CORRESPONDENCES BETWEEN KIM-BASED SYMBOLIC PROSODIC LABELS AND PARAMETERS OF THE FUJISAKI MODEL

Hartmut R. Pfitzinger
Hansjörg Mixdorff
Benno Peters

1 Introduction

The current study examines the relationship between prosodic labels assigned in the *Kiel Corpus of Spontaneous Speech IV* and parameters of the Fujisaki model of the production process of f0 (Fujisaki and Hirose 1984). By applying a quantitative model to f0 contours extracted from spontaneous speech of four speakers we make f0 peaks accessible to empirical analysis of the temporal alignment with segmental landmarks, e.g. the vowel onset of the accented syllable, and of the relationship between f0 interval and accent level class.

Tonal alignment is a research topic at least since Bruce 1977 who investigated the two distinctive prosodic patterns connected with primary stressed syllables in Swedish. He found temporal stability of the f0 contours. More specifically, "an early timing of the fall is interpreted as accent I and a late timing of the fall as accent II with a sharp shift of identification" (Bruce 1977, p. 147).

Recently, the question was investigated whether alignment differences across languages arise from a continuum of phonetic alignment realizations which fall within a single phonological category (Atterer and Ladd 2004) or are realizations of several phonological categories (Niebuhr and Ambrazaitis 2006).

Mixdorff and Fujisaki 2000 compared German ToBI labels with Fujisaki parameters. They found that tone labels were strongly correlated with accent commands, and the type of label (typically H*L and L*H) was clearly reflected by the onset and offset times of these accent commands. Only break indices (BI) 4 were almost unanimously aligned with phrase commands (96%), whereas only 55% of BI 3 boundaries were associated with phrase commands.

An alternative approach to the symbolic description of German intonation is the *Kiel Intonation Model* (KIM) (Kohler 1991a). The prosodic annotation scheme PROLAB (Kohler 1995) which is based on KIM comprises the functional and phonological distinction between early, medial, and late peaks (Kohler 1987; Kohler

1991b) and therefore motivates the current study since it also contains information on the level of word prominence.

Especially the question is addressed whether phonetic peak realizations in spontaneous speech which have been assigned to the three peak categories of KIM by labellers are aligned to segmental anchors in clearly different ways.

1.1 The Fujisaki Model

It has been shown for many languages including German, that the most important prosodic correlate, the fundamental frequency contour (f0), can be satisfactorily modelled using a quantitative model of the production process of f0, the so-called Fujisaki model (Fujisaki 2004). This model can be employed to infer a finite number of so-called commands from a given f0 contour which then in turn can be related to the structures and units underlying an observed utterance. The main components of the model are the phrase component which is associated with the slowly falling movement observed in a whole utterance, and the accent component which takes care of the fast changing movements that are associated with accented syllables, but also high boundary tones at phrase boundaries. The phrase component is derived from low-pass filtered impulse-wise phrase commands, defined by their magnitude Ap and onset time T0. The accent component is a low-pass filtered sequence of box-shaped accent commands which are defined by their onset time T1, offset time T2 and amplitude Aa.

The advantage of the f0 contour parameterization with the Fujisaki model is its data reduction to a small number of salient command values: as they represent smoothed and interpolated contours, it is possible to give unambiguous definitions of the points to be measured and analysed consistently and undisturbed by microprosodic events, which would be very difficult if not impossible especially in spontaneous speech.

1.2 Parameterization of spontaneous speech

Earlier works on spontaneous speech (Mixdorff and Pfitzinger 2005; Mixdorff, Pfitzinger, and Grauwinkel 2005) showed that it is generally characterized by shorter utterances and longer and more frequent pauses, as well as many non-terminal, rising phrase endings which complicate the estimation of the phrase component. In typical reading-style speech, the phrase component can be inferred from the f0 values in unaccented syllables. Since the automatic extraction (Mixdorff 2000) was developed on news-reading, the amount of manual post-processing is considerable. Besides, disfluencies and fillers render the analysis more difficult, since, for instance, disfluencies and speech sounds such as *"ummI really don't know … "*, can be connected in one and the same prosodic unit.

1.3 Video Task scenario

In the *Video Task scenario* (or *Lindenstrasse Daily Soap scenario*) of the *Kiel Corpus of Spontaneous Speech IV* (Kohler, Peters, and Scheffers 2006; Peters 2006), similar but non-identical video material was presented to two subjects sitting in separate quiet, sound-treated rooms. After the presentation, the subjects discussed differences and similarities of what they had seen and heard. They were not able to see each other and communicated via headphones and microphones placed in front of them. The processing of the speech data comprised the following steps: (1) orthographic transliteration including several special characters for e.g. breathing and pauses, (2) automatic generation of a phonematic transcription, (3) manual segmental labeling on the basis of the phonematic transcription, and (4) manual prosodic labeling using PROLAB. The corpus contains the transliteration and audio files (80 minutes, approx. 13,000 consecutive words) as well as time-aligned segmental and prosodic label files of 6 overlapping German dialogues (4 female and 2 male speaker pairs). Of the 12 speakers, 2 female and 2 male speakers each with approx. 160 seconds of speech were selected for the present study.

1.4 PROLAB

The concept of the prosodic labelling system PROLAB provides three different pitch peak synchronisations: *early* (f0 maximum located before the accented vowel),

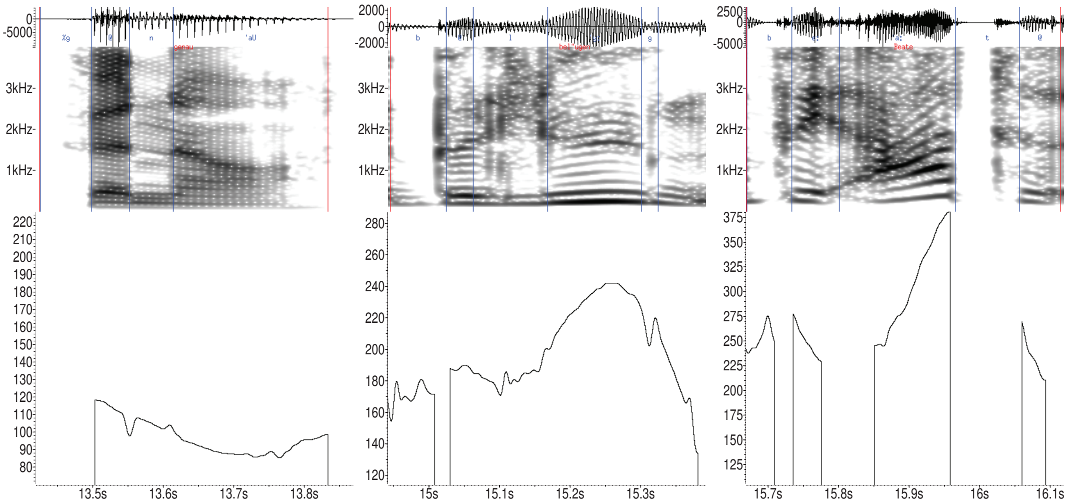

Fig. 1: Examples of an early (*left: l05uho51.wav*), medial (*middle: l02aha16.wav*), and late peak (*right: l02aha19.wav*) taken from the *Video Task scenario*. From *top* to *bottom*: Oscillogram, sonagram, and f0 contour. The y-scales of the f0-panels are adopted to the f0 range achieved during the respective signal file (*left:* male speaker, *middle/right:* female speaker).

medial (maximum within the accented vowel), and *late* (maximum late within, or after, the accented vowel). Fig. 1 shows three typical examples of an early, medial, and late peak taken from the *Kiel Corpus of Spontaneous Speech IV*. The prosodic labels are ‘)’, ‘^’, and ‘(’, respectively. PROLAB also provides four different accent levels: *reinforced accent*, *default accent*, *partial deaccentuation*, and *complete deaccentuation* symbolized with ‘3’, ‘2’, ‘1’, and ‘0’, respectively.

Labellers based their decisions on three criteria: (1) perceptual-phonetic assessment of the local f0 contour, (2) visual inspection of f0 values displayed in a window synchronous to the speech signal, and (3) functional-semantic categorization only within the range of a word whether an information is given (early), new (medial), or unexpected (late).

2 Method

The analysis of the relationship between the prosodic symbol layer and the signal layer requires (1) reliable extraction of fundamental frequency, (2) parameterization of f0 contours by means of the Fujisaki model, and (3) assignment of accent commands to prosodic symbols.

2.1 Fundamental frequency extraction

The f0 contours were extracted at a time step of 10 ms using four different algorithms: (1) *get_f0* of *ESPS waves* (Talkin 1995), (2) the autocorrelation-based f0 algorithm built-in in *praat* (Boersma 1993), (3) a time-domain pitch detector (Schaefer-Vincent 1982; Schaefer-Vincent 1983), and (4) a frequency-domain pitch detector based on harmonicity rules *mhspitch* (Scheffers 1983). The detection results were automatically compared and in case of differences the most accurately extracted f0 values were manually selected.

2.2 Fujisaki model parameter estimation

The parameterization followed the general guideline hitherto applied to reading-style speech, namely that accented syllables as well as high boundary tones should be associated with accent commands, and major phrase onsets (especially after each pause) should trigger a new phrase command. After a first pass with the automatic extraction algorithm all utterances were checked, and if necessary corrected using the interactive FujiParaEditor (Mixdorff 2009). As can be seen in Fig. 2, although the automatically estimated parameters yield a good overall approximation of the f0 contour, the last phrase command at 13.99 s is definitely too large, as a small accent command is missing on top of it. Note also the very small and short accent command at 13.53 s which is obviously not associated with an accented syllable

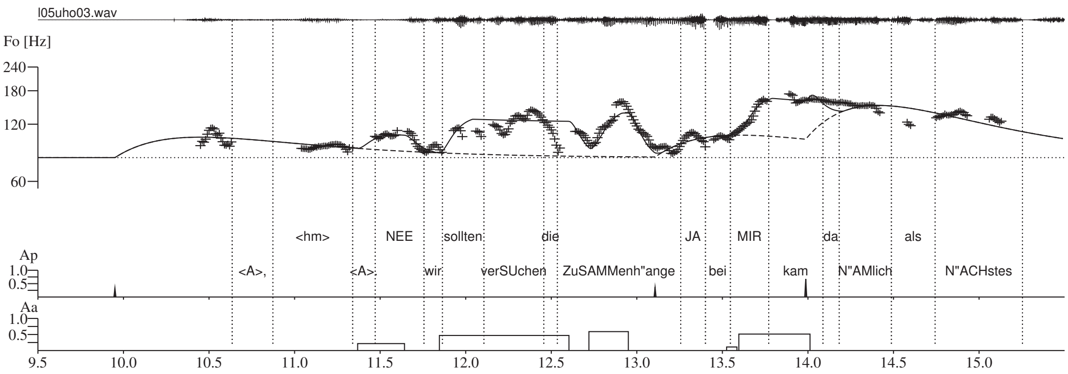

Fig. 2: Example of the automatic fundamental frequency contour modelling of the utterance: "Hm, nee wir sollten versuchen, die Zusammenhänge … ja bei mir kam da nämlich als nächstes …" — English translation: "Um, no, we should try to (find) the correspondences … yes, in my case the next was namely …". Accented syllables are set in capital letters.

but rather caused by micro-prosodic fluctuations in the f0 contour. Fig. 3 shows the same utterance after manual post-processing with the FujiParaEditor.

The labeller who performed the manual post-processing of Fujisaki parameters did not have access to PROLAB annotation. For each subject in the database a constant Fb was chosen based on the mean value established in the first automatic parameter estimation step. Alpha and beta were set to 2 and 20 Hz, respectively.

