Description of Polish speech rhythm using rhythm metrics and the time-delay approach: A comparative study

Agnieszka Wagner

Institute of Linguistics, Adam Mickiewicz University in Poznan, Poland

`wagner@amu.edu.pl`

Abstract

The goal of this study is to provide a multidimensional description of rhythmic structure of Polish utterances. For this purpose a time-delay approach proposed in [1] is applied and results of qualitative and quantitative analyses based on time-delay plots are compared with results obtained with selected rhythm metrics. The study shows that description that relies on a combination of rhythmic scores is inconclusive and difficult to interpret, because it does not account for rhythmic structuring nor grouping. The time-delay approach, on the contrary, appears to be very efficient in exploring short-time and long-term timing variability that determine Polish speech rhythm.

Index Terms: Polish speech rhythm, time-delay approach, meter, grouping, prominence, rhythm metrics

1. Introduction

1.1. Analysis and description of speech rhythm

For many years, research on speech rhythm was based on *isochrony* paradigm [2, 3] and concept of *rhythmic classification of languages* (i.e. the distinction between syllable-timed, stress-timed and mora-timed languages). The fact that instrumental studies failed to bring evidence for stress- and syllable-based isochrony caused that in rhythm research the focus moved from duration measurements of syllables and feet to investigation of phonemic and phonological factors which affect the timing of syllables and feet. Observed differences in vowel reduction, stress-based lengthening and syllable complexity between stress- and syllable-timed languages motivated the development of *rhythm metrics* – formulas that measure durational variability in consonantal (C) and vocalic (V) intervals. The most widely used metrics include: %V-ΔC [4], PVIs [5] and Varcos [6]. They have been applied in studies on rhythmic classification of languages [7], acquisition of timing patterns in L1 [8] and L2 [9, 10, 11, 12, 13, 14], detection of speech impairments [15] or dialect discrimination [16]. The proponents of rhythm metrics provided experimental results showing that they can be regarded as acoustic correlates of perceived speech rhythm and supporting the hypothesis of rhythmic classification of languages. However, rhythm metrics have been the subject of considerable criticism. First of all, it can be argued that what metrics describe is not rhythm, but timing of utterances [17]. Secondly, metric scores vary considerably within languages, because they are sensitive to a number of factors such as tempo, speech style or method of measurement [18, 19]. Thirdly, rhythmic classification based on the stress-timing vs. syllable-timing dichotomy applies to *some languages in some conditions* (for example, different metrics provide different classifications of the same language [18]) and there are languages, for example Polish, that can not be assigned to any class on the basis of the metric scores and are consequently labelled as "intermediate". At the same time, metrics are incapable of giving information on distinctive features of such intermediate rhythm. In the end, metrics such as PVIs or %V-ΔC describe single aspect of rhythm, i.e. variability in time domain, but "rhythm cannot be described as a one-dimensional property of speech, e.g. as more or less variable or more or less stress timed" ([1], p.145). Rhythm can be regarded as a perceptual impression of a *structure* consisting of more or less prominent (strong, weak) events (beats or syllables) which are *grouped* in a particular manner to form *perceptually distinct patterns* such as iambs or trochees. It is doubtful whether processes underlying rhythmic structuring and grouping can be explained by means of rhythm metrics, but there is evidence from [1] that they can be successfully explored using a multidimensional approach which takes into account short-term and long-term timing variability including, among others, relative timing of functionally distinct transitions (to stressed, to unstressed and to phrase final syllables), acceleration and deceleration tendencies, compensatory shortening and time shrinking (i.e. a psychoacoustic phenomenon which results in perception of decelerating sequences as isochronous). A multidimensional account of rhythm incorporates various levels of the prosodic hierarchy and dimensions other than duration/timing, because intensity, F0 and spectral features also constitute important correlates of prominence. The approach taken in this study makes it possible to analyze and describe speech rhythm in such a multidimensional manner. It is based on *time-delay plots* which provide "a useful tool in order to explore timing relations that are perceived as typical rhythms in speech" and which are "directly interpretable along similar rhythm-related dimensions as have been detected in typological analyses" ([1], p. 155). The time-delay approach seems to be particularly useful to investigate fine-grained timing differences related to rhythmic structure and grouping in non-prototypical – rhythmically "intermediate" or unclassified – languages, such as Polish.

