
Acoustic-phonetic realisation of polish syllable prominence: a corpus study of spontaneous speech

Akustyczno-fonetyczna realizacja wzmocnienia polskiej sylaby: korpusowe badanie mowy spontanicznej

Zofia Malisz, Petra Wagner

Universität Bielefeld, Germany
{zofia.malisz,petra.wagner}@uni-bielefeld.de

ABSTRACT

Polish presents an interesting case for testing alternative phonetic implementations of prominence: It has fixed lexical stress on the penultimate, it has been difficult to classify within the classic ‘stress-timing’ vs. ‘syllable-timing’ dichotomy [1, 2, *inter alia*] and its stress is regarded as ‘weakly expressed’ [3]. We investigate acoustic correlates of Polish prominence patterns in a corpus of spontaneous, task-oriented dialogue. Results indicate clear differences to prior analyses of more controlled data, with intensity but also duration and pitch movement being main indicators of prominence.

STRESZCZENIE

Polski przedstawia ciekawą możliwość przetestowania odmiennych struktur fonetycznej implementacji prominencji. Polski posiada stały akcent wyrazowy na przedostatniej sylabie, trudno zaklasyfikować go w ramach typologii rytmu sylabicznego lub opartego na akcencie, a akcent jest uznawany za słaby akustycznie. W niniejszej pracy rozpatrujemy akustyczne korelaty prominencji w języku polskim na podstawie korpusu dialogów zadaniowych. Głównymi przebieżnikami prominencji okazały się parametry oparte na intensywności, ale także iloczasię oraz częstotliwości podstawowej.

1. Introduction

Prosodic prominence is commonly regarded as the perceptual salience of a linguistic unit within its linguistic context. However, we are far from having a consensus on how it is measured subjectively and how it relates to objectively measurable acoustic events or linguistic prosodic structures such as lexical and sentence stress or prosodic focus – not even in widely studied Germanic languages such as English, German, Swedish or Dutch. At least, there is wide agreement that prominence perception is influenced by both top-down expectancies (mostly shaped by linguistic structures but also ‘paralinguistics’) and bottom-up processing, i.e. the interpretation of the acoustic signal [4–8]. Naturally, pure reliance on top-down expectancies in speech processing is implausible, since it would prevent listeners from detecting errors in stress placement or deviations from the ‘citation form’, e.g. occurring in non-native speech or due to stylistic diversity.

The acoustic-phonetic correlates of prominence are manifold and most likely language specific. Among the correlates suggested are pitch excursion and shape, e.g. in German, higher and late peaks tend to correlate with a higher degree of perceived prominence compared to early, flat ones, e.g. [9]. Despite the fact that pitch accents have a clear impact on perceptual prominence in several Germanic languages, its influence has sometimes been hard to detect in corpus studies, where its impact may be overruled by vocalic duration and intensity related measures, e.g. [10–12]. Vowel and syllable duration have successively been shown to be the most robust cue to perceptual prominence in many Germanic languages, while overall intensity seems to play an unclear, or even negligible, role e.g. in acquisition studies of German stress patterns [13]. Another acoustic measure that has been shown to correlate with prominence is an increase of ‘vocal effort’, resulting in a faster closing of the vocal folds thereby boosting the higher frequency component of the source spectrum [14–17]. This metric has become known as ‘spectral tilt’ and has been shown to correlate with perceptual prominence in several (Germanic) languages [12, 15, 16, 18, 19]. Since spectral tilt results from differences in the observed glottal shape parameters, it has also been used as an acoustic correlate of voice quality (modal vs. breathy and tense). It is a matter of perspective whether measures of spectral tilt and their effect on stress reflect greater articulatory effort or voice quality changes that contribute to the perception of stress/prominence. However as Prieto [17] notes, it is also difficult to disentangle the effect of stress on spectral tilt from other effects, such as vowel quality: the changes between higher and lower formant amplitudes are also related to changes in vowel formant frequencies. Possibly for this reason, spectral tilt has proven to be hard to capture as a correlate of stress cross-linguistically. Differences in vowel quality in reduced positions covary with stress in languages such as English and Dutch [15, 18] and with vowel tenseness in German [20].