2.3 Mapping of accent commands to prosodic labels

Each prosodic label corresponding to an accented vowel was assigned a Fujisaki accent command if the distance between the vowel onset and the accent command center were smallest of all competing accent commands within a maximum distance

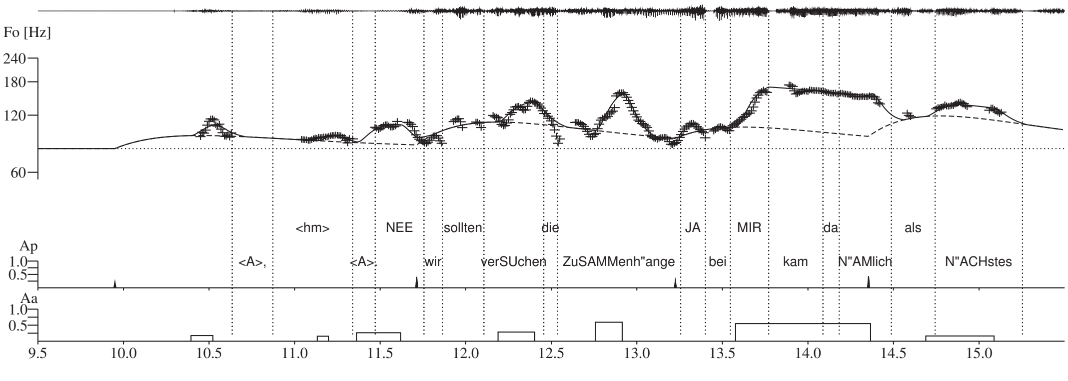

Fig. 3: The same utterance as in Fig. 2 after manual post-processing with the FujiParaEditor.

of ±300 ms. We informally tested several modifications of these assignment criteria including the variation of the maximum distance between 200 ms and 500 ms. It turned out that a larger maximum distance mainly increased the number of assigned labels but also the standard deviations of accent command onsets and offsets. However, it had little effect on their means.

3 Results

In total the subcorpus comprised 456 labels for early, medial, and late peaks of which 87.3% were subjected to further analysis (see Table 1). No early reinforced accent ‘3)’ was found. While all 31 reinforced accents co-occur with Fujisaki accent commands, 88.3% of the default accents and 79.6% of the partially deaccented accents are accompanied by accent commands.

Table 1: PROLAB’s labels assigned to accented syllables, numbers of occurrence and numbers as well as percentage of aligned accent commands.

accent	count	aligned accent commands	%
3^	17	17	100%
3(	14	14	100%
2)	82	65	79.3%
2^	160	140	87.5%
2(	90	88	97.8%
1)	24	19	79.2%
1^	46	34	73.9%
1(	23	21	91.3%
$\sum$	456	398	87.3%

3.1 Accent command amplitude and accent label

Fig. 4 shows the mean accent command amplitudes corresponding to the 8 accent labels. The reinforced accent 3 has a mean accent command amplitude of 0.62 while the default accent has 0.35 and the partially deaccented accent has 0.25. These values represent f0 intervals and can be transformed into the more convenient semitone values 9.7, 5.5, and 3.9, respectively.

An ANOVA with the independent variables *accent level* and *peak synchronisation* (Table 2) shows that the effect of accent level on accent command amplitude is highly significant while the peak synchronisation has no significant influence.

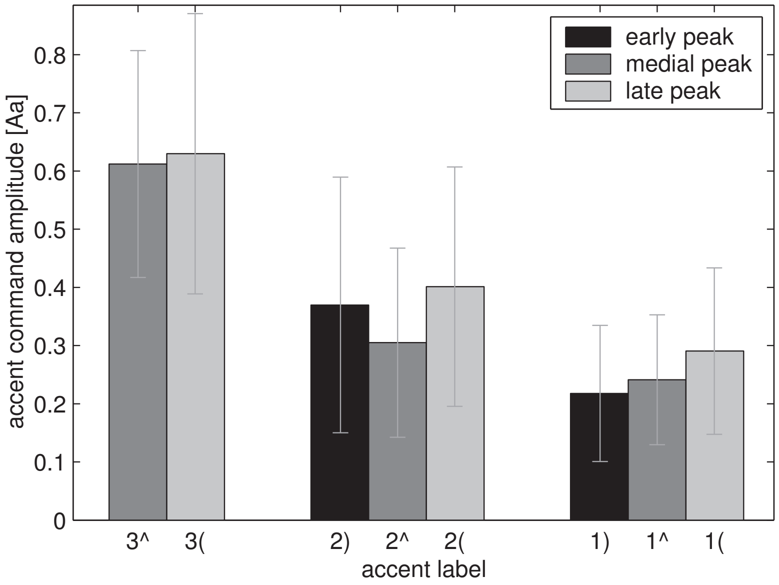

Fig. 4: Mean accent command amplitudes and standard deviations of the 8 accent labels.

However, posthoc multiple comparison (Scheffé as well as Bonferroni) reveals that late peaks are associated with significantly higher accent command amplitudes than early or medial peaks (p=0.019).

3.2 Accent command duration and vowel duration

Fig. 5 shows that durations of vowels in accented syllables and of their corresponding accent commands are completely uncorrelated. Even if the data points of accent commands longer than 300 ms or 400 ms would be removed from the estimation of Pearson's r the correlations would remain very weak. This is in accordance with the results of Mixdorff who found that syllable durations and accent command durations are uncorrelated.

Table 2: Analysis of variance of the effect of accent level and peak synchronisation class on accent command amplitude (Aa) of the Fujisaki model.

	df	F	p	significance
accent level	2	42.272	0.00	***
early/medial/late	2	1.878	0.15	n.s.
interaction	3	1.120	0.34	n.s.

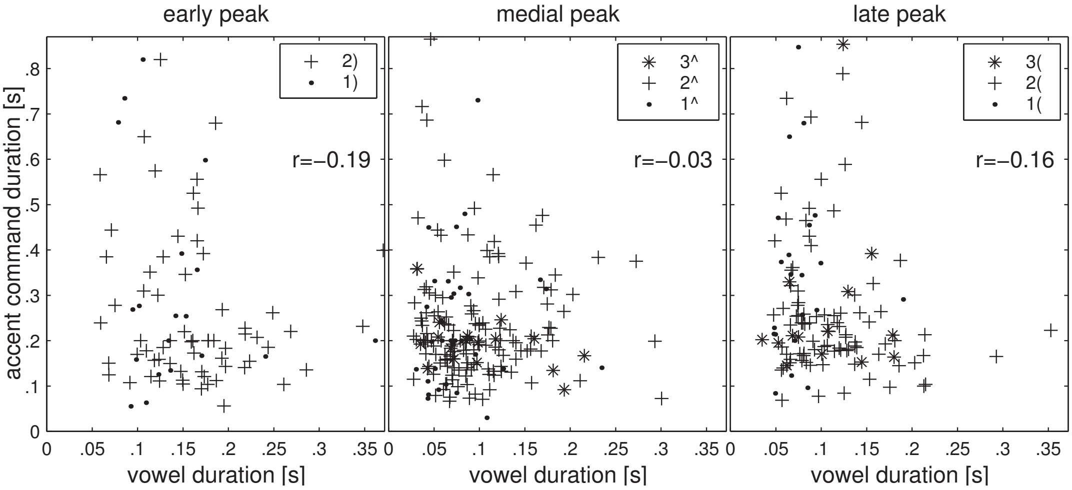

Fig. 5: Scatter plots between vowel durations of accented syllables and their corresponding accent commands.

3.3 Accent command timing and accent label

It was assessed by multivariate GLM whether the independent variables *accent level* and *peak synchronisation* showed correlations with the dependent variables *accent command onset time*, *offset time*, and *center time*, that is, the mean of onset and offset time. GLM were applied since the dependent variables were to some degree correlated. Table 3 shows that both independent variables significantly correlate with the timing of accent commands. Posthoc multiple comparison showed (i) that accent level has no significant effect on accent onset time (p=0.503) but a very significant effect on accent offset time (p=0.003), namely, a higher accent level leads to a later offset time, and (ii) that peak synchronisation highly significantly influences all accent command times (p=0.00): early peaks, on average, lead to 194 ms earlier onset times and 156 ms earlier offset times than medial peaks, and late peaks, on average, lead to 99 ms later onset times and 144 ms later offset times.

Table 3: Multivariate GLM analysis of the effect of accent level and peak synchronisation class on accent command onset time (T1), center, and offset time (T2) of the Fujisaki model.

	F	p	significance
accent level	2.365	0.03	*
early/medial/late	38.112	0.00	***
interaction	1.127	0.34	n.s.

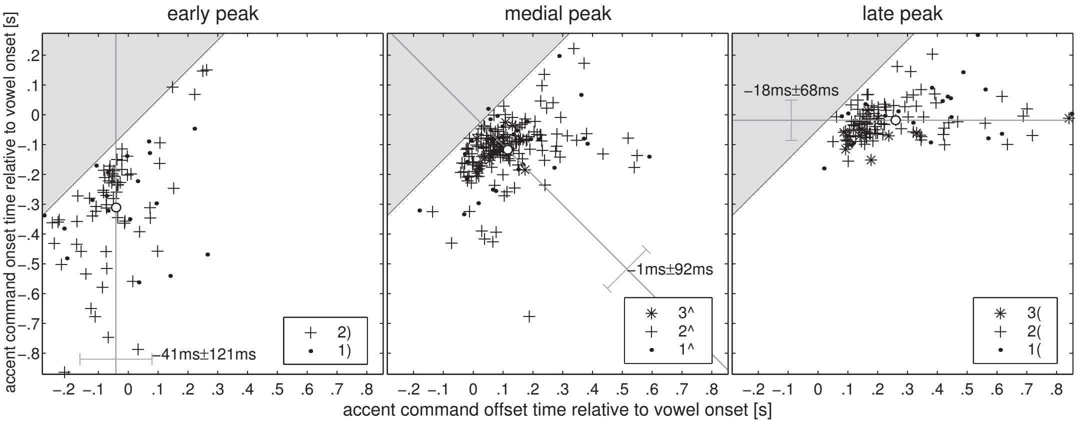

Fig. 6: Scatter plots between accent command offset and onset both relative to the vowel onset of the corresponding accented syllable. Standard deviations are *grey*. Black-bordered white dots show the means of the three distributions. The triangular area contains accent commands with a duration of 50 ms or less.

Fig. 6 shows scatter plots between the offset and onset of all accent commands both relative to the onset of the corresponding accented vowel onset. It is evident, that the accent command **offset** of early peaks is approximately 41 ms earlier than the vowel onset, while the accent command **onset** of late peaks is 18 ms earlier.

The observed phonetic realisations labelled as medial peaks include many so-called hat-patterns created by a sequence of two accented syllables, the first with a rising and the second with a falling f0 pattern. In these cases, the underlying accent command is actually temporally aligned with two syllables, making the alignment decision ambiguous. Nonetheless, it is quite obvious for most of the medial peaks that the corresponding accent command **centers** are aligned with the accented vowel onset (1 ms earlier).

Fig. 7 shows the distributions of accent command onset, center, and offset times relative to the vowel onset. The accent command onset times of early peaks show a large variance while the variance of the offset times is considerably smaller. The opposite applies to late peaks. The alignment of the accent command center times leads to the smallest variance for medial peaks.

Fig. 8 shows the mean alignments between the accented vowel segments annotated by prosodic labels of the PROLAB scheme and mean accent commands of the Fujisaki model. Especially the accent command timing of the medial peaks, displayed in the center panels, suggests an additional, but weaker anchoring of the accent command offset time to the vowel offset.