1.2. Polish speech rhythm

As regards phonological properties, Polish can be regarded as rhythmically "mixed" – it has fixed lexical stress, no vocalic reduction in unstressed syllables and subtle stress-related and accentual lengthening, which are considered typical features of syllable-timed rhythm, whereas high phonotactic complexity and presence of compensatory shortening points to stress-timing. Former studies provided evidence for accentual lengthening of vowels [20, 21, 22] and some support for isochrony within narrow rhythm units [23]. The results of recent corpus-based analyses [24, 25] showed that accentual lengthening is limited to vowels and syllables associated with major prominences (phrase accents), whereas durational marking of minor prominences (which coincide with word stress) is very subtle, if any. In [26] overall intensity was regarded as the main acoustic correlate of stress and in [27] – pitch movements. In [25] intensity and F0 features correlated significantly only with major prominences. Prosodic phrase

boundaries in Polish are signaled most of all by increased duration of the phrase-final and the penultimate syllable and vowel (associated most of the time with phrase/nuclear accent), but F0 features also play a significant role [22, 29]. The results of a multidimensional analysis of Polish rhythm in [1] showed only subtle lengthening of stressed syllables, tendency for foot final lengthening, deceleration throughout the foot, some compensatory shortening effects (comparing binary and longer feet) and potential evidence for time shrinking phenomenon that can contribute to impression of isochrony. These results clearly show that Polish can not be easily classified based on the stress-timed/syllable-timed dichotomy. The accounts of Polish rhythm based on rhythm metrics are inconclusive: According to PVIs [5], Polish is close to syllable-timed languages, but, according to $\%V - \Delta C$, it is grouped with stress-timed English and Dutch [4].

1.3. Objectives of the study

The objective of the study is to compare two approaches to analysis of speech rhythm in Polish: rhythm metrics [4, 5, 6] and the time-delay approach [1], and to provide explanation of factors underlying the "mixed rhythm effect" in Polish, because existing descriptions are neither informative nor satisfactory in this respect. For this purpose the study explores short-term timing variability in the realization of functionally different prosodic transitions: to stressed, to unstressed, to phrase final syllables, and long-term timing characteristics of rhythmical grouping at various levels of prosodic hierarchy – feet (of different sizes) and prosodic phrases. Rhythmic characteristics of Polish are also compared cross-linguistically by referring to results reported in the literature.

2. Methodology

2.1. Speech data

The speech material includes recordings of a literary fairy tale "The teapot" (by H. Ch. Andersen), read by five speakers – all coming from Poznan and presenting the Poznan-Cracow pronunciation. The text consists of 19 phonetically and prosodically rich sentences (491 syllables). Recordings were carried out in a sound-treated booth, directly to a disk and with a sampling frequency of 16 kHz. The subjects were asked to read the text once (sentence after sentence), at their own pace. Sentences containing disfluencies or mispronunciations were re-recorded. The recorded material constitutes part of a speech database created for the purpose of studying speech rhythm in native and non-native Polish [30]. The whole speech material was subject to automatic phonetic transcription and alignment [31] which were verified and manually corrected following standard segmentation criteria. Syllable boundaries were determined as in [32]. Prosodic annotation consisted in labeling four levels of prominence (unstressed, stressed but unaccented, accented and nuclear accented) and two levels of phrasing (intermediate and intonational phrase) [33, 34, 35]. Annotation and duration measurements were carried out in *Praat*. For statistical analyses *Statistica 10* was used.