Very few authors pointed out the influence of syllable initial consonantal strengthening in signalling prominence [9, 21].

Often, investigations of prosodic prominence are difficult to compare, given various scales used (binary, multilevel, continuous), the different linguistic units (syllables, words), the different annotators (expert vs. naïve), different languages or speaking styles (read, spontaneous). In a recent evaluation of annotation formats, Arnold, Möbius and Wagner [22, 23] pointed out the advantages of multilevel scales over binary ones, since these are more informative than binary scales where inter-annotator agreement is often misinterpreted as a high degree of prominence. Continuous scales seem to be less demanding than multilevel scales, while naïve listeners show a lot more inter-annotator consistency when judging word prominence rather than syllable prominence.

We therefore see a multitude of options and approaches and issues to keep in mind when aiming to unfold facts about a language that has been so far understudied in terms of its acoustic and linguistic correlates of prominence.

In this paper, prominence has been operationalised using a 3-level scale and the annotations were carried out on syllables. The annotators were expert listeners, and annotations were based on spontaneous, task-oriented dialogues.

2. The Phonetic Implementation of Polish Prominence

A number of studies on Polish lexical stress are available. Traditionally, Polish lexical stress has been described as ‘dynamic’, that is, acoustically primarily correlated with overall intensity. It was assumed that stressed syllables are articulated with greater vocal effort and perceived as louder and it is the relative differences in loudness that define Polish stressed and unstressed syllables rather than pitch movements or duration. Consequently, early observational studies such as Dłuska [24] claimed that a slight rise in loudness is the primary correlate of Polish stress.

Jassem [25], however, in 1962, has shown on the basis of acoustic measurements that it is in fact pitch movement that is ranked as the most salient correlate of lexical stress in Polish. His study involved spontaneous and read material, including isolated words and sentences.

Dogil [3] collected recordings of three speakers who replied to questions designed to elicit broad, narrow and no focus on a target word in a sentence. His results showed that in the position of no focus, primary stress in the target word is characterised by the highest F0 with a sharp pitch slope. The results appear to confirm Jassem’s findings. Under broad focus however, as [3] propose, “a position for the association with the nuclear pitch-accent morpheme of a sentence” is only “pointed to” by lexical stress. This means that lexical stress in Polish is best represented by a model where it is context dependent, ‘potential’ and strongly interacts with the intonational structure of a sentence, such as the one suggested by [26].

Notably, in all the above studies, duration has no or only weak influence on stressed vowels, contrary to e.g. most Germanic languages. Jassem [25] estimated the duration ratio of stressed to unstressed vowels at 1.17. Nowak [27] in a large corpus study on vowel reduction in Polish found a similar relationship of 1.22. Klessa’s [28] analysis of a corpus of spontaneous and read speech built for speech synthesis purposes also quotes values that amount to a ratio of approx. 1.2. However, when vowels in prepausal syllables were excluded, the ratio in Klessa’s work equals 1.1, while for English, this value equals 2 for monophthongs [29].

Secondary stress has received some attention and is impressionistically agreed to exist. In words longer than three syllables, secondary stress falls on the first syllable. Acoustically, [3] showed that relatively longer duration and a fully articulated vowel characterise syllables receiving secondary stress. However, a perceptual study by Steffen-Batogowa [30] has found no systematic evidence of secondary stress. The acoustic status of secondary stress in Polish has also been questioned recently [31]. A common process occurs, as described by Dogil [32], where under narrow focus, primary stress shifts from the canonical penult onto the first syllable, i.e.: “in Polish a single word, when under focus, switches the prominence values of primary and secondary stress” [3:p286].