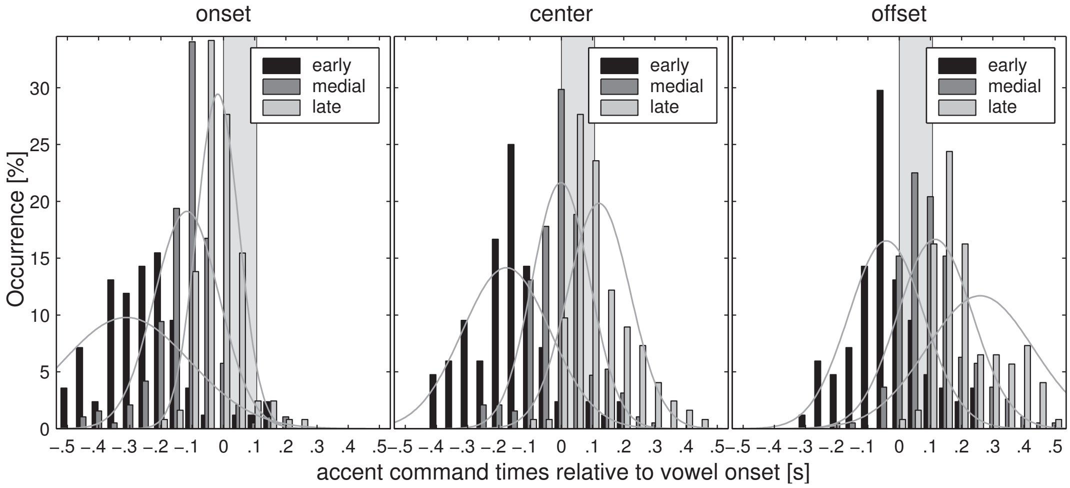

Fig. 7: Histograms and fitted Gaussian distributions of the distance between the accented vowel onset and onset, center, and offset times of the corresponding Fujisaki accent commands. The mean vowel is represented by a light grey rectangle.

4 Discussion

A good agreement between labeller-rated accent levels and accent command amplitudes, as well as peak alignment (early, medial, late) with accent command timing was found. Our results suggest that the accented vowel onset is synchronized with accent command offsets in case of early peaks, with center points of the accent commands in case of medial peaks, and with accent command onsets in case of late peaks (Fig. 8). However, there is a considerable overlap between the phonetic realisations of medial and late peaks (Fig. 7) which seems to be in line with the results of the categorical pitch perception experiments (Kohler 1987; Kohler 1991b) in which the distinction between the medial and late peak is less marked than between the early and medial peak.

Niebuhr and Ambrazaitis 2006 presented alignment values for the starting point of an f0 rise with the onset of the accented vowel in medial peaks and late peaks. According to them, f0 starts to rise 29 ms before the vowel onset in medial peaks and 54 ms in late peaks. The deviations from our results might be due to the fact that the applied criteria of anchor point extraction in the f0 contours are very different, and to the fact that Niebuhr and Ambrazaitis hand-picked only 31 accents in total from the whole corpus while we considered almost 400 accents from four out of the twelve speakers. The latter possibly also is the reason for our standard deviations to be considerably larger.

The conclusion of Niebuhr 2006 that the anchor point in the speech signal is the instant of increase in intensity instead of the accented vowel onset is not con-

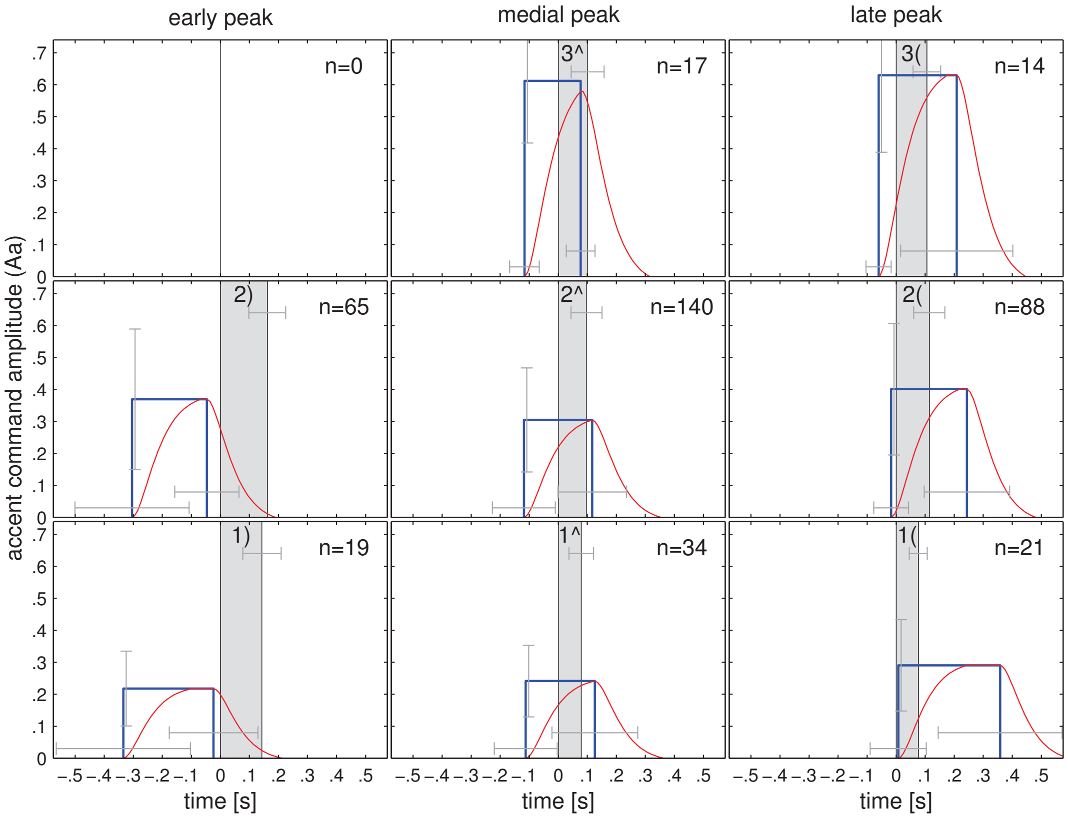

Fig. 8: The panels show mean accent commands of the Fujisaki model corresponding to the phonological classes of KIM: early, medial, and late peak (*left, center, and right panels, respectively*) each with reinforced, default, and partially deaccented accent level (*top, middle, and bottom row, respectively*). Corresponding vowel segments (*light grey*) are annotated by symbolic labels of the prosodic annotation scheme PROLAB. Their onsets are set to the zero point of the x-axis. All standard deviations are indicated by grey whiskers.

tradictory to our findings since in most cases both of these anchor points coincide. However, the question if another anchor point is possibly even better suited to reduce the observed standard deviations remains open and will be tackled in a future investigation.

5 Conclusion

Our findings concerning the relationship between the symbol layer and the signal layer of German intonational peaks in spontaneous speech can be summarized as follows:

1. Early peaks are characterized by an accent command offset time of 41 ms earlier than the accented vowel onset.
2. Medial peaks show an exact alignment of the accent command center point and the accented vowel onset and possibly an alignment of the accent command offset time with the vowel offset.
3. Late peaks are characterized by an accent command onset time of 18 ms earlier than the accented vowel onset.
4. Accent command amplitudes are significantly different for the three accent levels: Reinforced accent is realized with an f0 interval of 9.7 semitones, default accent with 5.5 semitones, and partial deaccentuation with 3.9 semitones.

Potential future applications of this work are automatic prosodic labelling within the PROLAB scheme, and f0 contour generation from PROLAB labels by means of the Fujisaki model.

Acknowledgements

Many thanks to Michel Scheffers for his invaluable discussions and support with the multiple extraction and unification of f0, and to Julia Arp, Benjamin Hamer, Anja Hübner, Thomas Jacobsen, and Nina Redlingshöfer for manual post-processing of f0 and of the accent and phrase commands.

REFERENCES

Abercrombie, D. (1967). *Elements of General Phonetics*. Edinburgh: Edinburgh University Press.

Aho, E. (2008a). Variation i dialektprosodi i de svenska dialekterna i Finland. In J. Lindström, P. Kukkonen, C. Lindholm, and Å. Mickwitz (Eds.), *Svenskan i Finland 10*, pp. 35–42. (Nordica Helsingiensia 11) Helsingfors Universitet, Institutionen för nordiska språk och nordisk litteratur.

Aho, E. (2008b). Om prosodin i Liljendaldialekten. In H. Palmén, C. Sandström, and J.-O. Östman (Eds.), *Dialekt i östra Nyland. Fältarbete i Liljendal med omnejd*, pp. 37–52. (Dialektforskning 1). Helsingfors Universitet, Institutionen för nordiska språk och nordisk litteratur.

Aho, E. (forthcoming). Sulvan prosodiasta. In J.-O. Östman (Ed.), *Dialektforskning: Solf.*

Ambrazaitis, G. (2007a). Expressing 'confirmation' in Swedish: the interplay of word and utterance prosody. In J. Trouvain and W. J. Barry (Eds.), *Proceedings of the 16th International Congress of Phonetic Sciences, 6–10 August, Saarbrücken, Germany*, pp. 1093–1096.

Ambrazaitis, G. (2007b). Swedish word accents in a 'confirmation' context. In *Proceedings of FONETIK 2007*, pp. 49–52. (TMH-QPSR 50) Stockholm: Royal Institute of Technology – Speech, Music and Hearing.

Anderson, A. H., M. Bader, E. G. Bard, E. Boyle, G. Doherty, S. Garrod, et al. (1991). The HCRC Map Task Corpus. *Language and Speech 34*, 351–366.

Andreeva, B., W. J. Barry, and I. Steiner (2007). Producing phrasal prominence in German. In J. Trouvain and W. J. Barry (Eds.), *Proceedings of the 16th International Congress of Phonetic Sciences, 6–10 August, Saarbrücken, Germany*, pp. 1209–1212.

Anumanchipalli, G., K. Prahallad, and A. Black (2008). Significance of early tagged contextual graphemes in grapheme based speech synthesis and recognition systems. In *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2008), March 30 – April 4, Las Vegas, NV*, pp. 4645–4648.

Asakawa, S. et al. (2007). Beigo boin hatsuon ni taisuru zettaiteki hyouka to soutaiteki hyouka kan no soukan bunseki (An analysis of the correlate between the absolute evaluation and relative evaluation on American English vowels). [In Japanese] In *Proceedings of the Acoustic Society of Japan Spring Meeting*, pp. 373–374.

Asu, E. L. and F. Nolan (1999). The effect of intonation on pitch cues to the Estonian quantity contrast. In J. J. Ohala, Y. Hasegawa, M. Ohala, D. Granville, and A. C. Bailey (Eds.), *Proceedings of the 14th International Congress of Phonetic Sciences, San Francisco, 1–7 August,* pp. 1873–1876.

Asu, E. L. and F. Nolan (2007). The analysis of low accentuation in Estonian. *Language and Speech 50*, 567–588.

Atterer, M. and D. R. Ladd (2004). On the phonetics and phonology of “segmental anchoring” of F0: evidence from German. *Journal of Phonetics 32*, 177–197.

Barbosa, P. A. (1996). At least two macrorhythmic units are necessary for modeling Brazilian Portuguese duration: emphasis on segmental duration generation. *Cadernos de Estudos Lingüísticos 31*, pp. 33–53.

Barbosa, P. A. (2001). Generating duration from a cognitively plausible model of rhythm production. In *Proceedings of the Seventh European Conference on Speech Communication and Technology (Eurospeech 2001), September 3–7, Aalborg, Denmark,* vol. 2, pp. 967–970.

Barry, W. J., B. Andreeva, M. Russo, S. Dimitrova, and T. Kostadinova (2003). Do rhythm measures tell us anything about language type? In M. J. Solé, D. Recasens, and J. Romero (Eds.), *Proceedings of the 15th International Congress of Phonetic Sciences, 3–9 August, Barcelona*, pp. 2693–2696.

Barry, W. J., B. Andreeva, and I. Steiner (2007). The phonetic exponency of phrasal accentuation in French and German. In *Proceedings of Interspeech 2007, August 27–31, Antwerp, Belgium*, pp. 1010–1013.

Black, A. (2006). CLUSTERGEN: A Statistical Parametric Synthesizer using Trajectory Modeling. In *Proceedings of Interspeech 2006 – ICSLP*, *September 17–21, Pittsburgh, PA*, pp. 1762–1765.

Black, A. W., H. Zen, and K. Tokuda (2007). Statistical parametric speech synthesis. In *Proceedings of the 2007 International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2007), April 15–20, Honolulu, Hawaii*, pp. 1229–1232.

Boersma, P. (1993). Accurate short-term analysis of the fundamental frequency and the harmonics-to-noise ratio of a sampled sound. In *IFA Proceedings 17*, pp. 97–110. Amsterdam: Institute of Phonetic Sciences.