2.2. Rhythm metrics

Segmentation into vocalic and consonantal intervals was based on the phonetic transcription and alignment. All vowels were marked as vocalic intervals and all consonants (except for post-vocalic glides) – as consonantal intervals. A vocalic interval could contain a single vowel or 2-3 subsequent vowels, or a vowel followed by a glide. Intervals could span across syllable and word boundaries. As in [18] prepausal intervals were not excluded from measurements and segments separated by a pause were treated as two distinct intervals. For each sentence we calculated the following metrics:

- **%V** – the proportion of vocalic intervals, **ΔV** and **ΔC** – the standard deviation of the duration of vocalic and consonantal intervals respectively [4]
- **rPVI-V** (raw Pairwise Variability Index) and **nPVI-V** (vocalic normalized Pairwise Variability Index): the mean of the duration differences between successive C intervals and the mean of the duration differences between successive V intervals divided by the sum of the same intervals respectively [5]
- **VarcoV/VarcoC**: standard deviation of vocalic/consonantal interval duration divided by mean vocalic/consonantal duration [6]

2.3. Time delay approach

This approach relies on time delay plots which are used to visualize relative timing of *functionally different transitions*, e.g. to stressed, to unstressed and to phrase final syllables. In time-delay plots, the duration of $syllable_i$ is plotted on the x-axis against the duration of $syllable_{i+1}$ plotted on the y-axis. Time delay plots can be applied to explore both short-term (syllable-level) and long-term (foot- and phrase-level) timing variability, and offer the possibility of interpreting continuous data in both continuous and categorical manner. By subtracting the duration of $syllable_i$ from the duration of $syllable_{i+1}$, we get information whether the transition is locally *accelerating*, *decelerating* or *isochronous* (continuous description). Decelerating transitions are plotted above the diagonal, accelerating transitions below it and the isochronous ones – along the diagonal (Figure 1).
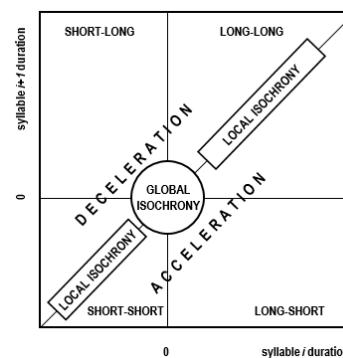

Figure 1: *Interpretation of a time-delay plot (based on [1]).*

Transitions can also be grouped into one of the four categories: short-short, long-long (locally isochronous transitions), short-long and long-short (an alternating rhythm). For example, if durations of $syllable_i$ and $syllable_{i+1}$ are both above/below the mean, the transition is categorized as long-long/short-short. A tendency towards global isochrony is indicated by concentration of data points in the center of the time-delay plot. The plots can be interpreted quantitatively, e.g. by performing one-factorial ANOVA with the duration difference $syllable_{i+1} - syllable_i$ as dependent variable and transition type as predictor variable (see also [1]). On a higher

level of rhythmic-prosodic organization, i.e. foot level, time-delay plots can be applied to visualize relative timing relations within feet and to observe *compensatory shortening* which contributes to impression of isochrony in stress-timed rhythm.

3. Results

3.1. Quantification using rhythm metrics

Comparison of the metric scores obtained in the current study with those reported in [18] (both studies used the same method of consonantal and vocalic interval measurements) shows that from among six rhythmically different languages, i.e. German, English, Spanish, Italian, Greek and Korean, Polish is characterized by the least variability in vocalic interval duration − it has the lowest nPVI and VarcoV. In terms of amount of vocalic speech (%V) Polish is ranked in between languages traditionally considered stress-timed (German, English) and those regarded as rhythmically unclassified (Greek, Korean) and syllable-timed (Italian, Spanish). High values of the raw consonantal metrics, i.e. ΔC and rPVI, indicate similarity to stress-timed German and English. On the contrary, VarcoC ranks Polish in between Greek, syllable-timed Italian and Spanish on the one hand (low VarcoC – low variability in consonantal interval duration) and stress-timed German and English, and Korean on the other (high VarcoC).

As regards the effect of speech rate (measured in the number of C and V intervals per second, cf. [36]) on the metric scores, significant inverse correlations were found for ΔV, ΔC and rPVI, indicating instability of these metrics. For the purpose of quantification of the actual distances between Polish and the six rhythmically different languages in the rhythm space determined by the most stable metrics, i.e. VarcoV-VarcoC and %V-nPVI (the latter, contrary to %V-VarcoV, were not significantly correlated with each other), Euclidean distances were calculated (Table 1). As they show, according to %V-nPVI, Polish is close to syllable-timed Italian and Spanish, but VarcoV-VarcoC place Polish the closest to German.