The covariance of spectral tilt with stress related vowel quality or tenseness in Germanic languages, mentioned in Section 1, should be methodologically less problematic for Polish. Crosswhite [33] showed that an acoustic measure linked to spectral tilt (the difference between the perceived loudness in phons and sound intensity level in dB), was significantly affected by stress in Polish, Macedonian and Bulgarian. We are not aware of any other studies related to the effect of the slope of the spectrum on stress in Polish.

Given the above evidence, it can be hypothesised that prominent syllables in Polish should exhibit differences across pitch and spectral tilt parameters, possibly overall intensity measures, and only weakly depend on duration. Overall prominence patterns in running non-read speech have not been systematically studied so far to our knowledge. The fullest account of Polish prosodic structure using continuous read speech material is given by Demenko [34] where different constituencies that contribute to the perception of prominence were considered. Our study gives an account of acoustic correlates of prominence patterns in an annotated corpus of spontaneous task-oriented dialogues.

3. Method

Building on the hypotheses stated above, we investigated acoustic correlates of Polish prominence using the acoustic measures stated in the previous section and compared them with annotations by native expert listeners.

3.1. Corpus data

A subset of DiaGest2, a Polish multimodal corpus of task-oriented dialogues [35], was used. The data in the present study come from 4 speakers of standard Polish (2 female and 2 male undergraduates) whose task was to instruct a dialogue partner in a paper folding task.

3.1.1. Syllabic segmentation

Syllabic boundaries were marked according to sonority principles. The Maximal Onset Principle was not used, resulting in the closing of syllables in case of medial clusters such as in: ‘miasto’ → mias.to, ‘mokry’ → mok.ry. There are several problems with segmenting Polish syllables. Examples include cases such as proclitics plus nouns: ‘z okna’ (out of the window), ‘w wodzie’ (in the water). In the present corpus they were segmented as ‘zok.na’ and ‘wwo.dzie’. Symmetrical cases such as ‘oko’ (eye) were labeled as o.ko. In general broad phonetic transcription principles were applied in the labeling of syllables, Polish SAMPA conventions were used.

3.1.2. Prominence and phrasal annotation

Rhythmic and phrasal structure was annotated using the Rhythm and Pitch (RaP) system [36], largely theory-independent and based on perceptual native judgements. Minor phrasal boundaries were delimited and rhythmic prominences were marked in the dataset by two expert phoneticians, native speakers of Polish.

The RaP minor phrase boundary is defined as a minimally perceptible disjuncture. It approximately corresponds to the ToBI break index ‘3’ [36]. The major phrase boundary marks clearly perceptible disjunctures and often coincides with major intonational phrase boundaries (ToBI break index ‘4’). However, the phrasal structure in this context follows metrical rather than intonational constituency and more strictly corresponds to breath groups. Tonal labels were not marked.

The experts identified all rhythmically stressed syllables on two prominence levels: minor and major prominence. A *major prominence* marked a syllable with a perceptually most salient prominence within a phrase. It has been observed that major prominences

coincided with phrasal pitch accents, however, since pitch accents were not marked, no systematic comparison was possible. *Minor prominences* were defined as syllables perceived as prominent but relatively less than major prominences. Consequently, the annotation procedure involved the identification of a major prominence within a RaP phrase first and then scanning the phrase for relatively weaker beats. This means the labeling was not done strictly sequentially but recursively, reflecting the relative structure of prominences in a studied phrase. Non-prominent syllables were not labeled.

3.2. Acoustic feature extraction

Pitch, RMS intensity and spectral slope parameters were extracted from the syllables in utterances judged by experts and naive non-native annotators. All prosodic features were extracted using scripts in Praat. The magnitude of pitch and intensity excursions was estimated by subtracting the mean phrase values from the mean value of a given syllable in that phrase and dividing by the standard deviation value of that phrase (z-score normalisation).

