were marked by listening and inspection of the speech waveform and the broad-band spectrum. The duration contour was then calculated by determining the syllabic z-score in the log duration domain. The syllabic duration means were calculated by adding the duration means of phones in a particular syllable. Phone means were available from a larger database created by the speaker.

4. Results of Analysis

4.1. Fujisaki Parameter Configurations

Figure 1 shows examples of analysis. Each panel displays from top to bottom: The speech waveform, the extracted (+ signs) and model-generated *FO* contours (solid line), the duration contour in terms of the syllabic z-score indicated as horizontal lines of the length of the respective syllables, the SAMPA transcription, and the underlying Fujisaki parameter configuration of impulse-wise phrase commands and step-wise accent commands.

Table 1: List of cases examined in the study. Narrowly focused words are set in bold type

Case	Sentence	Comment
1	Menemme laivalla Lemille.	In principle, Finnish syntax is very flexible. Therefore, narrowly focused items can be placed in the (default) utterance-final position.
2	Menemme laivalla Lemille .	
3	Menemme laivalla Lemille.	
4	Menemme laivalla Lemille ?	Echo-questions in Finnish are usually marked by particles 'ko' or 'kö' (Cases 7-8), but intonationally marked questions (Cases 4-6), though in traditional theory regarded as unidiomatic, can be observed in certain language communities.
5	Menemme laivalla Lemille ?	
6	Menemme laivalla Lemille ?	
7	Menemmekö laivalla Lemille ?	The regular yes/no questions in Finnish are marked by particle 'ko' or 'kö'
8	Menemmekö laivalla Lemille ?	
9	Menemmekö laivalla Lemille ?	
10	Menemmekö Lemille laivalla ?	Usually a narrowly focused item is placed utterance-finally.
11	Menemme Lemille laivalla, koska silta on rikki.	Context: "... because the bridge is broken."
12	Menemme Lemille laivalla, koska silta on rikki.	
13	Menemme Lemille laivalla , koska silta on rikki.	
14	Menemme laivalla, Lemille ei pääse muuten.	"We go by boat, to Lemi it's not very far."
15	Emme tarvitse autoa koska menemme laivalla Lemille.	Context: "We don't need a car because..."
16	Emme tarvitse autoa koska menemme laivalla Lemille .	
17	Emme tarvitse autoa koska menemme laivalla Lemille.	