Table 1. *Euclidean distances between Polish and six rhythmically different languages.*

rank	VarcoV-VarcoC	%V- nPVI
1	German (7,3)	Spanish (7,4)
2	Spanish (9,3)	Italian (7,6)
3	Italian (10,7)	Greek (11,2)
4	English (10,8)	Korean (12,4)
5	Korean (14,2)	German (13,6)
6	Greek (14,3)	English (17,9)

Unlike VarcoV-VarcoC, the combination of %V-nPVI provides some support for the rhythm typology – values of these metrics (Figure 2) and the Euclidean distances based on them separate syllable-timed Spanish and Italian from stress-timed German and English. The metrics do not provide strong evidence for an "intermediate" rhythm in Polish: while values of %V and VarcoC are indeed intermediate between languages traditionally regarded as stress- and syllable-timed, nPVI and VarcoV place Polish in a very different region from all the other languages (Figure 2).

Generally, it can be assumed that low nPVI and VarcoV in Polish reflect lack of reduction of unstressed vowels on the

one hand, and lack of significant stress-based lengthening of vowels on the other (see also discussion in section 1.2).
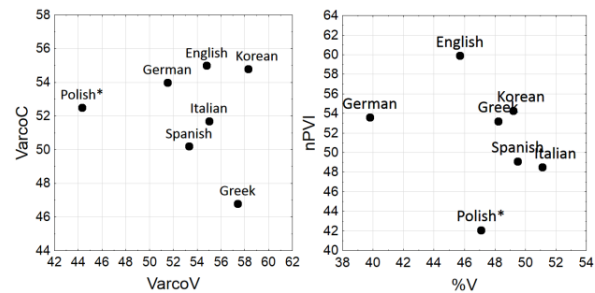

Figure 2: *Rhythm space determined by VarcoV-VarcoC (left) and %V-nPVI (right) for the six languages in [18] and Polish (*current study).*

As for VarcoC in Polish, its values (Figure 2) reflect the phonotactic structure of the utterances: high percentage (about 50%) of simple CV syllables (as in syllable-timed languages), high variation in complex syllable structures (indicating a tendency towards stress-timing) and frequency of very complex syllables even higher than in stress-timed languages.

3.2. Analysis with time-delay plots

3.2.1. Short-term timing variability

Table 2 presents concentrations of functionally different transitions in the four relative timing quadrants: Transitions to stressed syllables are concentrated mainly in the short-long and short-short quadrants, transitions to unstressed syllables – in the short-short quadrant, and transitions to phrase final syllables – in the long-long and short-long quadrants.

Table 2. *Concentrations of functionally different transitions in the four relative timing quadrants (based on z-score normalized syllable durations).*

transition	short-long	long-long	long-short	short-short
to stressed	37%	15%	16%	32%
to unstressed	21%	7%	23%	49%
to phrase final	29%	57%	3%	11%
total:	28%	18%	19%	35%

Generally, it seems that Polish, unlike English or French, *favors short-short sequences* the most. The predominance of short-short sequences in transitions to unstressed syllables, and long-long sequences in transitions to phrase final syllables, indicates a tendency towards *local isochrony* (like in syllable-timed Italian, [1]). At the same time, the high count of short-long transitions to stressed syllables, and to a lesser extent to phrase final syllables, indicates a tendency towards *alternation* (like in stress-timed English, [1]): Such timing patterns may contribute to the impression of an *intermediate* or *mixed* rhythm in Polish. Final lengthening is confined to the syllable at the edge of the phrase, but very often previous syllable is also lengthened (as indicated by high percentage of long-long transitions, see also Figure 3) due to co-occurrence with stress. One-factorial ANOVA showed significant *effect of transition type on relative timing of the sequence of syllables* expressed by the difference $syllabledur_{i+1} - syllabledur_i$ (indicating local acceleration or deceleration): $F=50.9$, $df=2$, $p<0.01$. Post-hoc

