Agnieszka Wagner

Stability of timing patterns in Polish. Experimental verification of the rhythm class hypothesis

Abstract

This paper investigates the stability of timing patterns in Polish across different speaking rates, phonotactic and poetic metrical structures and elicitation methods. Timing patterns of utterances from a recently developed *Polish rhythmic database* were described by means of *rhythm metrics*: PVI, ΔV, ΔC and %V and Varco. The rhythm scores were subjected to statistical analyses that showed a significant effect of the investigated factors on the timing patterns, which was reflected in the differences in the rhythm scores and in the position of Polish data in a two-dimensional "rhythm space". Analyses of cross-linguistic and within-language durational variability and an attempt to determine rhythm class affiliation of the Polish language (a "rhythmic outlier") failed to provide support for the *rhythm class hypothesis*. It is claimed that speech rhythm can not be adequately described by means of measures which rely solely on durational variability, and that studies which apply rhythm metrics to investigate speech timing or rhythm should take into account numerous limitations of such an approach.

1. Introduction and Background

For many years the research on speech rhythm has been dominated by the notion of *isochrony* (Pike, 1945; Abercrombie, 1967), which was the central concept underlying the so-called *rhythm class hypothesis*. According to this hypothesis every language, on the basis of its stable contrastive timing properties, can be assigned to one of two main rhythmic categories[1]: stress-timing (e.g., English, Dutch) or syllable-timing (e.g., Spanish, Italian). However, since neither strict foot- nor syllable-isochrony in speech production has been empirically proven, the rhythm class hypothesis lost its significance for some time. The concept of syllable- and stress-timed rhythm was revived by the introduction of so-called *rhythm metrics* – formulas designed to quantify linguistic rhythm based on Dauer's claim (1983, 1987) that perceived differences in rhythms of various languages have their origin in segmental phonology and phonotactics. The most commonly used metrics include:

- Pairwise Variability Indices – PVIs (Grabe & Low, 2002) which express the level of durational variability in successive intervals – vocalic (rate-normalized index, nPVI-V) and consonantal (raw index, rPVI-C)

[1] The third rhythmic category of *mora-timing*, exemplified by Japanese, is not discussed here.

- %V – percentage of vocalic intervals, ΔV – standard deviation of vocalic interval duration, and ΔC – standard deviation of consonantal interval duration (Ramus et al., 1999)
- Variation Coefficients, Varcos (Dellwo, 2010) for ΔC (VarcoC) and ΔV (VarcoV)

In parallel with an increasing interest in and application of the rhythm metrics to study topics as diverse as language, dialect or speaker discrimination, acquisition of rhythm in L2 and temporal characteristics of disordered speech, evidence has piled up against the categorical and metric-based approach to linguistic rhythm. First of all, production, perception and machine classification studies have provided, at most, limited support for the rhythmic typology. Arvaniti and Rodriquez (2013) showed that, at the perception level, timing differences on their own are insufficient for rhythmic discrimination of languages and that timing is not processed independently of other prosodic features. Loukina et al. (2011) reported only moderate success of rhythm metrics in separation and identification of languages in machine classification experiments. In a similar study, Horton and Arvaniti (2013) concluded that "the metrics model the rhythm typology only to the extent that they correlate with tempo" (p. 46). Secondly, rhythm metrics lack efficiency in providing an unambiguous rhythm class affiliation for a number of so called "unclassified" languages and show considerable sensitivity to speaking rates, text materials, speaking styles and speakers (Dellwo, 2010; White & Mattys, 2007a; Wiget et al., 2010). Consequently, the within-language variability in timing, as indicated by the metrics, can sometimes be greater than that between languages (Arvaniti, 2012). What presents another problem is that metrics do not well capture the intended properties of languages: "Systemic differences in the phonologies of diverse languages are not well reflected in the metrics […], nor are subtle (yet perceptible) realisational differences due to the dialect-specific implementation of the prosodic hierarchy and articulatory strategy" (Rathcke & Smith, 2015, p. 27).

The present study applies rhythm metrics to investigate the stability of timing patterns in Polish, a language regarded as rhythmically *unclassified* or *mixed*, and to provide insights into its possible rhythm class affiliation. The study also investigates variation in rhythm scores associated with different speaking rates, phonotactic and poetic metrical structures and elicitation methods. On the basis of the findings reported in the literature, a significant within-language variation in the rhythm scores is expected, e.g., higher scores (indicating greater variability in duration) at slower than at normal and faster tempi, higher scores for phonotactically complex than phonotactically simple sentences, or lower scores for controlled (read) speech as opposed to more natural (spontaneous) speech. The results of the current study can have important implications for the rhythm class hypothesis and/or the assumption according to which rhythm metrics reflect stable contrastive properties of languages that contribute to perceptually different rhythms (see also Wiget et al., 2010).