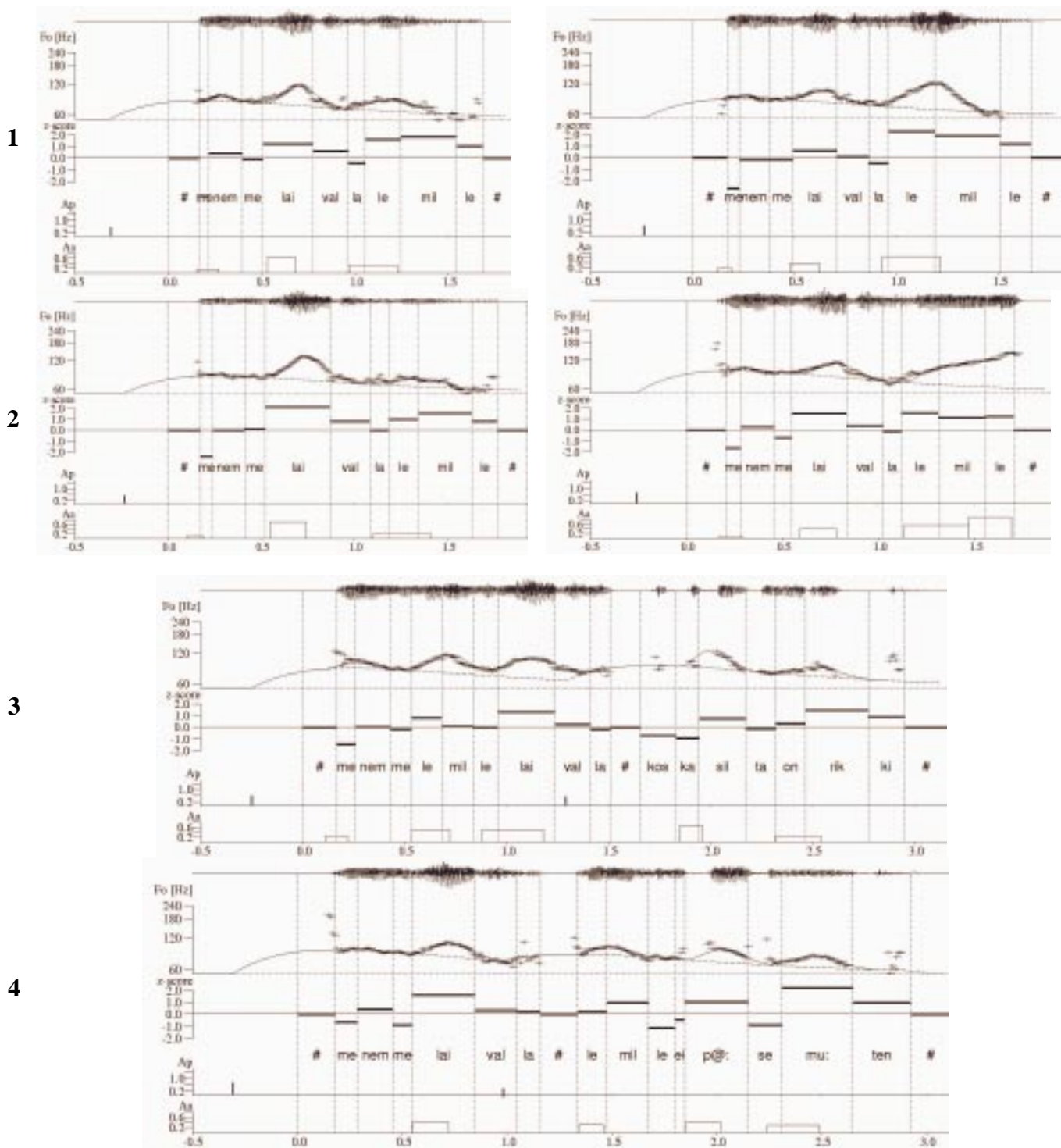


Figure 1: Examples of analysis. Case 1: declarative, broad focus (row 1, left); Case 2: declarative, narrow focus on 'Lemille' (row 1, right); Case 3: declarative, narrow focus on 'laivalla' (row 2, left); Case 4: echo-question, broad focus (row 2, right); Case 11: two-phrase utterance (row 3); Case 14: two phrase-utterance, phrase boundary after 'laivalla' (row 4).

The left panel in row 1 shows Case 1, a statement with broad focus. If we compare this example with Case 2 (narrow focus on 'Lemille', row 1 right), and Case 3 (narrow focus on 'laivalla', row 2 left), we see that every word in the utterance was assigned an accent command aligned with the first (default accent) syllable of the word, although the

accent on the verb 'menemme' is neglectably small. It can also be seen that the accent command on 'Lemille' is rather long and extends over the two syllables [la] and [le]. Considering the timing of the accent commands, all accents belong to the L+H* type [4]. Perceptually, however, due to their early timing with respect to the accented syllable, even

Table 2: Mean syllable durations in ms for selected contexts averaged over seven repetitions. Note that the word order of "laivalla Lemille" is reversed in Cases 11 and 13.

Case	me	nem	me	lai	val	la	le	mil	le
1	115	197	95	249	171	113	201	280	234
2	91	186	101	243	172	154	243	375	190
3	102	210	118	351	223	119	196	286	178
7	112	159	100	244	172	110	191	276	249
8	110	160	112	330	199	114	195	298	187
9	110	156	108	234	173	144	239	365	195
11	106	173	94	276	181	181	168	156	131
13	89	168	98	341	211	196	162	160	178
14	104	169	125	285	218	312	149	183	75
15	110	181	98	234	169	110	198	259	193
17	108	172	106	242	169	144	240	360	201

the pre-final accents have falling tone switches as opposed to the perceptually rising ones in German.

Compared with the broad focus condition, narrowly focused items exhibit higher accent command amplitudes and increased syllabic durations. The accent command amplitude *Aa* is almost doubled in the case of the utterance-final 'Lemille', and the accent syllable duration increased by 20%, whereas the increase in *Aa* is much smaller in the utterance-medial 'laivalla' which is rather marked by its stretched accent syllable (+40% compared with Case 1).