comparisons revealed that this effect is due to a strict trend towards *deceleration in transitions to stressed* (mean=0.5, σ=1.3) *and phrase final syllables* (mean=0.45, σ=1.5) on the one hand, and *acceleration in transitions to unstressed syllables* (mean=-0.07, σ=1.1) on the other. Figure 3 shows relative timing of transitions from unstressed to stressed syllables, with a distinction between lexical stress, pitch accent and nuclear accent, and transitions to phrase-final syllables.
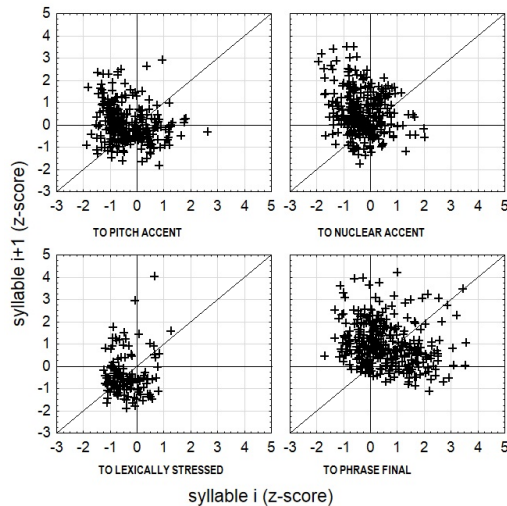

Figure 3: *Relative timing in functionally different transitions.*

Transitions to pitch accents have very similar distribution to the general one (when all prominent syllables are classified as stressed, see Table 2), whereas transitions to nuclear accents and lexically stressed syllables are concentrated mostly in the short-long and short-short quadrants respectively. These distributions indicate that durational marking is reserved mostly for nuclear accents – major prominences. One-factorial ANOVA showed significant differences in the relative timing between the three transition types, i.e. to lexically stressed, pitch accented and nuclear accented syllables: F=23, df=2, p<0.01. There is a strong *deceleration trend that increases with the level of prominence* from lexical stress (mean=0.02, σ=1) to nuclear accent (mean=0.86, σ=1.3), with pitch accents in between (mean=0.4, σ=1.2).

3.2.2. Long-term timing variability

The goal of the analysis of long-term characteristics of feet of different sizes is to "detect timing regularities that listeners can learn in order to form certain expectations concerning upcoming rhythmical events. Such long-term expectations are what we defined as *meter*" ([1], p. 162). It can be seen in Figure 4a, that in binary non-final feet, the stressed and the following unstressed syllable are almost identical in length, which indicates a tendency towards local isochrony. In this respect, Polish differs from both stress-timed English (-> increased duration on foot-initial stressed syllables) and syllable-timed French (-> a strong tendency towards foot-final lengthening) [1]. Binary phrase-final feet in Polish have a very similar timing pattern to that observed in French i.e., foot-final lengthening [1]. Ternary and quaternary feet have a tendency of acceleration after the foot-initial, stressed syllable and deceleration at foot boundary. Foot internal syllables (syl2 in ternary and syl3 in quaternary feet) are significantly shorter than foot-initial and foot-final syllables.
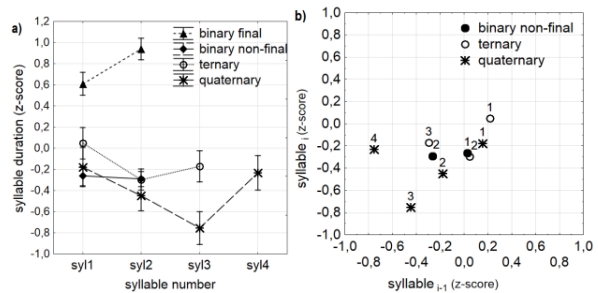

Figure 4: *Long-term timing patterns across different feet.*

These timing patterns are similar to those found in German longer feet [1]. Deceleration at foot boundary may cause the effect of *time shrinking* of the pre-final syllable and may lead to the *impression of isochrony and a lack of variation* in longer feet. Time shrinking may also concern stressed syllables in binary phrase-final feet. What can be traced in Figure 4b is the *compensatory shortening* – distribution of average relative durations of foot-internal syllables in binary non-final, ternary and quaternary feet shows tendency to shorten syllables with increasing foot length, but foot-final syllables are not affected by this phenomenon.

