ISCA Archive
http://www.isca-speech.org/archive

INTERSPEECH 2012
ISCA's 13th Annual Conference
Portland, OR, USA
September 9-13, 2012

Language differences in the perceptual weight of prominence-lending properties

Bistra Andreeva, William Barry, Magdalena Wolska

Department of Computational Linguistics & Phonetics,
Saarland University, Germany
{andreeva, wbarry, magda}@coli.uni-saarland.de

Abstract

A Bulgarian and a German sentence were presented to Bulgarian and German listeners together with a question which either expected an early narrow focus or a late narrow focus. The answering sentences were manipulated so that the word in the late-focused position ranged from completely de-accented to strongly accented. The early focused position was neutral, allowing late-focus perception with late strong accentuation and early focus with de-accentuation of the late-focus position. Accentuation strength of the late position was varied by changing the duration, intensity and *f0* values individually between accented and low de-accented. Subjects were asked to judge the suitability of the answers to the question. Results show the relative contribution of the three parameters to the acceptability of the word in the late focus position as focally accented or de-accented. Differences between Bulgarian and German in the relative weighting of the parameters are revealed.

Index Terms: focal accentuation, perception, parameter weighting, cross-language differences

10.21437/Interspeech.2012-635

1. Introduction

The phonetic basis of prominence (i.e., both from lexical stress and accentuation within utterances) has long been accepted as comprising the relative duration, *f0* (difference or movement), intensity and spectral properties of the (vocalic) unit. Excitation quality as an additional possible correlate (cf. [1] and [2]) has rarely been considered. Of the traditional parameters, duration and *f0* have been shown experimentally (for English) to be more important in perceived prominence than intensity and degree of spectral reduction [3, 4]. However, the contribution of *f0* to prominence has not been borne out in analyses of large speech databases [5, 6]. In agreement with received wisdom, the simple dB measure of syllable strength did not prove important, but more refined measures of signal energy suggest a revision of earlier assumptions. Van Kuijk & Boves [5] found that, classifying (lexically) stressed and unstressed vowels separately for each vowel phoneme, either a combined value of intensity and duration, duration alone, or a spectral tilt measure performed best. In linguistically careful controlled data Sluijter & van Heuven [7] and Sluijter et al. [8] also found both acoustic-correlate and perceptual validity for spectral tilt in experiments on stress and accent. Kochanski et al. [6] found that their acoustic "loudness" measure (based on [9]) was the primary correlate of accentuation, more important even than duration. In contrast, Streefkerk et al. [10], using the same database as in [5], found that the traditionally more important parameters, *f0* range and duration were the best predictors of perceived prominence. These studies, however, all had binarily labelled databases (±prominent auditory judgements /±lexical stress derived from the lexicon) rather than differentiated judgements of greater or lesser prominence to base their analyses on. We have to note discrepant results between the studies, but cannot say whether they stem from differences in material (full-band vs. telephone-quality speech) or differences in the language material (Dutch vs. English), or different approaches to the analysis (auditory judged prominence vs. lexical stress).

Much more differentiated estimates of relative prominence have been used by Fant & Kruckenberg [11, 12], and Eriksson [13, 14]. Indications of the contribution of energy, or effort (and corresponding acoustic measures) to perceived prominence were again found. On the basis of syllable by syllable judgments on a 0-30 scale, Fant & Kruckenberg [11] found the strongest single correlation between prominence and their spectral tilt parameter, with duration also very robust. Vocal effort as measured by subglottal pressure covaried strongly with spectral tilt. Eriksson [13] also found „vocal effort", pitch and duration to be the main predictive acoustic properties for prominence judged on a continuous scale, but they accounted for no more than half the variance. "Top-down" processing alone could have given the same result. Comparing Swedish listeners and English listeners judging Swedish Eriksson [14] showed the effect of prior expectations (prediction of prominence from text alone). English listeners had higher correlations between prominence and acoustic parameters (i.e., less top-down influence, understandably). However, beside the difference in the level of top-down effects, there were also clear indications of different parameter weightings underlying the prominence judgements: Swedish listeners attached more importance to effort, whereas the English listeners attached roughly equal weight to effort, pitch and duration.