Another striking difference from German is the fact that pre-focal as well as post-focal tone switches are never completely deleted. Phrase boundaries are also realized in a different manner, as can be seen in rows 3 and 4 showing two-phrase utterances from Cases 11 and 14: At the phrase boundary, the *F0* contour falls after the accent syllable [lai] of the phrase-final 'laivalla' (accent type L+H* as well), and does not remain on a plateau that would justify the assignment of a boundary tone H%. Although the contour does not drop to the lower edge of the speaker's *F0* range, it does not exhibit a marked continuation rise either, but a slight upward movement that is aligned with the upcoming phrase command preceding the second phrase.

In the second phrase (Case 11), the accent command on the phrase-final 'rikki' starts early in the syllable 'on' and is longer than that on the phrase-initial 'silta'. The characteristic 'long' accent pattern on the utterance-final words, as well as the slightly different accent shape in pre-boundary position cannot be adequately described by Välimaa-Blum's ToBI definition, which would assign an 'L+H*' to all of these cases.

The 'unidiomatic' intonationally marked echo-question (Case 4, row 2 right) exhibits a gradual rise on the phrase-final 'Lemille' that extends over the syllables [le] and [mil] and is further enforced by an additional accent command. This pattern is very similar to what can be observed in German.

4.2. Syllable durations

Table 2 lists the mean syllable durations for selected cases averaged over seven repetitions. The instances where a syllable is narrowly focused are indicated by bold type. The influence of the narrow focus can be seen in Cases 3, 8 and 13 on the first syllable of 'laivalla' and, though less clearly, in Cases 2 and 9 for 'Lemille' (on both the [le] and [mil]

syllables). Interestingly, the narrow focus on 'laivalla' in context 17 is hardly indicated in the temporal structure at all.

Phrase-final lengthening can be discerned extremely clearly in Case 14 on the last syllable [la] of 'laivalla'. Furthermore the last syllable [le] of the non-final word 'Lemille' in the same sentence is much shorter if compared to the same syllable in phrase-final position (Cases 1 and 15, for instance). Word order also accounts for some temporal adjustments, e.g. 'Lemille's [mil] syllable is considerably shorter in Cases 11 and 13 where "laivalla Lemille" is inverted to "Lemille laivalla".

5. Discussion and Conclusions

The current paper presented a small study on the prosody of Finnish. In general, we have shown that the methodology of MFGI can be applied to this language. Lexical accent syllables are aligned with accent commands whose amplitude increases in the presence of narrow focus. Accentuation and pre-boundary effects lengthen a syllable whereas lexical accent syllables are shortened under pre-focal conditions.

Different from German, tone switches - at least in the read material analyzed in this study - are generally falling and do not undergo complete deletion in pre-focal and post-focal position. Due to the former observation, in principle, the direction of the tone switch is not a contrastive property of intonemes in Finnish, unless one counts in the 'unidiomatic' echo-questions. First informal inspection of conversational material, however, indicates that non-terminal rises and rising echo-questions indeed occur in talk. Hence they should be taken into account when defining the intonational system of Finnish.

Although the relatively limited amount of data examined only permits tentative conclusions, our results suggest that certain phenomena are difficult to capture by a ToBI style representation. Future research will concern the analysis of more spontaneous data and listening tests with prosodically manipulated resynthesized speech.

6. References

- [1] Mixdorff, H., *Intonation Patterns of German - Model-based. Quantitative Analysis and Synthesis of F0-Contours*. D.Eng. thesis TU Dresden, (<http://www.tfh-berlin.de/~mixdorff/thesis.htm>), 1998.
- [2] Fujisaki, H. and Hirose, K., "Analysis of voice fundamental frequency contours for declarative sentences of Japanese", in *Journal of the Acoustical Society of Japan (E)*, 5(4): 233-241, 1984.
- [3] Donegan, P. and Stampe, D., "Rhythm and the Holistic Organization of Language Structure", in Richardson, J. (eds.) Chicago Linguistic Society, pages 337-353, 1983.
- [4] Välimaa-Blum, R., "A Pitch Accent Analysis of Intonation in Finnish", in *Ural-Altische Jahrbücher N.F.*, vol. 12, pages 82-94, 1993.
- [5] Iivonen, A., "Intonation in Finnish", in Daniel Hirst and Albert Di Cristo (eds.) *Intonation systems - A survey of twenty languages*, pages 311-327, Cambridge University Press, 1998.
- [6] Mixdorff, H., "A novel approach to the fully automatic extraction of Fujisaki model parameters", in *Proceedings ICASSP 2000*, vol. 3, 1281-1284, Istanbul, Turkey, 2000.