Boersma, P. and D. Weenink (2007–2008). Praat: doing phonetics by computer (Versions 4.6–5.0) [Computer program]. Retrieved 2007–2008 from http://www.praat.org/.

Bolinger, D. L. (1965). Pitch accent and sentence rhythm. In I. Abe and T. Kanekiyo (Eds.), *Forms of English: Accent, Morpheme, Order*, pp. 138–180. Cambridge, MA: Harvard University Press.

Bolinger, D. (1972). Accent is predictable (if you’re a mind-reader). *Language 48*, 633–644.

Bondarko, L. V. (1998). *Fonetika sovremennogo russkogo yazyka.* Sankt-Peterburg: Izd. S.-Peterburgskogo universiteta.

Borden, G. J., K. S. Harris, and L. J. Raphael (1994). *Speech Science Primer. 3rd Ed.* Maryland: Williams & Wilkins.

Brady, P. T. (1968). A statistical analysis of on-off patterns in 16 conversations. *The Bell System Technical Journal 47*, 73–91.

Bruce, G. (1974). Tonaccentregler för sammansatta ord i några sydsvenska stadsmål. In C. Platzack (Ed.), *Svenskans beskrivning 8*, 62–75.

Bruce, G. (1977). *Swedish Word Accents in Sentence Perspective*. (Travaux de l'Institut de Linguistique de Lund XII) Lund: Gleerup.

Bruce, G. (1998). *Allmän och svensk prosodi*. (Praktisk Lingvistik 16) Lund University, Institutionen för Lingvistik.

Bruce, G. (2001). Secondary stress and pitch accent synchronization in Swedish. In W. A. van Dommelen and T. Fretheim (Eds.), *Nordic Prosody: Proceedings of the VIIIth Conference, Trondheim 2000,* pp. 33–44. Frankfurt am Main: Peter Lang.

Bruce, G. (2004). An intonational typology of Swedish. In B. Bel and I. Marlien (Eds.), *Proceedings of Speech Prosody 2004, March 23–26, Nara, Japan,* pp. 175–178.

Bruce, G. (2005). Intonational prominence in varieties of Swedish revisited. In S.-A. Jun (Ed.), *Prosodic Typology. The Phonology of Intonation and Phrasing*, pp. 410–429. Oxford: Oxford University Press.

Bruce, G. (2007). Components of a prosodic typology of Swedish intonation. In T. Riad and C. Gussenhoven (Eds.), *Tones and Tunes*, *Vol. 1, Typological Studies in Word and Sentence Prosody*, pp 113–146. Berlin: Mouton de Gruyter.

Bruce, G. and E. Gårding (1978). A prosodic typology for Swedish dialects. In E. Gårding, G. Bruce, and R. Bannert (Eds.), *Nordic Prosody: Papers from a Symposium*, pp. 219–228. Lund: Lund University, Department of Linguistics.

Bruce, G. and B. Granström (1993). Prosodic modelling in Swedish speech synthesis. A prosodic typology for Swedish dialects. *Speech Communication 13*, 63–73.

Bruce, G., B. Granström, and S. Schötz (2007). Simulating Intonational Varieties of Swedish. In J. Trouvain and W. J. Barry (Eds.), *Proceedings of the 16th International Congress of Phonetic Sciences, 6–10 August, Saarbrücken, Germany*, pp. 1237–1240.

Buder, E. H. and J. L. Edrington (2008). Conversational prosodic interactivity when one partner has aphasia. In P. A. Barbosa, S. Madureira, and C. Reis (Eds.), *Speech Prosody 2008: Fourth Conference on Speech Prosody, Campinas, Brazil*, pp. 501–504.

Caspers, J. (2003). Local speech melody as a limiting factor in the turn-taking system in Dutch. *Journal of Phonetics 31*, 251–276.

Catford, J. C. (2001). *A Practical Introduction to Phonetics*. Cambridge: Cambridge University Press.

Cathcart, N., J. Carletta, and E. Klein (2003). A shallow model of backchannel continuers in spoken dialogue. In *Proceedings of the Tenth Conference of the*

European Chapter of the Association for Computational Linguistics, pp. 51–58. Budapest, Hungary.

Chiba, T. and M. Kajiyama (1941). *The Vowel: Its Nature and Structure*, Tokyo: Kaiseikan.

Clark, J. and C. Yallop (1990). *An Introduction to Phonetics and Phonology*, Massachusetts: Basil Blackwell.

Couper-Kuhlen, E. (1986). *An Introduction to English Prosody.* London: Edward Arnold.

Cruttenden, A. (1986). *Intonation.* Cambridge: Cambridge University Press.

Cruttenden, A. (1997). *Intonation.* Cambridge: Cambridge University Press.

Crystal, D. (1975). *The English Tone of Voice.* London: Edward Arnold.

Dabbs, J. M., Jr. and R. B. Ruback (1984). Vocal patterns in male and female groups. *Personality and Social Psychology Bulletin 10*, 518–525.

Dang, J. and K. Honda (2004). Construction and control of a physiological articulatory model. *The Journal of the Acoustical Society of America 115*, 853–870.

Dauer, R. (1983). Stress-timing and syllable-timing reanalyzed. *Journal of Phonetics 11*, 51–62.

Dauer, R. M. (1987). Phonetic and phonological components of language rhythm. *Proceedings of the 11th International Congress of Phonetic Sciences, August 1–7, Tallinn, Estonia, vol. 5,* pp. 447–450.

de Boer, B. and P. K. Kuhl (2003). Investigating the role of infant-directed speech with a computer model. *Acoustics Research Letters Online 4*, 129–134.

de Cheveigné, A. and H. Kawahara (2002). YIN, a fundamental frequency estimator for speech and music. *The Journal of the Acoustical Society of America 111*, 1917–1930.

De Graaf, T. and F. J. Koopmans-van Beinum (1982/3). Vowel contrast reduction in Japanese compared to Dutch. In *Proceedings of the Institute of Phonetic Sciences,* Amsterdam, pp. 27–38.

de Lacy, P. (1999). Tone and Prominence. Ms, University of Massachusetts, Amherst. ROA# 333 [http://roa.rutgers.edu].

del Giudice, A., R. Shosted, K. Davidson, M. Salihie, and A. Arvaniti (2007). Comparing methods for locating pitch ”elbows”. In J. Trouvain and W. J. Barry (Eds.), *Proceedings of the 16th International Congress of Phonetic Sciences, 6–10 August, Saarbrücken, Germany*, pp. 1117–1120.

Deutsch, D., T. Henthorn, E. Marvin, and H. Xu (2006). Absolute pitch among American and Chinese conservatory students: prevalence differences, and evidence for a speech-related critical period. *The Journal of the Acoustical Society of America 119*, 719–722.

Di Cristo, A. (1978/1985). *De la Microprosodie à l'Intonosyntaxe*. Thèse de Doctorat d'Etat, Université de Provence, Aix, 1978. Publications de l'Université de Provence, 2 vol., 1985.

Di Cristo, A. (1998). Intonation in French. In D. Hirst and A. Di Cristo (Eds.), *Intonation Systems. A Survey of Twenty Languages*, pp. 195–218. Cambridge: Cambridge University Press.

Duncan, S., Jr. (1972). Some signals and rules for taking speaking turns in conversations. *Journal of Personality and Social Psychology 23*, 283–292.

Edlund, J. and M. Heldner (2005). Exploring prosody in interaction control. *Phonetica 62*, 215–226.

Eek, A. (1980a). Estonian quantity: notes on the perception of duration. In A. Eek (Ed.), *Estonian Papers in Phonetics 1979*, pp. 5–29. Tallinn: Institute of Language and Literature.

Eek, A. (1980b). Further information on the perception of Estonian quantity. *Estonian Papers in Phonetics 1979*, 31–57.

Eek, A. and E. Meister (2003). Foneetilisi katseid ja arutlusi kvantiteedi alalt (I). Häälikukestusi muutvad kontekstid ja välde. *Keel ja Kirjandus 11*, 815–837 and *12*, 904–918.

Eek, A. and E. Meister (2004). Foneetilisi katseid ja arutlusi kvantiteedi alalt (II). Takt, silp ja välde. *Keel ja Kirjandus 4*, 251–271, and *5*, 336–357.

Eklund, R. and A. Lindström (2001). Xenophones: an investigation of phone set expansion in Swedish and implications for speech recognition and speech synthesis. *Speech Communication 35*, 81–102.

Elenius, K. (1999). Two Swedish SpeechDat databases – some experiences and results. *Proceedings of Eurospeech '99, Budapest*, pp. 2243–2246.

Elert, C.-C. (1964). *Phonologic Studies of Quantity in Swedish Based on Material from Stockholm Speakers*. Uppsala: Almqvist & Wiksell.

Elert, C.-C. (1994). Indelning och gränser inom området för den nu talade svenskan – en aktuell dialektografi. In L. E. Edlund (Ed.), *Kulturgränser – myt eller verklighet*, pp. 215–228. Umeå: Diabas.

Ellis, R. (1994). *The Study of Second Language Acquisition.* Oxford: Oxford University Press.

Engstrand, O. (1995). Phonetic interpretation of the word accent contrast in Swedish. *Phonetica 52*, 171–179.

Engstrand, O. (1997). Phonetic interpretation of the word accent contrast in Swedish: evidence from spontaneous speech. *Phonetica 54*, 61–75.

Engstrand, O., R. Bannert, G. Bruce, C-C. Elert, and A. Eriksson (1997). Phonetics and phonology of Swedish dialects around the year 2000: a research plan. *Papers from Fonetik 97,* pp 97–100. (PHONUM 4) Umeå: Umeå University, Department of Phonetics.

Eriksson, A. (1991). *Aspects of Swedish Speech Rhythm.* Göteborg: University of Göteborg.

Fanselow, G. and D. Ćavar (2002). Distributed Deletion. In A. Alexiadou (Ed.), *Theoretical Approaches to Universals*, pp. 65–107. Amsterdam: Benjamins.

Fanselow, G. and C. Féry (in prep.). Discontinuous Noun Phrases. To appear at Cambridge University Press.

Fant, G., A. Kruckenberg, and J. Liljencrants (2000). Acoustic-phonetic analysis of prominence in Swedish. In A. Botinis (Ed.), *Intonation: Analysis, Modelling and Technology*, pp. 55–86. Dordrecht: Kluwer Academic Publishers.

Fernald, A. and T. Simon (1984). Expanded intonation contours in mothers' speech to newborns. *Developmental Psychology 20*, 104–113.

Féry, C., A. Paslawska, and G. Fanselow (2007). Nominal Split Constructions in Ukrainian. *Journal of Slavic Linguistics 15*, 3–48.

Fintoft, K. (1987). Toneme patterns in Norwegian and in Swedish dialects. In R. Channon and L. Shockey (Eds.), *In Honor of Ilse Lehiste: Ilse Lehiste pühendusteos*, pp. 33–50. Dordrecht: Foris Publications.

Fisher, C. and H. Tokura (1996). Acoustic cues to grammatical structure in infant-directed speech: cross-linguistic evidence. *Child Development 67*, 3192–3218.

Flege, J. E. (1995). Second language speech learning: theory, findings, and problems. In W. Strange (Ed.), *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*, pp. 233–277. Timonium, MD: York Press.

Fónagy, I. and K. Magdics (1960). Speed of utterance in phrases of different lengths. *Language and Speech 3*, 179–192.

Fox, R. A. and I. Lehiste (1987). Discrimination of duration ratios by native English and Estonian listeners. *Journal of Phonetics 15*, 349–363.

Fox, R. A. and I. Lehiste (1989). Discrimination of duration ratios in bisyllabic tokens by native English and Estonian listeners. *Journal of Phonetics 17*, 167–174.

Francis, A. L., V. Ciocca, and B. K. Chit Ng (2003). On the (non)categorical perception of lexical tones. *Perception and Psychophysics 65*, 1029–1044.

Francis, A. L., V. Ciocca, L. Ma, and K. Fenn (2008). Perceptual learning of Cantonese lexical tones by tone and non-tone language speakers. *Journal of Phonetics 36*, 268–294.