Recent *production* studies have shown significant differences between languages in the degree to which the four basic properties (duration, *f0*, intensity and spectral properties) are exploited to signal phrasal prominence [15, 16]. Some differences in *perceptual sensitivity* across languages have also been shown in [17], but no *functional* effect was found in an information-structural task. In Cumming [18], however, functionally relevant differences between Swiss German, Swiss French and Standard French were found in the perceptual weighting of *f0* and duration signaling syntactic juncture in a digit and letter grouping task. The discrepancy between the perceptual studies may result from the task difference, but it may stem from the use of unmanipulated stimuli in the former study, in which the properties varied in a correlated way, while the latter used resynthesized stimuli with independently modified parameters.

This experiment therefore modifies the acoustic properties (duration, intensity and *f0*) independently of one another to provide a more differentiated array of stimuli to tease out a functionally relevant effect of differences in perceptual sensitivity in the information-structural task.

2. Stimulus material

A Bulgarian and a German sentence: "*Igrax na dama bez kaka ti*" (I played draughts without your older sister) and "*Der Mann fuhr den Wagen vor*" (The man brought the car round) were spoken in response to two questions requiring i) "*kaka*" (older sister) or "*Wagen*" (car) to be in narrow focus and therefore strongly accented and ii) "*dama*" (draughts) or "*Mann*" (man) to be in narrow focus with the post-nuclear "kaka" or "Wagen" de-accented.

The second part of the two versions of each sentence ("*...bez kaka ti*" and "*… fuhr den Wagen vor*") were spliced on to the first part of a broad-focus rendering of the corresponding sentence. The accentuation of the first half was thus neutral in its prediction of the accentual pattern of the second part. A de-accented "*kaka*" or "*Wagen*" made the prenuclear accent on "*dama*" or "*Mann*" completely acceptable as an early narrow focus (ENF). With a late narrow focus (LNF) on "*kaka*" or "*Wagen*", the prenuclear accent from the broad focus rendering remains a, prenuclear accent, which is an acceptable alternative to a de-accented "*dama*" or "*Mann*" in Bulgarian and German for an LNF utterance.

Signal manipulation was carried out as a single-step weakening or strengthening procedure for each parameter separately. The LNF values for intensity (*int*) and *f0* were reduced to those of the de-accented version (A(ccented)*int* to D(e-accented)*int* and A*f0* to D*f0*) and the ENF values of the postnuclear de-accented versions were increased to the LNF values (D*int* to A*int* and D*f0* to A*f0*). In the case of duration (*dur*), the de-accented values sounded very unnatural in any combination except the D values for intensity and *f0*. Therefore an intermediate duration was selected for D*dur* which was clearly shorter than A*dur* but remained acceptable in combination with A*int* and A*f0*. The reason for manipulating both the accented and the de-accented syllables was to offer stimuli with and without natural non-phonological vowel-quality reduction. The manipulations resulted in 8 combinations of A and D parameter values for each version (LNF and ENF).

Stimuli were offered in groups of 4 for comparative judgment of acceptability in the given question-answer context. In each group one parameter was kept constant in all four stimuli and the other two were varied, individually and together: e.g. for the LNF condition with accented *kaka/Wagen* as point of departure the four stimuli were:

(1) (a) A*dur* A*int* A*f0*[1], (b) A*dur* D*int* A*f0*,
(c) A*dur* A*int* D*f0*, (d) A*dur* D*int* D*f0*.

Theoretically, stimulus (a) in the group would receive the highest acceptability judgment in the context of a question expecting an LNF answer and the lowest acceptability after a question expecting an ENF answer. Responses to stimuli (b) and (c) would indicate the degree to which D*int* and D*f0* alone reduce or improve acceptability in the two contexts. Responses to stimulus (d) should indicate the effect of combined weakening of the parameters.

Each group was offered twice, once with a question suited to an LNF answer ("*Bez kogo igra na dama*?" Without whom did you play draughts? "*Was fuhr der Mann vor*?" What did the man bring round?) and once with a question suited to an ENF answer with a de-accented target ("*Na kakvo igra bez kaka ti*?" What did you play without your older sister? "*Wer fuhr den Wagen vor*?" Who brought the car round?). The total number of groups offered for judgment of the "best suited" answer to the question was 24: 2 stimuli (LNF-eliciting, ENF-eliciting) x 2 values (A, D) x 2 questions (for early and late focus) x 3 parameters (duration, intensity, *f0*).