All prepausal syllables were excluded based on the RaP phrasal structure labels described in section 3.1.2. This was done because of the universal pre-boundary lengthening processes also evidenced in Polish [35] that might confound duration effects in our study. Moreover, substantial pitch movements signaling phrase endings are also observed in Polish, steep rises in these positions are characteristic in colloquial speech and final falls in more formal varieties. Any hesitations, filled pauses etc were also excluded.

Spectral tilt parameters consisted of H1A1, H1A2 and H1A3. H1H2 was found to highly correlate with H1A3 and was excluded from further analysis. To determine the slope of the spectrum the difference between the first harmonic (H1) and the maximum harmonic in three formant peaks (A1-A3) is calculated. It has to be noted that our measures capture the spectral tilt over the whole syllable including the consonants and therefore approximately capture glottal effort of a syllable rather than voice qualities.

4. Results

In order to assess the dependence of prominence judgements on prosodic features we first calculated correlations between each prosodic measure and the prominence values labeled by native annotators. The results are presented in Table 1 for a) all prominence labels ordered from no prominence to minor prominence and major prominence and for b) only minor and major prominence labels.

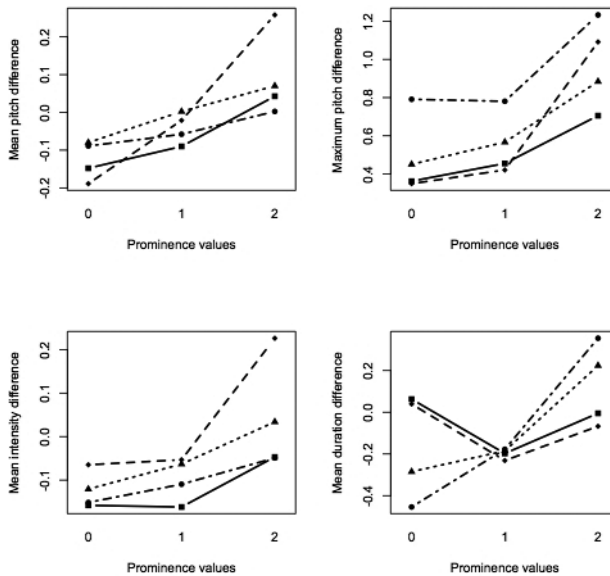
For all prominence levels in this spontaneous material, the correlation strengths are low but statistically significant for all features. Maximum pitch difference, mean intensity difference and syllable duration positively correlate with increasing levels of prominence. Spectral slope measures correlate negatively: the more prominent the syllable, the more ‘effortful’ it is and the flatter the slope is. Spectral slope does not correlate significantly with the minor vs. major prominence level progression.

Most notably, no acoustic feature significantly correlated with the ‘no prominence’ to ‘minor prominence’ judgements.

Table 1: Spearman rank correlation coefficients between prominence values and prosodic features. Values in italics were all significant at $p < .001$.

	Prominence level 0 to 1 and 2	Prominence level 1 to 2	All
Mean pitch difference	0.10	0.07	0.10
Max pitch difference	0.13	0.17	0.15
Mean intensity difference	0.10	0.13	0.12
Syllable duration	0.10	0.18	0.12
H1A1	-0.03	-0.08	-0.07
H1A2	-0.11	-0.05	-0.12
H1A3	-0.09	-0.02	-0.10

To further illustrate the data, Figure 1 presents the mean values of prosodic parameters that represent the most widely claimed determinants of prominence: pitch measures, mean intensity and syllable duration across prominence levels.

**Figure 1: Mean values of four main prosodic features known to correlate with prominence, by subject.**

Three Generalised Linear Mixed Models (GLMM) with a logit link-function were formulated for the following binomial dependent variables: a) the ‘no prominence’ ($N = 742$) vs. ‘minor prominence’ ($N = 300$) classes, b) ‘no prominence’ vs. ‘minor prominence’ and ‘major prominence’ ($N = 587$) and c) ‘minor prominence’ vs. ‘major prominence’ ($N = 287$). The fixed factors included *Mean Intensity Difference*, *Maximum Pitch Difference*, *Syllable Duration* and the three spectral slope measures. Subject identifiers and syllable SAMPA labels were entered as random factors to account for individual variances of the subjects and the phonetic structure of any given syllable in the corpus.