Fretheim, T. (1981). Intonational Phrasing in Norwegian. *Nordic Journal of Linguistics 4*, 111–137.

Fretheim, T. (1987). Phonetically Low Tone–Phonologically High Tone, and Vice Versa. *Nordic Journal of Linguistics 10*, 35–58.

Fretheim, T. (1992). Themehood, rhemehood and Norwegian focus structure. *Folia Linguistica 26*, 111–160.

Fretheim, T. and R. A. Nilsen (1989). Romsdal Intonation: Where East and West Norwegian Pitch Contours Meet. In J. Niemi (Ed.), *Papers from the Eleventh Scandinavian Conference of Linguistics*, pp. 442–458. Joensuu: Joensuun yliopisto.

Fretheim, T. and R. A. Nilsen (1991). In defence of [±foc]. In Y. No and M. Libucha (Eds.), *ESCOL '90. Proceedings of the Seventh Eastern States Conference on Linguistics*, pp. 102–111. Columbus, OH: The Ohio State University.

Fujisaki, H. (2004). Information, prosody, and modeling with emphasis on tonal features of speech. In *Proceedings of the 2nd International Conference on Speech Prosody, March 23–26, Nara, Japan*, pp. 1–10.

Fujisaki, H. and K. Hirose (1984). Analysis of voice fundamental frequency contours for declarative sentences of Japanese. *Journal of the Acoustical Society of Japan (E) 5*(4), 233–241.

Gårding, E. (1977). *The Scandinavian Word Accents*. (Travaux de l'Institut de Linguistique de Lund 11) Lund: Gleerup.

Gårding, E. (1982). Swedish Prosody. *Phonetica 39*, 288–301.

Gårding, E. (1983). Superposition as an invariant feature of intonation. Paper read at a symposium on "Invariance and Variability", MIT, 1983.

Gårding, E. (1989). Intonation in Swedish. In *Working Papers 35*, pp. 63–88. Lund: Lund University, Department of Linguistics.

Gibbon, D. (1998). Intonation in German. In D. Hirst and A. Di Cristo (Eds.), *Intonation Systems. A Survey of Twenty Languages*, pp. 78–95. Cambridge: Cambridge University Press.

Goldsmith, J. A. (1979). *Autosegmental Phonology*. PhD dissertation, MIT. Distributed by IUCL. New York: Garland Press.

Goldsmith, J. A. (1990). *Autosegmental & Metrical Phonology*. Oxford, Cambridge: Basil Blackwell.

Goldstein, L., I. Chitoran, and E. Selkirk (2007). Syllable structure as coupled oscillator modes: Evidence from Georgian vs. Tashlhiyt Berber. In J. Trouvain and W. J. Barry (Eds.), *Proceedings of the 16th International Congress of Phonetic Sciences, 6–10 August, Saarbrücken, Germany*, pp. 241–244.

Golston, C. and R. Wiese (1996). Zero morphology and constraint interaction: subtraction and epenthesis in German dialects. In G. Booij and J. van Marle (Eds.), *Yearbook of Morphology 1995*, pp. 143–159. Berlin: Springer.

Grice, M. and S. Baumann (2002). Deutsche Intonation und GToBI. *Linguistische Berichte 191*, 267–298.

Grice M., S. Baumann, and R. Benzmüller (2005). German intonation in autosegmental-metrical phonology. In S.-A. Jun (Ed.), *Prosodic Typology. The Phonology of Intonation and Phrasing*, pp. 55–83. Oxford: Oxford University Press.

Grieser D. L. and P. K. Kuhl (1988). Maternal speech to infants in a tonal language: Support for universal prosodic features in motherese. *Developmental Psychology 24*, 14–20.

Gussenhoven, C. (2004). *The Phonology of Tone and Intonation.* Cambridge: Cambridge University Press.

Hakulinen, L. (1961). *The Structure and Development of the Finnish Language.* Indiana University. (Transl. John Atkinson).

Halliday, M. A. K. (1967). *Intonation and Grammar in British English*. The Hague: Mouton.

Hansson, P. (2003). *Prosodic Phrasing in Spontaneous Swedish*. (Travaux de l'Institut de Linguistique de Lund 43) Lund.

Hardy, H., K. Baker, H. Bonneau-Maynard, L. Devillers, S. Rosset, and T. Strzalkowski (2003). Semantic and dialogic annotation for automated multilingual customer service. In *Proceedings of EUROSPEECH 2003, September 1–4, Geneva, Switzerland*, pp. 201–204.

Haspelmath, M. (2006). Against markedness (and what to replace it with). *Journal of Linguistics 42*, 25–70.

Haugen, E. (1967). On the rules of Norwegian tonality. *Language 43*, 185–202.

Hedelin, P., A. Jonsson, and P. Lindblad (1987). Svenskt uttalslexikon. 3 ed. *Tech Report 4*, Chalmers University of Technology.

Heldner, M. (2001). *Focal accent – f0 movement and beyond.* (PHONUM 8) Umeå University.

Helgason, P. (2006). SMTC – A Swedish Map Task Corpus. In *Proceedings from Fonetik 2006,* pp. 57–60. (Working Papers 52) Lund: Lund University, Centre for Languages & Literature, Department of Linguistics & Phonetics.

Hewings, M. (1995). The English intonation of native speakers and Indonesian learners: a comparative study. *RELC Journal 26*, 27–46.

Hockett, C. (1958). *A Course in Modern Linguistics*. New York: MacMillan.

Horne, M. (1994). Generating prosodic structure for synthesis of Swedish intonation. *Working Papers, Department of Linguistics, Lund University* 43, pp. 72–75.

House, D. (1990). *Tonal Perception in Speech.*(Travaux de l'Institut de linguistique de Lund 24) Lund: Lund University Press.

Iivonen, A. (1984). On explaining the initial fundamental frequency in Finnish utterances. In C.-C. Elert, I. Johansson, and E. Strangert (Eds.), *Nordic Prosody III,* pp. 107–119. (Acta Universitatis Umensis, Umeå Studies in the Humanities 59) Stockholm: Almqvist & Wiksell.

Iivonen, A. (1994). Implications and applications of a psychoacoustical F1/F2 vowel chart. In R. Aulanko and A.-M. Korpijaakko-Huuhka (Eds.), *Proceedings of the Third Congress of the International Clinical Phonetics and Linguistics Association*, *9–11 August, Helsinki,* pp. 83–91. (Publications of the Department of Phonetics 39) Helsinki: University of Helsinki.

Iivonen, A. (1998). Intonation in Finnish. In D. Hirst and A. Di Cristo (Eds.), *Intonation Systems: A Survey of Twenty Languages*, pp. 311–327. Cambridge: Cambridge University Press.

Iivonen, A., T. Seppänen, K. Noponen, and J. Toivanen (2004). Puhujan temporaalisen äänialan visualisointisovellus. [Visualisation application of speaker's temporal voice range] *Puhe ja kieli 24*, 31–39.

Imai, S. (1983). Cepstral analysis/synthesis on the Mel frequency scale. In *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '83), Boston, MA*, pp. 93–96.

Jacobson J. L., D. C. Boersma, R. B. Fields, and K. L. Olson (1983). Paralinguistic features of adult speech to infants and small children. *Child Development 54*, 436–442.

Jaffe, J. and S. Feldstein (1970). *Rhythms of dialogue*. New York, NY: Academic Press.

Jurafsky, D., E. Shriberg, and D. Biasca (1997). Switchboard SWBD-DAMSL Annotation Coders Manual, Draft 13. Retrieved 2008-12-06, from http://www.stanford.edu/~jurafsky/ws97/manual.august1.html

Jurafsky, D., E. Shriberg, B. A. Fox, and T. Curl (1998). Lexical, prosodic, and syntactic cues for dialog acts. In *Proceedings of ACL/COLING-98 Workshop on Discourse Relations and Discourse Markers,* pp. 114–120. Montreal, Canada.

Kabak, B. and F. Plank (2007). The role of phonology in word class distinction. Handout, Association for Linguistic Typology – 7th Biennial Meeting (ALT 7), 25–28 September, Paris.

Kärkkäinen, H., V. de Silva, and R. Ullakonoja (2006a). Havaintoja perustaajuudesta spontaanin puheen jaksottelussa suomessa ja venäjässä. In R. Aulanko, L. Wahlberg, and M. Vainio (Eds.), *Fonetiikan päivät 2006 – The Phonetics Symposium 2006*, pp. 84–92. (Publications of the Department of Speech Sciences 53) Helsinki: University of Helsinki.

Kärkkäinen, H., V. de Silva, R. Ullakonoja, and H. Anttila (2006b). Delenie na intonatsionnye edinitsy i realizatsiya aktsentnoj vydelennosti v finskoj spontannoj rechi. In L.V. Bondarko (Ed.). *Materialy XXXV Mezhdunarodnoj filologicheskoj konferentsii. Sektsiya fonetiki 20,* pp. 3–15. Sankt-Peterburg: Filologicheskij fakul'tet Sankt-Peterburgskogo gosudarstvennogo universiteta.

Kazenin, K. (submitted). Splits and Morphological Complexity: The Case of Caucasian Languages. Submitted to *Linguistic Typology*.

Kenstowicz, M. (1994). *Phonology in Generative Grammar.* Cambridge MA, Oxford UK: Blackwell.

Kettunen, L. (1925). Untersuchung über die livische Sprache I. Phonetische Einführung. Sprachproben. *Acta et Commentationes Universitatis Dorpatensis B VIII*. Tartu.

Kettunen, L. (1937). Vielä liivin intonaatioista. *Virittäjä 41*, 313–317, 380.

Kim, H. and J. Cole (2005). The stress foot as a unit of planned timing: Evidence from shortening in the prosodic phrase. In *Proceedings of the 9th European*

Conference on Speech Communication and Technology (Interspeech 2005 – Eurospeech), Lisboa, Portugal, pp. 2365–2368.

Kim, Y. (2006). Variation and Finnish influence in Finland Swedish dialect intonation. In *Proceedings from Fonetik 2006,* pp. 77–80. (Working Papers 52) Lund: Lund University, Centre for Languages & Literature, Department of Linguistics & Phonetics.

Kiparsky, P. (2006). Livonian stød. Manuscript. Stanford University. (http://www.stanford.edu/~kiparsky/Papers/livonian.pdf)

Kirchhoff, K. and J. A. Bilmes (1999). Statistical acoustic indications of coarticulation. In J. J. Ohala, Y. Hasegawa, M. Ohala, D. Granville, and A. C. Bailey (Eds.), *Proceedings of the 14th International Congress of Phonetic Sciences, San Francisco, 1–7 August,* pp. 1729–1732.

Kirchhoff, K. and S. Schimmel (2005). Statistical properties of infant-directed vesus adult-directed speech: insights from speech recognition. *The Journal of the Acoustical Society of America 117*, 2238–2246.

Klatt, D. K. (1976). Linguistic uses of segmental duration in English: acoustic and perceptual evidence. *The Journal of the Acoustical Society of America 59*, 1208–1221.

Kohler, K. (1987). Categorical pitch perception. *Proceedings of the 11th International Congress of Phonetic Sciences, August 1–7, Tallinn, Estonia, vol. 5,* pp. 331–333.

Kohler, K. J. (1991a). A model of German intonation. In *Arbeitsberichte (AIPUK) 25*, pp. 295–360. Kiel: IPDS, Christian-Albrechts-University.

Kohler, K. J. (1991b). Terminal intonation patterns in single-accent utterances of German: Phonetics, phonology and semantics. In *Arbeitsberichte (AIPUK) 25*, pp. 115–185. Kiel: IPDS, Christian-Albrechts-University.

Kohler, K. J. (1995). PROLAB — the Kiel system of prosodic labelling. In K. Elenius and P. Branderud (Eds.), *Proceedings of the XIIIth International Congress of Phonetic Sciences, 13–19 August, Stockholm, Sweden, vol. 3*, pp. 162–165.