3. Subjects and task

Twenty-six German and twenty-seven Bulgarian native speakers took part in the experiment. The stimuli were presented over headphones, and icons for the question and each of the 4 answer variants in the group appeared on a computer monitor. The subject could play the question and any answer by double-clicking the icons. There was no constraint on the number of times question and answers could be heard. Subjects were told to move each of the answer icons up towards the upper limit line, defined as "suits the question perfectly" or down towards the base line, defined as "does not suit the question at all" according to the degree to which they judged the answer suited the question. Thus the measurements were on a continuous scale, from 0 to 100.

4. Analysis procedures

Preliminary study of the response patterns revealed a much greater degree of variation between the Bulgarian subjects than between the German subjects. Correlations were calculated for each subject between stimuli with AAA and DDD values for answers to questions expecting an ENF and an LNF answer. Systematic responses should result in a significant negative correlation. Twelve of the Bulgarian subjects revealed non-significant and in many cases negligible correlations and had therefore to be excluded from further analysis. Further tests are planned to ascertain possible reasons for this quasi-random perceptual behaviour.

Repeated measures Analyses of Variance were performed for the remaining 15 Bulgarian and 26 German subjects on each of the 24 stimulus groups with stimulus as the intra-subject variable. Greenhouse-Geisser correction was used if sphericity assumption was not met (Mauchly's test). α was set at 0.05. Where a significant main effect was found, i.e. the acceptability of the sentence as an answer to the question differed significantly across the stimuli in the group, five paired comparisons between stimuli pairs within the group were carried out (paired-samples t-test). Taking example group (1) in section 2 above as an example, the responses to stimulus (a) were compared with those for each of stimuli (b), (c) and (d). Also, responses to stimuli (b) and (c) (single-parameter manipulation) were compared with those to stimulus (d) (two-parameter manipulation). The significance level was set after Bonferroni correction to $\alpha = 0.01$.

The first three comparisons indicated whether or not one or both the single-parameter manipulations or whether only the two-parameter manipulation had an effect on the level of acceptance. The comparison of the one- and two-parameter manipulated stimuli was performed to indicate whether one parameter had a greater effect than the other in cases where both stimuli (b) and (c) differed significantly from (a).

In short, the comparisons offered the possibility to determine a hierarchy of parameter importance in terms of the contribution to a narrow focus accent or a de-accented element.

[1] Stimuli will be defined by their parameter values (A or D) in the order duration, intensity, *f0*. Thus the stimulus ADA defines the stimulus with high duration value, low intensity value and high *f0* value.

5. Results

One clear difference between the Bulgarian and German subjects was immediately apparent after the ANOVA stage of the analysis (cf. table 1). Whereas the German subjects revealed a significant main effect for *all the stimulus groups*, independent of which parameter was kept constant and which were manipulated, the Bulgarian subjects showed *no significant effects* for any of the stimulus-groups in the *f0-constant* conditions. In other words the low or high *f0* on "*kaka*" determined the relative acceptability of the answer to both LNF- and ENF-eliciting questions with no significant effect of any intensity or duration manipulation. Conversely, the significant effects found in the duration-constant and intensity-constant stimulus groups are exclusively due to the changes in acceptability induced by the *f0* manipulation (either lowering or raising).

Stimulus group	ENF		LNF	
	BG	D	BG	D
AAA-AAD-ADA-ADD	.001	.000	.003	.000
DAA-DAD-DDA-DDD	.000	.000	.001	.000
AAA-AAD-DAA-DAD	.000	.000	.007	.000
ADA-ADD-DDA-DDD	.000	.000	.000	.000
AAA-ADA-DAA-DDA	n.s.	.002	n.s.	.018
AAD-ADD-DAD-DDD	n.s.	.000	n.s.	.002