No significant differences were found for model a) comparing unmarked syllables with syllables bearing ‘minor prominence’ labels. Table 2 shows the estimates of model. b) after refitting, comparing non-prominent syllables to all prominent ones. Table 2 presents the coefficients of model c), after refitting, where the two prominence marked levels are compared to each other.

Table 2: Binomial GLMM estimates for ‘no prominence’ vs. ‘minor prominence’ and ‘major prominence’.

Fixed effects:	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	-0.542	0.16	-3.308	0.0009 ***
Max pitch difference	0.174	0.08	2.135	0.03 *
Mean intensity difference	0.589	0.196	2.996	0.003 **
Random effects:				
Resid. Variance	Var.		Std. Dev.	
	5.3 (syllable label)		2.3	
Model predicted probabilities:				
	C-index	Dxy		
	0.93	0.86		

5. Discussion

We suggest that acoustic correlates of prominence in Polish manifest themselves largely in phrase accentuation structure. Overall intensity, duration and pitch movement are good correlates of phrase accent. In this study, syllables labeled as ‘major prominence’ represent realised phrasal accents and were shown to exhibit a much larger and significant difference in these acoustic dimensions, similarly to results in [34].

A model containing the no-marking class and weakly prominent class did not return any significant results for all studied acoustic dimensions. Syllables labeled as ‘minor prominence’ in the present corpus did not carry pitch accent and largely coincided with lexically stressed syllables, whether primary or secondary stress, under no focus. Our results therefore support the previously reported weak effect of lexical stress on acoustic features in Polish. Spectral tilt has been hypothesised to express lexical stress better than overall intensity ([37] for Dutch). We have not found evidence for that, although a statistical tendency existed in the analysis (not reported) and will be investigated further with more data and refined measurements of spectral slope e.g. over the vocalic portion of the syllable.

Our results support the notion that the fixed penultimate stress pattern in Polish is determined linguistically as a highly influential ‘expectation’ that is perceived by native listeners but not attested acoustically in a clear way.

Results such as [38] confirm in an Event-Related Potential (ERP) study that stimuli with phonologically deviant stress patterns evoke strong positive potentials in Polish subjects as opposed to any stimuli with the default stress pattern.

Differences across all levels, best representing the general correlates of prominence in Polish, indicate intensity difference and maximum pitch difference as the main determinants of overall prominence. Weak stress effects on duration seem to

Table 3: Binomial GLMM estimates for ‘minor prominence’ vs. ‘major prominence’.

Fixed effects:	Estimate	Std.Error	z value	Pr(> z)
(Intercept)	0.05954	0.17043	0.349	0.726801
Syllable duration	0.468	0.13	3.668	0.0002 ***
Max pitch difference	0.319	0.10	3.086	0.002 **
Mean intensity difference	0.678	0.25	2.742	0.006 **
Random effects:				
Resid. Variance	Var.		Std. Dev.	
	1.02 (syllable label)		1.00	
	0.025 (subject)		0.16	
Model predicted probabilities:		C-index	Dxy	
		0.91	0.83	

cancel out the apparent accentuation effects in this dimension. Given the high values of predicted probability measures, it can be expected that an automatic classification of prominence in Polish based on these features is possible.

6. Conclusion

Given the present material of spontaneous task-oriented dialogues we were able to show that pitch, intensity and duration variability covaries with perceptual prominence in Polish. Future directions involve incorporating more data, other prosodic parameters and measures.

A subsequent extension will compare the present findings to those based on other speaking styles as well as annotations by non-native and non-expert listeners using a continuous prominence annotation scale.