Kohler, K. J., B. Peters, and M. Scheffers (Eds.) (2006). *The Kiel Corpus of Spontaneous Speech IV, German: Video Task Scenario (Kiel-DVD1).* Kiel: IPDS, Christian-Albrechts-University.

Kohonen, T. (1984). *Self-Organization and Associative Memory.* Berlin: Springer.

Koiso, H., Y. Horiuchi, S. Tutiya, A. Ichikawa, and Y. Den (1998). An analysis of turn-taking and backchannels based on prosodic and syntactic features in Japanese map task dialogs. *Language and Speech 41*, 295–321.

Kominek, J. and A. Black (2003). CMU ARCTIC databases for speech synthesis. Carnegie Mellon University, Language Technologies Institute, *Tech Report* CMU-LTI-03-177.

Kopell, N. (1988). Toward a theory of modelling central pattern generators. In A. H. Cohen, S. Rossignol, and S. Grillner (Eds.), *Neural Control of Rhythmic Movements in Vertebrates*, pp. 369–413. New York: John Wiley and Sons.

Koreman, J., B. Andreeva, and W. J. Barry (2008). Accentuation cues in French and German. In *Proceedings of the 4th Conference on Speech Prosody*, 6–9 May, Campinas, Brazil, pp. 613–616.

Kristoffersen, G. (1993). An autosegmental analysis of East Norwegian pitch accent. In B. Granström and L. Nord (Eds.), *Nordic Prosody VI,* pp. 109–122. Stockholm: Almquist & Wiksell.

Kristoffersen, G. (1995). Første opponent. *Norsk Lingvistisk Tidsskrift 13*, 91–105.

Kristoffersen, G. (2000). *The Phonology of Norwegian*. Oxford: Oxford University Press.

Kristoffersen, G. (2003). *Norsk prosodi*. Manuscript (third preliminary edition).

Kristoffersen, G. (2007). Dialect variation in East Norwegian tone. In T. Riad and C. Gussenhoven (Eds.), *Tones and Tunes, Vol. 1: Typological and Comparative Studies in Word and Sentence Prosody,* pp. 91–111. Berlin: Mouton de Gruyter.

Krull, D. (1993). Word-prosodic features in Estonian conversational speech: some preliminary results. In *Experiments in speech processes*, pp. 45–54. (PERILUS XVII) Stockholm: Stockholm University, Department of Linguistics.

Krull, D. (1997). Prepausal lengthening in Estonian: evidence from conversational speech. In I. Lehiste and J. Ross (Eds.), *Estonian Prosody: Papers from a Symposium. Proceedings of the International Symposium on Estonian Prosody*, pp. 136–148. Tallinn: Institute of Estonian Language.

Krull, D. (1998). Perception of Estonian word prosody. A study of words extracted from conversational speech. *Linguistica Uralica 24*, 167–171.

Kuhl, P. K., J. E. Andruski, I. A. Chistovich, L. A. Chistovich, E. V. Kozhevnikova, V. L. Ryskina, E. I. Stolyarova, U. Sundberg, and F. Lacerda (1997). Cross-language analysis of phonetic units in language addressed to infants. *Science 277*, 684–686.

Ladd, D. R. (1996). *Intonational Phonology*. (Cambridge Studies in Linguistics 79) Cambridge: Cambridge University Press.

Ladefoged, P. (1967). *Three Areas of Experimental Phonetics*. London: Oxford University Press.

Lahiri, A., A. Wetterlin, and E. Jönsson-Steiner (2005). Lexical specification of tone in North Germanic. *Nordic Journal of Linguistics 28*, 61–96.

Lahiri, A., A. Wetterlin, and E. Jönsson-Steiner (2006). Scandinavian lexical tone: prefixes and compounds. In G. Bruce and M. Horne (Eds.), *Nordic Prosody: Proceedings of the IXth Conference, Lund 2004*, pp. 167–173. Frankfurt am Main: Peter Lang.

Laskowski, K., J. Edlund, and M. Heldner (2008a). An instantaneous vector representation of delta pitch for speaker-change prediction in conversational dialogue systems. In *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2008), March 30 – April 4, Las Vegas, NV*, pp. 5041–5044

Laskowski, K., J. Edlund, and M. Heldner (2008b). Machine learning of prosodic sequences using the fundamental frequency variation spectrum. In *Proceedings of the 4th Conference on Speech Prosody*, *6–9 May, Campinas, Brazil,* pp. 151–154.

Laskowski, K., M. Wölfel, M. Heldner, and J. Edlund (2008c). Computing the fundamental frequency variation spectrum in conversational spoken dialogue systems. In *Proceedings of the 155th Meeting of the Acoustical Society of America, 5th EAA Forum Acusticum, and 9th SFA Congrès Français d'Acoustique (Acoustics 2008), 30 June – 4 July, Paris, France*, pp. 3305–3310.

Laver, J. (1994). *Principles of Phonetics.* Cambridge: Cambridge University Press.

Lee, Y.-S., D. A. Vakoch, and L. H. Wurm (1996). Tone perception in Cantonese and Mandarin: a cross-linguistic comparison. *Journal of Psycholinguistic Research 25*, 527–542.

Lehiste, I. (1960). Segmental and syllabic quantity in Estonian. *American Studies in Uralic Linguistics 1*, 21–82. Bloomington: Indiana University.

Lehiste, I. (1970–1975). Experiments with synthetic speech concerning quantity in Estonian. In: V. Hallap (Ed.), *Congressus Tertius Internationalis Fenno-Ugristarum Tallinnae habitus 17.–23. VIII 1970. Pars I Acta linguistica*, pp. 254–269. Tallinn: Valgus.

Lehiste, I. (1976). Suprasegmental Features of Speech. In N. Lass (Ed.), *Contemporary Issues in Experimental Phonetics*, pp. 225–239. New York/San Francisco/London: Academic Press.

Lehiste, I. (1977). Isochrony reconsidered. *Journal of Phonetics 5*, 253–263.

Lehiste, I. (1997). Search for phonetic correlates in Estonian prosody. In: I. Lehiste and J. Ross (Eds.), *Estonian Prosody: Papers from a Symposium. Proceedings of the International Symposium on Estonian Prosody*, pp. 11–35. Tallinn: Institute of Estonian Language.

Lehiste, I. (2003). Prosodic change in progress: from quantity language to accent language. In P. Fikkert and H. Jacobs (Eds.), *Development in Prosodic Systems*, pp. 47–65. Berlin, New York: Mouton de Gruyter.

Lehiste, I. and D. G. Danforth (1977). Foneettisten vihjeiden hierarkia viron kvantiteetin havaitsemisessa. *Virittäjä 81*, 404–411.

Lehiste, I., P. Teras, V. Ernštreits, P. Lippus, K. Pajusalu, T. Tuisk, and T.-R. Viitso (2008). *Livonian Prosody.* (Mémoires de la Société Finno-Ougrienne 255) Helsinki: Société Finno-Ougrienne.

Lehtonen, J. (1970). *Aspects of Quantity in Standard Finnish.* (Studia Philologica Jyväskyläensia VI) Jyväskylä: University of Jyväskylä.

Liiv, G. (1961). Eesti keele kolme vältusastme vokaalide kestus ja meloodiatüübid. *Keel ja Kirjandus 7*, 412–424, and *8*, 480–490.

Lindblom, B. E. F. (1968). Temporal organization of syllable production. *Quarterly Progress and Status Report*, 9(2–3), pp. 1–5.

Lippus, P., K. Pajusalu and J. Allik (2007). The tonal component in perception of the Estonian quantity. In J. Trouvain and W. J. Barry (Eds.), *Proceedings of the 16th International Congress of Phonetic Sciences, 6–10 August, 2007, Saarbrücken, Germany*, pp. 1049–1052.

Liu, F. and Y. Xu (2005). Parallel encoding of focus and interrogative meaning in Mandarin intonation. *Phonetica 62*, 70–87.

Lively, S. E., J. S. Logan, and D. B. Pisoni (1993). Greater variety also improves phonetic category learning in foreign speakers. *The Journal of the Acoustical Society of America 94*, 1242–1255.

Local, J. K. and J. Kelly (1986). Projection and 'silences': notes on phonetic and conversational structure. *Human Studies 9*, 185–204.

Local, J. K., J. Kelly, and W. H. G. Wells (1986). Towards a phonology of conversation: turn-taking in Tyneside English. *Journal of Linguistics 22*, 411–437.

Lorentz, O. (2008). Tonelagsbasis i norsk. *Maal og Minne 1* (2008), 50–68.

Lundgren, A. (2005). HMM-baserad talsyntes. Master's Thesis. Stockholm: Royal Institute of Technology, Department of Speech, Music and Hearing.

Marcus, S. M. (1981). Acoustic determinants of perceptual center (P-center) location. *Perception and Psychophysics 30*, 247–256.

Markusa, D. (1991). Syllable tonemes in Latvian. In *Proceedings of the XIIth International Congress of Phonetic Sciences, August 19–24, Aix-en-Provence, France, Vol. 2*, pp. 242–245.

Matte, A. C. F. (2004). Relating Emotional Content to Speech Rate in Brazilian Portuguese In B. Bel and I. Marlien (Eds.), *Proceedings of Speech Prosody 2004, March 23–26, Nara, Japan,* pp. 663–666.

Matusevich, M. I. (1976). *Sovremennyj russkij yazyk. Fonetika*. Moskva: Izd. Prosveschenie.

Meister, E. and S. Werner (2006). Intrinsic microprosodic variations in Estonian and Finnish: acoustic analysis. In R. Aulanko, L. Wahlberg, and M. Vainio (Eds.), *Fonetiikan päivät 2006 – The Phonetics Symposium 2006*, pp. 103–112. (Publications of the Department of Speech Sciences 53) Helsinki: University of Helsinki

Meyer, E. A. (1937). *Die Intonation im Schwedischen: Die Sveamundarten.* (Studies Scand. Philol. 10) Stockholm universitet.

Mixdorff, H. (2000). A novel approach to the fully automatic extraction of Fujisaki model parameters. In *Proceedings of the IEEE International*

Conference on Acoustics, Speech, and Signal Processing (ICASSP2000), vol. 3, Istanbul, pp. 1281–1284.

Mixdorff, H. (1/10/2009). FujiParaEditor: http://www.tfhberlin.de/~mixdorff/thesis/fujisaki.html. TFH Berlin University of Applied Sciences.

Mixdorff, H. and H. Fujisaki (2000). A quantitative description of German prosody offering symbolic labels as a by-product. In *Proceedings of ICSLP 2000, Beijing, vol. 2*, pp. 98–101.

Mixdorff, H. and H. R. Pfitzinger (2005). Analysing fundamental frequency contours and speech rate in Map Task dialogs. *Speech Communication 46*, 310–325.

Mixdorff, H., H. R. Pfitzinger, and K. Grauwinkel (2005). Towards objective measures for comparing speaking styles. In *Proceedings of the 10th International Conference on Speech and Computer, Patras, Greece*, pp.131–134.

Nadeina, T. M. (2004). Nauchnaya deyatel'nost' professora L. B. Zlatoustovoj. In G. E. Kedrova and V. V. Potapov (Eds.), *Yazyk i rech': problemy i resheniya: Sbornik nauchnyh trudov k yubileyu professora L. V. Zlatoustovoj*, pp. 7–23. Moskva: MAKS Press.

Nam, H. and E. Saltzman (2003). A competitive, coupled oscillator model of syllable structure. In M. J. Solé, D. Recasens, and J. Romero (Eds.), *Proceedings of the 15th International Congress of Phonetic Sciences, Barcelona*, *3–9 August,* pp. 2253–2256.

Niebuhr, O. (2006). The role of the accented-vowel onset in the perception of German early and medial peaks. In *Proceedings of the 3rd International Conference on Speech Prosody, Dresden*, pp. 109–112.

Niebuhr, O. and G. Ambrazaitis (2006). Alignment of medial and late peaks in German spontaneous speech. In *Proceedings of the 3rd International Conference on Speech Prosody, Dresden*, pp. 161–164.