Table 1: *Main effect (p) across the stimuli in the group.*

	accented		de-accented	
	ENF	LNF	ENF	LNF
duration	AAA < AAD 19.4 63.0	AAA > AAD 76.1 39.4	DDD > DAA 83.1 22.5	DDD < DAA 47.8 86.2
	AAA = ADA 19.4 21.7	AAA = ADA 76.1 80.3	DDD = DAD 83.1 74.7	DDD = DAD 47.8 41.7
	AAA < ADD 19.4 72.8	AAA > ADD 76.136.8	DDD > DDA 83.1 24.9	DDD < DDA 47.8 85.7
	ADD = AAD 72.8 63.0	ADD = AAD 36.8 39.4	DAA < DAD 22.5 74.7	DAA > DAD 86.2 41.7
	ADD > ADA 72.8 21.7	ADD < ADA 36.8 80.3	DAA = DDA 22.5 24.9	DAA = DDA 86.2 85.7
intensity	AAA < AAD 21.0 68.3	AAA = AAD 80.2 43.0	DDD > ADA 79.5 22.5	DDD < ADA 36.3 83.0
	AAA = DAA 21.0 29.7	AAA = DAA 80.2 82.3	DDD = ADD 79.5 77.9	DDD = ADD 36.3 34.7
	AAA < DAD 21.0 67.3	AAA = DAD 80.2 50.4	DDD > DDA 79.5 29.5	DDD < DDA 36.3 84.3
	DAD = AAD 67.3 61.3	DAD = AAD 50.443.0	ADA < ADD 22.5 77.9	ADA > ADD 83.0 34.7
	DAD > DAA 67.3 29.7	DAD = DAA 50.4 82.3	ADA = DDA 22.5 29.5	ADA = DDA 83.0 84.3

Table 2a: *Perceptual patterns for Bulgarian ('>' and '<' represent a significant difference, p≤ .01, means for the compared pairs are given).*

This picture is borne out in detail by the paired comparisons. Table 2a shows that the expected *low* acceptance of the AAA stimulus in the ENF condition (target word de-accented) is significantly lower than AAD (with lowered *f0*) but does not differ from ADA (reduced intensity) or DAA (reduced duration). In the LNF condition (target word accented) the *high* acceptance of the AAA stimulus is significantly lowered by the *f0* manipulated AAD but is again not affected by the intensity manipulated ADA (but note that in the intensity-constant group *none* of the stimuli are judged significantly differently). The converse is true for the DDD stimuli. If *f0* has the D value, raising either or both the other parameters has no effect whereas a change in the *f0* value immediately changes the acceptability (for better or for worse depending on the question asked) in the duration-constant and intensity-constant stimulus groups.

	accented		de-accented	
	ENF	LNF	ENF	LNF
duration	AAA < AAD 18.1 72.0	AAA > AAD 92.2 26.0	DDD > DAA 90.9 18.2	DDD < DAA 30.7 90.1
	AAA = ADA 18.1 20.7	AAA > ADA 92.2 87.9	DDD > DAD 90.9 77.2	DDD = DAD 30.7 33.7
	AAA < ADD 18.1 86.0	AAA > ADD 92.2 29.0	DDD > DDA 90.9 22.0	DDD < DDA 30.7 84.5
	ADD > AAD 86.0 72.0	ADD = AAD 29.0 26.0	DAA < DAD 18.2 77.2	DAA > DAD 90.1 33.7
	ADD > ADA 86.0 20.7	ADD < ADA 29.0 87.9	DAA = DDA 18.2 22.0	DAA = DDA 90.1 84.5
intensity	AAA < AAD 18.6 77.8	AAA > AAD 86.7 34.0	DDD > ADA 88.8 17.7	DDD < ADA 29.7 88.2
	AAA = DAA 18.6 17.0	AAA = DAA 86.7 87.7	DDD > ADD 88.8 82.3	DDD = ADD 29.7 34.0
	AAA < DAD 18.6 86.5	AAA > DAD 86.7 27.1	DDD > DDA 88.8 23.5	DDD < DDA 29.7 88.6
	DAD = AAD 86.5 77.8	DAD < AAD 27.1 34.0	ADA < ADD 17.7 82.3	ADA > ADD 88.234.0
	DAD > DAA 86.5 17.0	DAD < DAA 27.1 87.7	ADA = DDA 17.7 23.5	ADA = DDA 88.2 88.6
f0	AAA < ADA 15.1 24.6	AAA = ADA 86.0 84.4	DDD > AAD 84.3 66.3	DDD < AAD 36.5 47.3
	AAA = DAA 15.1 17.3	AAA = DAA 86.0 90.5	DDD = ADD 84.3 80.3	DDD = ADD 36.5 30.9
	AAA < DDA 15.1 26.1	AAA = DDA 86.081.0	DDD > DAD 84.3 68.4	DDD = DAD 36.5 40.9
	DDA = ADA 26.1 24.6	DDA = ADA 81.0 84.4	AAD < ADD 66.3 80.3	AAD > ADD 47.3 30.9
	DDA = DAA 26.1 17.3	DDA < DAA 81.0 90.5	AAD = DAD 66.3 68.4	AAD = DAD 47.3 40.9