REFERENCES

- [1] Ramus, F., E. Dupoux, and J. Mehler. 2003. The psychological reality of rhythm classes: Perceptual studies. In *Proceedings of the 15th ICPhS, Barcelona, Spain*.
- [2] Malisz, Z. 2011. Tempo differentiated analysis of timing in Polish. *Proceedings of the 17th ICPhS, Hong Kong, China*. 1322–1225.
- [3] Dogil, G. 1999. The phonetic manifestation of word stress in Lithuanian, Polish and German and Spanish. In H. van der Hulst, ed, *Word prosodic systems in the languages of Europe*. Berlin: Mouton de Gruyter.
- [4] Fant, G. and A. Kruckenberg. 1989. Preliminaries to the study of Swedish prose reading and reading style. *STL-QPSR* 30 (2), 1–80.
- [5] Streefkerk, B. 2002. *Prominence – Acoustical and lexical/syntactic correlates*. Utrecht: LOT.
- [6] Eriksson, A., G. Thunberg and H. Traunmüller. 2001. Syllable prominence: A matter of vocal effort, phonetic distinctness and top-down processing. In *Proceedings of Eurospeech 2001*. 399–402.
- [7] Eriksson, A., E. Grabe and H. Traunmüller. 2002. Perception of syllable prominence by listeners with and without competence in the tested language. In *Proceedings Speech Prosody 2002, Aix-en-Provence*. 275–278.

- [8] Wagner, P. 2005. Great Expectations – Introspective vs. Perceptual prominence ratings and their acoustic correlates. In *Proceedings of Interspeech 2005, Lisbon*. 2381–2384.
- [9] Kohler, K. and O. Niebuhr. 2007. The phonetics of emphasis. *Proceedings of the 16th ICPhS, Saarbrücken, Germany*. 2145–2148.
- [10] Silipo, R. and S. Greenberg, S. 1999. Automatic Transcription of Prosodic Stress for Spontaneous English Discourse. In *Proceedings of ICPhS 1999, San Francisco*. 2351–2354.
- [11] Batliner, A., S. Steidl, B. Schuller, D. Seppi, T. Vogt, L. Devillers, L. Vidrascu, N. Amir, L. Kessous and V. Aharonson. 2007. The Impact of F0 extraction errors on the classification of prominence and emotion. In *Proceedings of 17th ICPhS, Saarbrücken, Germany*. 2201–2204.
- [12] Kochanski, G., E. Grabe, J. Coleman and B. Rosner. 2005. Loudness predicts prominence: fundamental frequency lends little. *Journal of the Acoustical Society of America* 118 (2), 1038–1054.
- [13] Schneider, K. and B. Möbius. 2007. Word stress correlates in spontaneous child-directed speech in German. In *Proceedings of Interspeech 2007, Antwerp, Belgium*. 1394–1397.
- [14] Sluijter A. 1995. *Phonetic correlates of stress and accent*. Leiden: Holland Institute of Generative Linguistics.
- [15] Sluijter, A. and V. van Heuven. 1996. Acoustic correlates of linguistic stress and accent in Dutch and American English. In *Spoken Language 1996. ICSLP 1996*. Vol. 2, 630–633.
- [16] Heldner, M. (2001). Spectral emphasis as an additional source of information in accent detection. In M. Bacchiani, J. Hirschberg, D. Litman and M. Ostendorf, eds., *Prosody 2001: ISCA Tutorial and Research Workshop on Prosody in Speech Recognition and Understanding*. Red Bank, NJ, USA: ISCA. 57–60.
- [17] Ortega-Llebaria, M. and P. Prieto. 2010. Acoustic correlates of stress in Central Catalan and Castilian Spanish. *Language and Speech* 54 (1), 73–97.
- [18] Campbell, N. and M. Beckman. 1997. Stress, prominence, and spectral tilt. In *INT-1997, Athens, Greece*. 67–70.
- [19] Portele, T., B. Heuft, C. Widera, P. Wagner. 2000. Perceptual Prominence. In W. Sendlmeier, ed, *Speech and Signals. Aspects of Speech Synthesis and Automatic Speech Recognition. Festschrift dedicated to Wolfgang Hess on his 60th birthday. Forum Phoneticum* 69, 97–116.
- [20] Marasek K. 1996. Glottal correlates of the word stress and the tense/lax opposition in the German vowels, In *Proceedings of ICSLP 1996*, 1573–1577.
- [21] Niebuhr, O. and A. Wolf. 2011. Low and high, short and long by crook or by hook? In *Proceedings of Interspeech 2011*. 1869–1872.
- [22] Arnold, D., P. Wagner and B. Möbius. 2011. Evaluating different rating scales for obtaining judgments of syllable prominence from naive listeners. In *Proceedings of 17th ICPhS 2011, Hong Kong*.
- [23] Arnold, D., B. Möbius and P. Wagner. 2011. Comparing word and syllable prominence rated by naive listeners. In *Proceedings of Interspeech 2011, Florence, Italy*.
- [24] Dłuska, M. 1950. *Fonetyka polska*. PWN: Warszawa.
- [25] Jassem, W. 1962. *Akcent języka polskiego*. Wrocław: Ossolineum.
- [26] Abercrombie, D. 1991. *Fifty years in phonetics*. Edinburgh: Edinburgh University Press.
- [27] Nowak, P. M. 2006. *Vowel reduction in Polish. Ph.D. dissertation*, University of California, Berkeley.
- [28] Klessa, K. 2006. *Analiza iloczasu głoskowego na potrzeby syntezy mowy polskiej*. Doctoral Dissertation. Poznań: Uniwersytet im. Adama Mickiewicza, Poznań.
- [29] Crystal, T. and A. House. 1998. Segmental durations in connected-speech signals: Syllabic stress. *Journal of the Acoustical Society of America* 83, 1574–1585.