Nilsen, R. A. (1992). *Intonasjon i interaksjon – sentrale spørsmål i norsk intonologi.* Lingvistisk institutt, Universitetet i Trondheim.

Nilsen, R. A. (1995). Svar frå doktoranden. *Norsk Lingvistisk Tidsskrift 13*, 111–122.

Noguchi, H. and Y. Den (1998). Prosody-based detection of the context of backchannel responses. In *Proceedings of the Fifth International Conference on Spoken Language Processing, Sydney, Australia (ICSLP '98),* pp. 487–490.

Nolan, F. (2003). Intonational equivalence: an experimental evaluation of pitch scales. In M. J. Solé, D. Recasens, and J. Romero (Eds.), *Proceedings of the 15th International Congress of Phonetic Sciences, Barcelona*, *3–9 August,* pp. 771–774.

Norwine, A. C. and O. J. Murphy (1938). Characteristic time intervals in telephonic conversation. *The Bell System Technical Journal 17*, 281–291.

O'Connor, J. D. (1973). *Phonetics*. Harmondsworth: Penguin.

O'Dell, M., M. Lennes, and T. Nieminen (2008). Hierarchical levels of rhythm in conversational speech. In P. A. Barbosa, S. Madureira, and C. Reis (Eds.), *Speech Prosody 2008: Fourth Conference on Speech Prosody, Campinas, Brazil*, pp. 355–358.

O'Dell, M. L., M. Lennes, S. Werner, and T. Nieminen (2007). Looking for rhythms in conversational speech. In J. Trouvain and W. J. Barry (Eds.), *Proceedings of the 16th International Congress of Phonetic Sciences, 6–10 August, Saarbrücken, Germany*, pp. 1201–1204.

O'Dell, M. and T. Nieminen (1998). Reasons for an underlying unity in rhythm dichotomy. *Linguistica Uralica 3*, 178–185.

O'Dell, M. L. and T. Nieminen (1999). Coupled oscillator model of speech rhythm. In J. J. Ohala, Y. Hasegawa, M. Ohala, D. Granville, and A. C. Bailey (Eds.), *Proceedings of the 14th International Congress of Phonetic Sciences, San Francisco, 1–7 August,* pp. 1075–1078.

O'Dell, M. and T. Nieminen (2001). Speech rhythms as cyclical activity. In S. Ojala and J. Tuomainen (Eds.), *21. Fonetiikan päivät Turku 4.–5.1.2001*, pp. 159–168. (Publications of the Department of Finnish and General Linguistics 67) Turku: University of Turku. (A corrected version of this article can be found at http://papers.legisign.org/articles/odell-nieminen2001.pdf.)

O'Dell, M. and T. Nieminen (2002a). How long is a stress group? *Cadernos de Estudos Lingüísticos 43,* 93–108.

O'Dell, M. and T. Nieminen (2002b). Rytmijakson pituus oskillaattorimallissa. In P. Korhonen (Ed.), *Fonetiikan Päivät 2002 / The Phonetics Symposium 2002*, pp. 195–204. Helsinki University of Technology, Laboratory of Acoustics and Audio Signal Processing.

O'Dell, M. and T. Nieminen (2006). Tahdin ajoitus suomessa oskillaattorimallin näkökulmasta. In R. Aulanko, L. Wahlberg, and M. Vainio (Eds.), *Fonetiikan päivät 2006 / The Phonetics Symposium 2006*, pp. 134–143. (Publications of the Department of Speech Sciences 53) Helsinki: University of Helsinki.

Ogden, R. (2001). Turn transition, creak and glottal stop in Finnish talk-in-interaction. *Journal of the International Phonetic Association 31*, 139–152.

Öhman, S. (1967). Word and sentence intonation: A quantitative model. *STL-QPSR 2–3*, 20–54.

Orlikoff, R. F. and R. J. Baken (1993). *Clinical Speech and Voice Measurement: Laboratory Exercises.* San Diego: Singular Publishing Group.

Pajupuu, H. and T.-R. Viitso (1986). Livonian polyphtongs. *Estonian Papers in Phonetics 1984–1985*, 96–130.

Papousek, M. and S. C. Hwang (1991). Tone and intonation in Mandarin babytalk to presyllabic infants: comparison with registers of adult conversation and foreign language instruction. *Applied Psycholinguistics 12*, 481–504.

Penttilä, A. (1963). *Suomen kielioppi.* Porvoo/Helsinki: WSOY.

Penttilä, A. and L. Posti (1941). Über die steigende und sog. Stossintonation im Livischen. *Finnisch-Ugrische Forschungen XVII*, 235–272.

Peters, B. (2006). Form und Funktion prosodischer Grenzen im Gespräch. Ph. D. thesis. Kiel: IPDS, Christian-Albrechts-University.

Peterson, G. E. and H. L. Barney (1952). Control methods used in a study of the vowels. *The Journal of the Acoustic Society of America 24,* 175–184.

Pike, K. L. (1945). The Intonation of American English. University of Michigan Press, Ann Arbor.

Pike, K. L. (1972). The intonation of American English. In D. Bolinger (Ed.), *Intonation*, pp. 53–83. Harmondsworth: Penguin.

Pollack, I. (1952). The information of elementary auditory displays. *The Journal of the Acoustical Society of America 24*, 745–749.

Pompino-Marschall, B. (1989). On the psychoacoustic nature of the P-center phenomenon. *Journal of Phonetics 17*, 175–192.

Posti, L. (1936). Liivin kielen intonaatioista. *Virittäjä 40*, 314–326, 384.

Posti, L. (1937). Lisää liivin intonaatioista. *Virittäjä 41*, 189–200, 263.

Praeger, K. J. (1995). *The Psychology of Intimacy.* New York: Guilford Press.

Prom-on, S., Y. Xu, and B. Thipakorn (2006a). Functional-oriented articulatory modeling of tones and intonations. In *Proceedings of the 3rd International Conference on Speech Prosody, May 2–5, Dresden, Germany*, PS2-14_0089.

Prom-on, S., Y. Xu, and B. Thipakorn (2006b). Quantitative Target Approximation model: simulating underlying mechanisms of tones and intonations. In *Proceedings of The 31st International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2006), May 14–19, Toulouse, France*. pp. I-749–752.

Prom-on, S., Y. Xu, and B. Thipakorn (submitted). Modeling tone and intonation in Mandarin and English as a process of target approximation.

Räsänen, O. J. (2007). Speech Segmentation and Clustering Methods for a New Speech Recognition Architecture. Master's thesis. Helsinki University of Technology, Laboratory of Acoustics and Audio Signal Processing, http://lib.tkk.fi/Dipl/2007/urn010123.pdf.

Reinhart, T. (1981). Pragmatics and linguistics: An analysis of sentence topics. *Philosophica 27*, 53–94.

Remmel, M. (1975). *The phonetic scope of Estonian: some specifications.* Preprint KKI-5. Academy of Sciences of the Estonian S.S.R. Academy of Sciences of the Estonian S.S.R. Tallinn: Institute of Language and Literature.

Riad, T. (1998). Towards a Scandinavian tone accent typology. In W. Kehrein and R. Wiese (Eds.), *Phonology and Morphology of the Germanic Languages*, pp. 77–109. (Linguistische Arbeiten 386) Tübingen: Niemeyer.

Riad, T. (2003). Diachrony of the Scandinavian accent typology. In P. Fikkert and H. Jacobs (Eds.), *Development in Prosodic Systems*, pp. 91–144. (Studies in Generative Grammar 58). Berlin/New York: Mouton de Gruyter.

Riad, T. (2005). Historien om tonaccenten. In C. Falk and L.-O. Delsing (Eds.), *Svenska språkets historia 8*, pp. 1–27. Lund: Studentlitteratur.

Riad, T. (2006). Scandinavian accent typology. *Sprachtypologie und Universalienforschung (STUF) 59*, 36–55.

Riad, T. (2008a). "Börk börk börk. Ehula hule de chokolad muus". *Språktidningen nr 2*, 34–39.

Riad, T. (2008b). *Prosodi i svenskans ordbildning*. Ms. Stockholm University.

Riad, T. and M. Segerup (2008). Phonological association of tone. Phonetic implications in West Swedish and East Norwegian. In *Proceedings Fonetik 2008,* pp. 93–96. Göteborg: University of Gothenburg, Department of Linguistics.

Rialland, A. (2001). Anticipatory raising in downstep realization: Evidence for preplanning in tone production. In S. Kaji (Ed.), *Proceedings of the Symposium Cross-Linguistic Studies of Tonal Phenomena: Tonogenesis, Japanese Accentology, and Other Topics*, pp. 301–321. Tokyo: Tokyo University of Foreign Studies, Institute for Languages and Cultures of Asia and Africa.

Rischel, J. (1960). Über die phonematische und morphophonematische Funktion der sogenannten Worttöne in Norwegischen. *Zeitschrift für Phonetik und allgemeine Sprachwissenschaft 13*, 177–185. [Norwegian translation printed in E. H. Jahr and O. Lorentz (Eds.), *Prosodi/Prosody*. (Studies in Norwegian Linguistics 2) Oslo: Novus, 1983.]

Rischel, J. (1963). Morphemic tone and word tone i Eastern Norwegian. *Phonetica 10*, 154–164. [Reprinted in E. H. Jahr and O. Lorentz (Eds.), *Prosodi/Prosody*. (Studies in Norwegian Linguistics 2) Oslo: Novus, 1983.]

Roll, M. (2004). Prosodic cues to the syntactic structure of subordinate clauses in Swedish. ms. Department of Linguistics and Phonetics, Lund University.

Roll, M. (2006). Prosodic cues to the syntactic structure of subordinate clauses in Swedish. In G. Bruce and M. Horne (Eds.), *Nordic Prosody: Proceedings of the IXth conference, Lund 2004*, pp. 195–204. Frankfurt am Main: Peter Lang.

Roll, M., M. Horne, and M. Lindgren (2009). Left-edge boundary tone and main clause verb effects on syntactic processing in embedded clauses – An ERP study. *Journal of Neurolinguistics 22*, 55–73.

Routarinne, S. (2003). *Tytöt äänessä. Parenteesit ja nouseva sävelkulku kertojan vuorovaikutuskeinoina*. Helsinki: Suomalaisen Kirjallisuuden Seura.

Saltzman, E., H. Nam, J. Krivokapic, and L. Goldstein (2008). A task-dynamic toolkit for modeling the effects of prosodic structure on articulation. In P. A. Barbosa, S. Madureira, and C. Reis (Eds.), *Speech Prosody 2008: Fourth Conference on Speech Prosody, Campinas, Brazil*, pp. 175–184.

Salvi, G. (2008). SAMPA codes used for the Swedish SpeechDat: STA. http://www.speech.kth.se/~giampi/downloads/sampa_latex_sv.pdf.

Schaefer-Vincent, K. (1982). Signicant points: Pitch period detection as a problem of segmentation. *Phonetica 39*, 241–253.

Schaefer-Vincent, K. (1983). Pitch period detection and chaining: Method and evaluation. *Phonetica 40*, 177–202.

Scheffers, M. T. M. (1983). Simulation of auditory analysis of pitch: An elaboration on the DWS pitch meter. *The Journal of the Acoustical Society of America 74*, 1716–1725.

Schegloff, E. (1982). Discourse as an interactional achievement: some uses of 'uh huh' and other things that come between sentences. In D. Tannen (Ed.), *Analyzing Discourse: Text and Talk*, pp. 71–93. Washington, D.C.: Georgetown University Press.

Schellenberg, E. G. and S. E. Trehub (2008). Is there an Asian advantage for pitch memory? *Music Perception 25*, 241–252.

Scherba, L. V. (1957). *Fonetika frantsuzskogo yazyka*. Moskva: Izd. literatury na inostrannyh yazykah.

Scherer, K. (1995). Expression of emotion in voice and music. *Journal of Voice 9*, 235–248.

Schuetze-Coburn, S., M. Shapley, and E. G. Weber (1991). Units of intonation in discourse: a comparison of acoustic and auditory analyses. *Language and Speech 34*, 207–234.