Table 2b: *Perceptual patterns for German (p≤ .01).*

Compared to this clear picture of *f0* dominance for BG subjects, German perceptual patterns are more complex, though consistent and clear. The *f0* effect is strong but it is less dominant. As table 2b (bottom section) shows: A constant *f0* value does not impede effects from the manipulation of the other parameters, in particular after an *ENF-eliciting* question. The ADA stimulus (reduced intensity) and the DDA stimulus (reduced duration and intensity) are significantly more acceptable than the maximally unacceptable AAA stimulus. Note, however, that DAA (reduced duration alone) does not differ significantly from either AAA or ADA. A similar pattern can be observed for the comparisons with the DDD stimulus group: DAD (increased intensity alone) is significantly less acceptable than the optimal DDD stimulus, as is the AAD stimulus (increased duration and intensity). ADD (manipulated duration alone – in this case increased) again has no significant effect on acceptability. This is apparent both from the non-significant comparison with DDD and the significantly lower acceptability of AAD compared to ADD. In summary, intensity but not duration manipulation can improve low *f0*-based acceptability scores and lower high *f0*-based scores when an ENF answer (target word de-accented) is expected.

In the LNF-eliciting context, the same effect is present but it is weaker. If *f0* has a high peak (A) on "*Wagen*" no reduction of duration or intensity can affect the high acceptability. The low acceptability of a low *f0* (D) in this context may, on the other hand, be significantly improved by increasing the intensity (AAD > DDD and AAD > ADD).

Despite the clear contribution of signal intensity to the acceptability or unacceptability of an answer to a given question, the stronger role of *f0* is apparent in the paired comparisons of intensity-constant and duration-constant stimuli (table 2b, top two sections). An A or D value for *f0* is always the primary determinant of basic acceptability (> 50%) or unacceptability (< 50%), independent of what is happening with the intensity or duration values. However, the *relative degree* of acceptability (or unacceptability) is influenced by intensity and duration.

In the duration-constant condition, intensity manipulation can significantly affect the acceptability of answers to ENF-eliciting questions, e.g. ADD > AAD (higher intensity makes a stimulus with acceptably low *f0* less acceptable); DDD > DAD (higher intensity again makes an acceptable stimulus less acceptable) and to LNF-eliciting questions, e.g. AAA > ADA (lower intensity makes an acceptable stimulus less acceptable). But note that DDD = DAD (higher intensity does not make the unacceptable stimulus less unacceptable).

In the intensity-constant condition, duration manipulation also significantly affects the acceptability of an answer in one case of an LNF-eliciting question: DAD < AAD (when *f0* is unfavourable with favourable intensity and duration values, reducing duration reduces acceptability).

6. Summary of results and discussion

There are significant differences between the two languages in the weighting of the prominence-lending properties in a prosody-based functionally oriented task. Bulgarian listeners base their judgments exclusively on *f0* whereas German listeners base their judgments on both *f0* and intensity, with *f0* clearly more important than intensity. However, a system-inherent or production-based explanation is not immediately obvious.

There are no direct links between production differences and perception differences. Analysis of the production patterns for ENF and LNF (de-)accentuation of target words in the kaka/Wagen positions [19] showed for BG and G:

(i) a comparable change of average *f0* on the stressed syllable from ENF to LNF (6.1 vs 5.5 semi-tones),

(ii) a greater change in intensity for BG than G (8.6 dB vs. 6.8 dB),

(iii) a much smaller % change in rhyme duration for BG than for G (12.7 vs. 18.6).

From these data, a theoretical scenario in which production patterns map onto perceptual sensitivity would predict roughly equal sensitivity to *f0* change, greater BG sensitivity to intensity change and greater G sensitivity to duration change. Both language groups' strong reaction to *f0* changes conform with such a prediction, but G's (secondary) sensitivity to intensity changes together with BG's apparent lack of sensitivity contradicts the prediction, as does G's lack of sensitivity to duration changes.

Differences in segmental phonological structure have no apparent bearing on the weighting of the properties either. Theoretically two predictive scenarios are feasible:

(i) Having no phonemic length contrast, Bulgarian, in contrast to German, might exploit duration more for accentuation, as is the case in Italian (lexically) and French (in affective emphasis phrasally),

(ii) Alternatively, the potential for lengthening phonemically long vowels in the accented position could predict the acceptance by the G subjects of greater duration in the focally accented /va:/ of "Wagen".