4. Discussion and conclusions

The results of the analysis with rhythm metrics indicated very low variability in vocalic interval duration (VarcoV and nPVI) in Polish comparing to six rhythmically different languages. In terms of variability in consonantal interval duration (VarcoC) and the amount of vocalic speech (%V) Polish can be regarded as *intermediate* between syllable- and stress-timed languages. Euclidean distances showed that according to %V-nPVI, Polish is close to Italian and Spanish, but VarcoV-VarcoC place Polish the closest to German. The description of Polish rhythm provided by the metrics is thus inconclusive and hard to interpret. On the contrary, time-delay approach appeared to be a very efficient method for visualization and analysis, both quantitative and qualitative, of the contribution of short-term and long-term timing variability to rhythmic structure and rhythmic grouping at different hierarchical levels (syllable, foot and phrase level). The time-delay analysis provided some evidence for "mixed" rhythm in Polish which is characterized by *as much local isochrony* (-> syllable-timing) *as alternation* (-> stress-timing). Speech rhythm in Polish is also determined by *deceleration* in transitions to *phrase-final and stressed syllables* (the trend increases with the level of prominence), and *acceleration* in transitions to *unstressed syllables*. The analysis of long-term timing variability brought some evidence of *time shrinking* and *compensatory shortening*, which may lead to the impression of syllable- and stress-based isochrony respectively. The analyses showed that rhythmic grouping is determined by *deceleration* at the foot and phrase boundary. As regards rhythmic structure in Polish, it is necessary to go beyond the dimension of timing, because the current study showed that durational marking of prominence is very weak, if any.

5. Acknowledgements

6. References

[1] Wagner, P. (2008). The rhythm of language and speech: Constraining factors, models, metrics and applications. Habilitationsschrift, University of Bonn.

[2] Pike, K. L. (1945). The Intonation of American English. Ann Arbor: University of Michigan.

[3] Abercrombie, D. (1967). *Elements of general phonetics* (Vol. 203). Edinburgh: Edinburgh University Press.

[4] Ramus, F., Nespor, M. & Mehler, J. (1999). Correlates of linguistic rhythm in the speech signal. *Cognition 73*(3). 1-28.

[5] Grabe, E., & Low, E. L. (2002). Durational variability in speech and the rhythm class hypothesis. *Papers in laboratory phonology, 7*(515-546).

[6] Dellwo, V. (2006). Rhythm and speech rate: A variation coefficient forΔ C. *Language and language processing*, 231-241.

[7] Mairano, P. and Romano, A. (2011). Rhythm metrics for 21 languages. In *Proceedings of the XVIIth International Congress of Phonetic Sciences*, 17-21 August 2011, Hong Kong, China. 1318-1321.

[8] Payne, E., Post, B., Astruc, L., Prieto, P. & del Mar Vanrell, M. (2012). Measuring child rhythm. *Language and Speech, 55*(2), 203-229.

[9] White, L. & Mattys, S. L. (2007). Calibrating rhythm: First language and second language studies. *Journal of Phonetics 3*(5). 501-522.

[10] Mok, P. P., & Dellwo, V. (2008). Comparing native and non-native speech rhythm using acoustic rhythmic measures: Cantonese, Beijing Mandarin and English. In *Proceedings of the Speech Prosody 2008 Conference* (pp. 423-426).

[11] Tortel, A., & Hirst, D. (2010). Rhythm metrics and the production of English L1/L2. In *Proceedings of Speech Prosody 2010*.

[12] Ordin, M., Polyanskaya, L. & Ulbrich, Ch. (2011). Acquisition of Timing Patterns in Second Language. In *Proceedings of INTERSPEECH 2011*, 27-31 August 2011, Florence, Italy. 1129-1132.

[13] Kinoshita, N., & Sheppard, C. (2011). Validating acoustic measures of speech rhythm for second language acquisition. In *Proceedings of the XIIth International Congress of the Phonetic Sciences* (Vol. 17, pp. 1086-1089).

[14] Li, A. and Post, B. (2012). L2 rhythm development by Mandarin Chinese learners of English. In *Proceedings of Perspectives on Rhythm and Timing*, 19-21 July 2012, Glasgow, UK.