- [30] Steffen-Batogowa, M. 2000. *Struktura akcentowa języka polskiego*. PWN: Warszawa.
- [31] Newlin-Lukowicz, L. 2012. Polish Stress: A phonetic investigation of phonological claims. In *The 36th Penn Linguistics Colloquium*, Philadelphia.
- [32] Dogil, G. 1980. Focus marking in Polish. *Linguistic Analysis* 6, 221–245.
- [33] Crosswhite, K. 2003. Spectral tilt as a cue to stress in Polish, Macedonian and Bulgarian. *Proceedings of the 15th International Conference of Phonetic Sciences, Barcelona*. 767–770.
- [34] Demenko, G. 1999. *Analiza cech suprasegmentalnych języka polskiego na potrzeby technologii mowy*. Poznań: Wydawnictwo Naukowe UAM.
- [35] Karpiński, M., E. Jarołowicz-Nowikow, Z. Malisz, M. Szczyszek and J. Juszczak, 2008. Rejestracja, transkrypcja i tagowanie mowy oraz gestów w narracji dzieci i dorosłych. (Recording, transcription and tagging of speech and gestures in child and adult narration). Poznań, *Investigationes Linguisticae XVI*.
- [36] Breen, M., L. C. Dilley, J. Kraemer and E. Gibson. In press. Inter-transcriber reliability for two systems of prosodic annotation: ToBI (Tones and Break Indices) and RaP (Rhythm and Pitch). *Corpus Linguistics and Linguistic Theory*.
- [37] Sluijter, A., V. van Heuven and J. A. Pacilly. 1997. Spectral tilt as a cue in the perception of lexical stress. *Journal of the Acoustical Society of America* 101 (1), 503–513.
- [38] Orzechowska, P., U. Domahs, J. Knaus and R. Wiese. 2012. Processing (un-)predictable word stress in Polish: an ERP study. Presented at the 9th Old World Conference in Phonology 2012, Berlin.