Selenius, E. (1972). *Västnyländsk ordaccent*. (Studier i nordisk filologi 59) Helsingfors: Svenska Litteratursällskapet i Finland.

Selkirk, E. (1995). Sentence prosody: intonation, stress, and phrasing. In J. A. Goldsmith (Ed.), *The Handbook of Phonological Theory*, pp. 550–569. Cambridge, MA, and Oxford, UK: Blackwell.

Selkirk, E. (2000). The interaction of constraints on prosodic phrasing. In M. Horne (Ed.), *Prosody Theory and Experiment Studies Presented to Gösta Bruce*, pp. 231–261. Dordrecht: Kluwer Academic Publishers.

Sellen, A. J. (1995). Remote conversations: the effects of mediating talk with technology. *Human-Computer Interaction 10*, 401–444.

Selting, M. (1996). On the interplay of syntax and prosody in the constitution of turn-constructional units and turns in conversation. *Pragmatics 6*, 357–388.

Shriberg, E., R. Bates, A. Stolcke, P. Taylor, D. Jurafsky, K. Ries, et al. (1998). Can prosody aid the automatic classification of dialog acts in conversational speech. *Language and Speech 41*, 443–492.

Sievers, E. (1893). *Grundzüge der Phonetik zur Einführung in das Studium der Lautlehre der indogermanischen Sprachen.* (4th Ed.) Leipzig: Breitkopf & Härtel.

Skrelin, P. and N. Volskaya (2004). Russian read and spontaneous speech: prosodic data analysis. In G. Bruce and M. Horne (Eds.), *Nordic Prosody: Proceedings of the IXth Conference*, *Lund 2004*, pp. 235–244. Frankfurt am Main: Peter Lang.

Slethei, K. (1996). *Grunnbok i fonetikk for språkstudenter.* Oslo: Cappelen Akademisk Forlag.

Sluijter, A. M. C. and V. J. van Heuven (1996). Acoustic correlates of linguistic stress and accent in Dutch and American English. In *Proceedings of the Fourth International Conference on Spoken Language Processing (ICSLP 96)*, pp. 630–633.

Smith, J. L. (2001). Lexical category and phonological contrast. In R. Kirchner, J. Pater, and W. Wikely (Eds.), *PETL 6: Proceedings of the Workshop on the Lexicon in Phonetics and Phonology*, pp. 61–72. Edmonton: University of Alberta.

Somervuo P. and T. Kohonen (2004). Self-organizing maps and Learning Vector Quantization for feature sequences. *Neural Processing Letters 10*, 151-159.

Sperber, D. and D. Wilson (1995). *Relevance. Communication and Cognition.* Oxford: Blackwell.

Stern, D.N., S. Spieker, R. K. Barnett, and K. MacKain (1983). The prosody of maternal speech: infant age and context related changes. *Journal of Child Language 10*, 1–15.

Suhonen, S. (1982). Über die Quantität der livischen Vokale. In: A. Iivonen, S. Suhonen, and P. Virtaranta (Eds.), *Voces amicorum Sovijärvi. In Honorem Antti Sovijärvi septuagesimum annum agentis die XXII mensis aprilis anno MCMLXXXII*, pp. 295–306. (Mémoires de la Société Finno-Ougrienne 181) Helsinki: Société Finno-Ougrienne.

Sulkala, H. and M. Karjalainen (1992). *Finnish*. London/New York: Routledge.

Suomi, K. (1980). *Voicing in English and Finnish Stops. A Typological Comparison with an Interlanguage Study of Two Languages in Contrast.* (Publications of the Department of Finnish and General Linguistics 10) Turku: University of Turku.

Suomi, K. (2007). On the tonal and temporal domains of accent in Finnish. *Journal of Phonetics 35*, 50–55.

Suomi, K., J. Toivanen, and R. Ylitalo (2003). Durational and tonal correlates of accent in Finnish. *Journal of Phonetics 31*, 113–138.

Suomi, K. and R. Ylitalo (2004). On durational correlates of word stress in Finnish. *Journal of Phonetics 32*, 35–63.

Svärd, N. (2001). Word accents in the Närpes dialect: Is there really only one accent? *Papers from FONETIK 2001*, pp. 160–163. (Working Papers 49) Lund: Lund University, Department of Linguistics.

Svetozarova, N. D. (1998). Intonation in Russian. In D. Hirst and A. Di Cristo (Eds.), *Intonation Systems. A Survey of Twenty Languages*, pp. 261–274. Cambridge: Cambridge University Press.

Sweet, H. (1877). *A Handbook of Phonetics*. Oxford: Clarendon Press.

Takahashi, T. and L. M. Beebe (1987). The development of pragmatic competence by Japanese learners of English. *JALT Journal (Japanese Association for Language Teaching) 8*, 131–155.

Talkin, D. (1995). A robust algorithm for pitch tracking (RAPT). In W. B. Kleijn and K. K. Paliwal (Eds.), *Speech Coding and Synthesis*, pp. 495–518. New York: Elsevier.

Tench, P. (1996). *The Intonation Systems of English.* London: Cassell.

’t Hart, J., R. Collier, and A. Cohen (1990). *A Perceptual Study of Intonation. An Experimental-phonetic Approach to Speech Melody*. Cambridge: Cambridge University Press.

Thiessen, E. D., E. A. Hill, and J. R. Saffran (2005). Infant-directed speech facilitates word segmentation. *Infancy 7*, 53–71.

Thomsen, V. (1890). *Beröringer mellem de finske og de baltiske (litauisk-lettiske) Sprog*. København.

Toivanen, J. (1995). *Perspectives of Intonation and Paralanguage: English, Finnish and Finnish English Interlanguage.* Unpublished Licentiate Thesis, Department of English, University of Oulu.

Toivanen, J. (2001). *Perspectives on Intonation: English, Finnish and English Spoken by Finns.* Frankfurt am Main: Peter Lang.

Toivanen, J. and T. Waaramaa (2005). Tone choice and voice quality of dispreferred turns in the English of Finns. *Logopedics Phoniatrics Vocology 30*, 181–184.

Tokuda, K., H. Zen, and A. Black (2002). An HMM-based speech synthesis system applied to English. In *Proceedings of 2002 IEEE Workshop on Speech Synthesis, 11–13 September, Santa Monica, CA,* pp. 227–230.

Traunmüller, H. and D. Krull (2003). The effect of local speaking rate on the perception of quantity in Estonian. *Phonetica 60*, 187–207.

Truckenbrodt, H. (1999). On the relation between syntactic phrases and phonological phrases. *Linguistic Inquiry 30*, 219–255.

Truckenbrodt, H. (2004). Final lowering in non-final position. *Journal of Phonetics 32*, 313–348.

Vainio, M. and J. Järvikivi (2006). Tonal features, intensity, and word order in the perception of prominence. *Journal of Phonetics 34*, 319–342.

Vainio, M. and J. Järvikivi (2007). Focus in production: tonal shape, intensity and word order. *The Journal of the Acoustical Society of America 121*, EL55–EL61.

Vainio M., J. Järvikivi, and S. Werner (2004). Intensity and word order in the perception and production of Finnish prosody. In T. Seppänen, K. Suomi, and J. Toivanen (Eds.), *Fonetiikan päivät 2004*, pp. 52–55. Oulu: Oulun yliopistopaino.

Välimaa-Blum, R. (1993). A pitch accent analysis of intonation in Finnish. *Ural-Altaische Jahrbücher 12*, 82–94.

Vallduví, E. and E. Engdahl (1996). The linguistic realisation of information packaging. *Linguistics 34*, 459–519.

Vallduví, E. and M. Vilkuna (1998). On Rheme and Kontrast. *Syntax and Semantics 29*, 79–108.

van de Weijer, J. (2001). Vowels in infant- and adult-directed speech. In A. Karlsson and J. van de Weijer (Eds.), *Papers from Fonetik 2001*, pp. 172–175. (Working Papers 49) Lund: Lund University, Department of Linguistics.

van Dommelen, W. A. and O. A. Husby (2008). Perception of Norwegian word tones by Chinese and German listeners. In *New Sounds 2007: Proceedings of the Fifth International Symposium on the Acquisition of Second Language Speech, November 25–28, Florianópolis, Brazil,* pp. 165–174.

Vihman, M. M. (1971). *Livonian Phonology, with an Appendix on Stød in Danish and Livonian*. Academic dissertation. University of California, Berkeley.

Viitso, T.-R. (1981). *Läänemeresoome fonoloogia küsimusi*. Tallinn: Eesti NSV TA, Eesti Keele ja Kirjanduse Instituut.

Vilkuna, M. (1989). *Free Word Order in Finnish: Its Syntax and Discourse Functions*. Helsinki: Suomalaisen Kirjallisuuden Seura.

Vilkuna, M. (1995). Discourse Configurationality in Finnish. In K. É. Kiss (Ed.), *Discourse Configurational Languages*, pp. 244–268. New York/Oxford: Oxford University Press.

Vol'skaya, N. B. and S. B. Stepanova (2005). O nekotoryh problemah sintagmaticheskogo chleneniya spontannogo teksta. *Materialy XXXIV nauchno-metodicheskoj konferencii SPBGU 2005*, pp. 16–24.

Volskaya, N. B., N. D. Svetozarova, and P. A. Skrelin (1997). Modelirovanie prosodicheskogo oformlenia russkogo teksta. *Bulleten foneticheskogo fonda russkogo jazyka 6*, 85–110

Wang, Y., A. Jongman, and J. Sereno (2003). Acoustic and perceptual evaluation of Mandarin tone productions before and after training. *The Journal of the Acoustical Society of America 113*, 1033–1043.

Wang, Y., M. Spence, A. Jongman, and J. Sereno (1999). Training American listeners to perceive Mandarin tones. *The Journal of the Acoustical Society of America 106*, 3649–3658.

Ward, N. (2004). Pragmatic functions of prosodic features in non-lexical utterances. In B. Bel and I. Marlien (Eds.), *Proceedings of Speech Prosody 2004, March 23–26, Nara, Japan,* pp. 325–328.

Ward, N. and W. Tsukahara (2000). Prosodic features which cue back-channel responses in English and Japanese. *Journal of Pragmatics 32*, 1177–1207.

Wayland, R. P. and S. G. Guion (2004). Training English and Chinese listeners to perceive Thai tones: a preliminary report. *Language Learning 54*, 681–712.

Wayland, R. P. and B. Li (2008). Effects of two training procedures in cross-language perception of tones. *Journal of Phonetics 36*, 250–267.

Wetterlin, A. (2007). The Lexical Specification of Norwegian Tonal Word Accents. PhD diss. University of Konstanz.

Wiedemann, F. J. (1861). *Joh. Andreas Sjögren's Livische Grammatik nebst Sprachproben*. Joh. Andreas Sjögren's Gesammelte Schriften, Band II, Theil I. St. Petersburg.

Wiik, K. (1989). *Liivin katko*. Turku: Turun Yliopisto, Fonetiikka.

Xu, Y. (1999). Effects of tone and focus on the formation and alignment of F0 contours. *Journal of Phonetics 27*, 55–105.

Xu, Y. and Q. E. Wang (2001). Pitch targets and their realization: evidence from Mandarin Chinese. *Speech Communication 33*, 319–337.

Yngve, V. H. (1970). On getting a word in edgewise. In *Papers from the Sixth Regional Meeting Chicago Linguistic Society,* pp. 567–578. Chicago, IL, USA: Chicago Linguistic Society.

Yudong, C. (2007). From tone to accent: the tonal transfer strategy for Chinese L2 learners of Spanish. In J. Trouvain and W. J. Barry (Eds.), *Proceedings of the 16th International Congress of Phonetic Sciences, 6–10 August, Saarbrücken, Germany*, pp. 1645–1648.

Zharkov I. V., S. L. Slobodyanuk, and N. D. Svetozarova (1994). Avtomatichesky akcentno-intonacionny transkriptor proizvol'nogo russkogo teksta. *Bulleten foneticheskogo fonda russkogo jazyka 5*, 58–71

Zlatoustova, L. V. (1962). *Foneticheskaya struktura slova v potoke rechi*. Kazan'.