[15] Liss, J., White, L., Mattys, S., Lansford, K., Lotto, A., Spitzer, S. & Caviness, J. (2009). Quantifying speech rhythm abnormalities in the dysarthrias. *Journal of Speech, Language, and Hearing Research* Vol.52 1334-1352.

[16] Leemann, A., Dellwo, V., Kolly, M. J., & Schmid, S. (2012). Rhythmic variability in Swiss German dialects. In *Proceedings of Speech Prosody 2012*, Shanghai, China.

[17] Arvaniti, A. (2009). Rhythm, timing and the timing of rhythm. *Phonetica, 66*(1-2), 46-63.

[18] Arvaniti, A. (2012). The usefulness of metrics in the quantification of speech rhythm. *Journal of Phonetics, 40*(3), 351-373.

[19] Wiget, L., White, L., Schuppler, B., Grenon, I., Rauch, O. & Mattys, S. L. (2010). How stable are acoustic metrics of contrastive speech rhythm? *J. Acoust. Soc. Am 127*(3). 1559-1569.

[20] Richter, L. (1978). Wpływ pozycji w zestroju akcentowym na czas trwania głosek. (The effect of the position in the accent group on the phoneme duration). *Lingua Posnaniensia*, vol. 21.

[21] Imiolczyk, J., Nowak, I., Demenko, G. (1994). High intelligibility text-to-speech synthesis for Polish. *Archives of Acoustics 19* (2), pp. 161-172.

[22] Demenko, G. (1999). Analysis of Polish Suprasegmentals for needs of Speech Technology. UAM: Poznań.

[23] Richter, L. (1987). Modelling of the rhythmic structure of utterances in Polish. *Studia Phonetica Posnaniensia* 1, 91–125.

[24] Klessa, K. (2012). Polish segmental duration: selected observations based on corpus data. *Speech and Language Technology*, vol. 14/15, pp. 95-104.

[25] Malisz, Z., & Wagner, P. (2012). Acoustic-phonetic realisation of Polish syllable prominence: a corpus study. *Speech and Language Technology*, vol. 14/15, pp. 105-114

[26] Dłuska, M. (1933). Próba badań nad trwaniem spółgłosek polskich w zależności od brzmienia. (An attempt at investigating duration of Polish consonants depending on their quality). *Slavia Occidentalis* vol. 12, pp. 288-297.

[27] Jassem, W. (1962). Akcent języka polskiego (Accent of Polish).

[28] Klessa, K. (2006). Analysis of segmental duration for needs of speech synthesis in Polish. Unpublished PhD dissertation, Adam Mickiewicz University, Poznań.

[29] Wagner, A. (2010). Acoustic cues for automatic determination of phrasing. In *Proceedings of Speech Prosody 2010*, 11-14 May 2010, Chicago, USA.

[30] Wagner, A. (2012). Speech rhythm in native and non-native Polish, In *Proc. of ISCA Workshop on Experimental Llinguistics*, Athens, 27-29 August 2012.

[31] Demenko G., Wypych M. & Baranowska E. (2003.) Implementation of Polish grapheme-tophoneme rules and extended SAMPA alphabet in Polish text-to-speech synthesis. *Speech and Language Technology*, vol. 7, pp. 79-96.

[32] Ostaszewska, D., & Tambor, J. (2000). Phonetics and phonology of contemporary Polish. Polish Scientific Publishers (PWN).

[33] Prieto, P., Vanrell, M. D. M., Astruc, L., Payne, E., & Post, B. (2012). Phonotactic and phrasal properties of speech rhythm. Evidence from Catalan, English, and Spanish. *Speech Communication, 54*(6), 681-702.

[34] Beckman, M. E., & Edwards, J. (1994). Articulatory evidence for differentiating stress categories. *Papers in laboratory phonology III: Phonological structure and phonetic form*, 7-33.

[35] Selkirk, E. O. (1995). Sentence prosody: Intonation, stress and phrasing. In Goldsmith, J., editor, Handbook of Phonological Theory. Oxford.

[36] Beňuš, Š. & Šimko, J. (2012). Rhythm and tempo in Slovak. In *Proceedings of the 6th International Conference on Speech Prosody*, Shanghai, China.