However, our perceptual data indicate that neither BG nor G weights duration as a contributing property in focal accentuation. This may seem to contradict Cumming's [18] finding of German perceptual sensitivity to syllable duration in a prosodic function, namely in a syntactic grouping task, However, there is no reason to expect focal prominence to use the same properties as syntactic grouping, a function that is strongly dependent on 'final lengthening'.

7. Acknowledgements

This research was funded by the German Research Council, grant (BA 737/10-1).

8. References

[1] Koreman, J., *Decoding Linguistic Information in the Glottal Airflow*. Dissertation, Nijmegen University, 1996.

[2] Marasek, K., *Electroglottographic Description of Voice Quality*. (= AIMS 3,2, Lehrstuhl für experimentelle Phonetik), Universität Stutgart, 1997.

[3] Fry, D.B., "Experiments in the perception of stress." *Lang. & Speech* 1, 126-152, 1958.

[4] Fry, D.B., "The dependence of stress judgments on vowel formant structure." *Proc. 6th ICPhS* Münster, 306-311, 1965.

[5] Kuijk, D. van & Boves, L (1999). "Acoustic characteristics of lexical stress in continuous telephone speech", *Speech Communication* 27, 95-112.

[6] Kochanski, G., Grabe, E., Coleman, J. & Rosner, B, "Loudness predicts prominence: Fundamental frequency lends little", *J. Acoustical Soc. America* 118, 1038-1054, 2005.

[7] Sluijter, A. M. C. & van Heuven, V.J., "Spectreal balance as an acoustic correlate of linguistic stress", *J. Acoustical Soc. America* 100, 503-513, 1996.

[8] Sluijter, A. M. C., van Heuven, V.J. & Pacilly, J. J. A., "Spectreal balance as a cue in the perception of linguistic stress", *J. Acoustical Soc. America* 101, 2471-2485, 1997.

[9] Stevens, S. S., "Perceived level of noise by Mark VII and decibels", *J. Acoustical Soc. America* 51, 575-602, 1971.

[10] Streefkerk, B M. et al., "Acoustical features as predictors for prominence in read aloud Dutch sentences used in ANN's.", *Proc. Eurospeech '99*, Budapest, 551-554, 1999.

[11] Fant, G. & Kruckenberg, A., "Prominence and accentuation. Acoustical correlates", P. Branderund & H. Traunmüller (eds.) *Proc. of FONETIK 98*, Dept. of Linguistics, Stockholm University, 142-145, 1998.

[12] Fant, G. & Kruckenberg, A., "Prominence correlates in Swedish prosody", *Proc. ICPhS XIV*, San Francisco, 1749-1752, 1999.

[13] Eriksson, A., Thunberg, G.C. & Traunmüller, H., "Syllable prominence: a matter of vocal effort, phonetic distinctness and top-down processing", *Proc. Eurospeech 2001*, 388-402, 2001.

[14] Eriksson, A., Grabe, E. & Traunmüller, H., "Perception of syllable prominenceby listeners with and without competence in the tested language", *Proc. Speech Prosody* 2002, 275-278, 2002.

[15] Koreman, J., Andreeva B. & Barry, W., Accentuation cues in French and German, In: P. A. Barbosa, S. Madureira, & C. Reis (eds.), *Proc. Speech Prosody* 2008, Campinas: Editora RG/CNPq, 613-616, 2008.

[16] Koreman J., Andreeva, B., Barry, W., Dommelen, W.v., Sikveland. R.-O., Cross-language differences in the production of phrasal prominence in Norwegian and German, Martti Vainio, Reijo Aulanko, and Olli Aaltonen (eds.), *Nordic Prosody, Proceedings of the Xth Conference*, Helsinki 2008, Frankfurt: Peter Lang, 139-150, 2009

[17] Barry, W. & Andreeva, B. Is it important for communication which parameters signal accentuation? *Proc. ICPhS XVII*, 2011, Hong Kong, 288-291, 2011.

[18] Cumming, Ruth E., *Speech Rhythm. The language specific integration of pitch and duration*. PhD. Dissertation, University of Cambridge, UK, 2010.

[19] Andreeva, B. & Barry, W., Fine phonetic detail in prosody. Cross-language differences need not inhibit communication. In: O. Niebuhr (ed.), *Prosodies - context, function, and communication*. Berlin/New York: de Gruyter. 259-288, (forthcoming).