2. Methodology

2.1. Data

The speech material for the study comes from the *Polish Rhythmic Database* (Wagner, Klessa & Bachan, 2016) which has been developed specifically for research on rhythm, timing and prosody in general, and is the largest database of this type (altogether 5 hours of speech) currently available for Polish. Only the *L1 Polish subcorpus* was used in the present study (Table 1). It contains recordings of 19 native speakers of Polish (11 females, 8 males) who provided readings of the story “The north wind and the sun”, 3 sets of 5 sentences each varying in their phonotactic complexity (“stress-timed” – containing the most complex syllable structures and instances of vowel hiatus, “syllable-timed” – only simple syllable structures, and “uncontrolled” – the most frequent syllable structures; cf. Arvaniti, 2012), fragments of 4 poems representing different poetic metrical structures (iamb, trochee, Polish hexameter and amphibrach), two read mini-dialogues and spontaneous monologue which was elicited using the diapix task (Bradlow et al., 2007). The story, sentences and poems were produced at 5 intended tempi (normal, fast, very fast, again normal tempo, slow, very slow, therefore 2 normal rate conditions).

<table>
<tr><th>Speech materials</th><th>Intended speaking rate</th></tr>
<tr><td>story “The north wind and the sun”</td><td rowspan="3">normal,
fast,
very fast,
normal (2nd repetition),
slow,
very slow</td></tr>
<tr><td>3 sets of 5 sentences:
syllable-timed, e.g. (CVCV)(CVCV)(CV)(CV)...
uncontrolled, e.g. (CVCCVCCCV)(CVCCV)(CVCCVCV)...
stress-timed, e.g. (CVCCCVC)(C)(CVCCV)(V)(VCCVC)...
(brackets indicate word boundaries)</td></tr>
<tr><td>poems:
trochee: _’ _/_’ _/
iamb: _ _’ /_ _’ /
amphibrach: _ _’_ /_ _’_/
Polish hexameter: _’ _/_’ _ _/_’ _/ + _’ _ _/_’ _ _/_’ _/</td></tr>
<tr><td>2 mini-dialogues</td><td rowspan="2">normal</td></tr>
<tr><td>spontaneous monologue</td></tr>
</table>

Table 1: Features of the Polish L1 Subcorpus from the Polish Rhythmic Database (Wagner et al., 2016).

All recordings were segmented into inter-pausal intervals, automatically transcribed and aligned at word, syllable and phoneme level, and the results were manually verified by four trained labelers and approved by an expert phonetician. The segmentation and annotation of the speech material was carried out in *Annotation Pro* (Klessa, 2016) which is integrated with the *Polish Rhythmic Database*. Vocalic and consonantal intervals were derived automatically from the phonemic annotations. For every inter-pausal interval 8 rhythm metrics – syllabic and vocalic nPVI, consonantal rPVI, Varcos, %V, ΔV and ΔC, were calculated as in Dellwo (2010) using *Annotation Pro* plugins. Prepausal intervals were included in the measurements (Arvaniti, 2012).

2.2. Analysis steps

In order to study timing patterns in different phonotactic structures and at different speaking rates, repeated-measures ANOVAs were performed with mean speaker scores for each metric from each sentence type as the dependent variable, tempo as a repeated-measures factor and phonotactic complexity as a categorical predictor. The same method of analysis was chosen to study timing patterns in different poetic metrical structures and in the story (the latter without a categorical predictor). The effect of elicitation method was analyzed by means of an ANCOVA separate slopes model with elicitation method (dialogue, spontaneous monologue, story, sentences and poems – the latter three only at a normal tempo, scores pooled across phonotactic and metrical structures) as a categorical predictor and tempo (in syl./sec.) as a continuous predictor (cf. Clopper & Smiljanic, 2015). Fisher's least significant difference (LSD) test was applied to perform pairwise comparisons following a significant overall test result. Statistical significance was defined at $p < .05$. Partial η^2 was calculated for each factor and each metric separately and used as an estimate of effect size. For ease of presentation and clarity of the text, the results of all the statistical analyses are presented in tables in Appendix.

In order to compare within- and between-language variation in the rhythm scores and to investigate the rhythm class affiliation of Polish, pairs of vocalic and consonantal metric scores were used as coordinates, and data for Polish and prototypical stress- and syllable-timed languages (based on Arvaniti, 2012) was plotted in a two-dimensional “rhythm space”. The actual distances between the data points were quantified by means of Euclidean distances.

3. Results

3.1. Timing patterns in Polish

3.1.1. Phonotactic complexity & tempo

Repeated-measures ANOVAs revealed main effect of tempo to be significant for all metrics except syllabic nPVI and Varcos, and significant effect of phonotactic structure for all metrics except ΔV. For tempo the effect size varied from marginal for nPVI-V [$\eta_p^2 = .11$] to substantial for rPVI-C [$\eta_p^2 = .91$] and deltas [**ΔV**, $\eta_p^2 = .76$; **ΔC**, $\eta_p^2 = .91$], and was larger for the raw than the rate-normalized metrics, indicating greater stability of the latter across different speaking rates. As concerns phonotactic complexity the effect size was greater than that of tempo for all metrics (see: Table 2 in Appendix). This shows that durational variability within sentence types (due to increasing and decreasing the tempo) is lower than that between sentence types (due to differences in the phonotactic structure). The interaction between tempo and structure was significant for some rPVI-C, nPVI-syl, %V and ΔC, and the effect was greater for the consonantal than the vocalic measures.

Generally, the scores of deltas and rPVI-C increased gradually with tempo deceleration and decreased with tempo acceleration, which corroborates previous results (Wagner & Dellwo, 2003; Clopper & Smiljanic, 2015). As for the effect of sentence structure, it could be observed that more complex phonotactics result in greater durational variability as indicated by significantly higher scores for "stress-timed" than for "syllable-timed" sentences (Figure 1). Fisher's LSD post-hoc test showed a significant difference among the three sentence types for all metrics except syllabic and vocalic nPVIs which grouped "uncontrolled" and "stress-timed" sentences together. The interaction effects indicated that speaking rate variation did not affect all sentence types equally and that at some tempos the differences due to phonotactic complexity were neutralized as indicated by the same scores of ΔC/rPVI-C at a normal speaking rate in the "uncontrolled" set and at fast rate in "stress-timed" sentences/slow tempo in "syllable-timed" set.

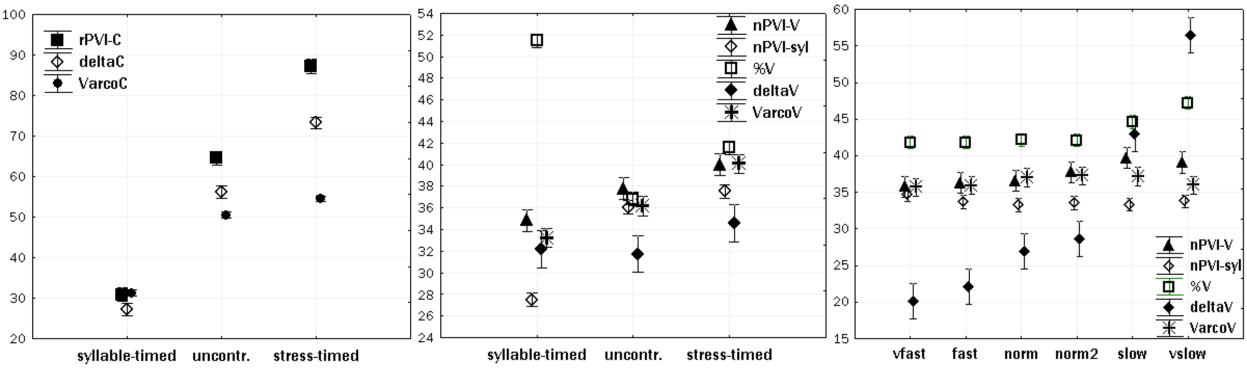

Figure 1: Means and 95% confidence intervals of consonantal and vocalic metrics in phonotactically simple (syllable-timed), complex (stress-timed) and uncontrolled sentences, and at different tempi (only for V-metrics; see the figure to the right).

3.1.2. Poetic metrical structure & tempo

The effect of tempo was significant for all metrics, but the size of the effect varied. Generally, rate-normalized metrics were affected less than raw measures and vocalic metrics less than the consonantal ones. The scores of rPVI-C and deltas showed again a significant increase in durational variability with tempo decrease (Figure 2).

All metrics also showed a main effect of poetic metrical structure. The fact that the size of the effect was larger than the tempo effect for the same set of data for all metrics except deltas and rPVI-C, indicates higher sensitivity of the latter to differences in speaking rate than in metrical structure (see: Table 3 in Appendix). In addition, nPVI-syl and VarcoC showed an interaction between poetic meter and tempo [for **nPVI-syl**, $F(15, 330) = 2.29$; **VarcoC**, $F(15, 330) = 2.41$].

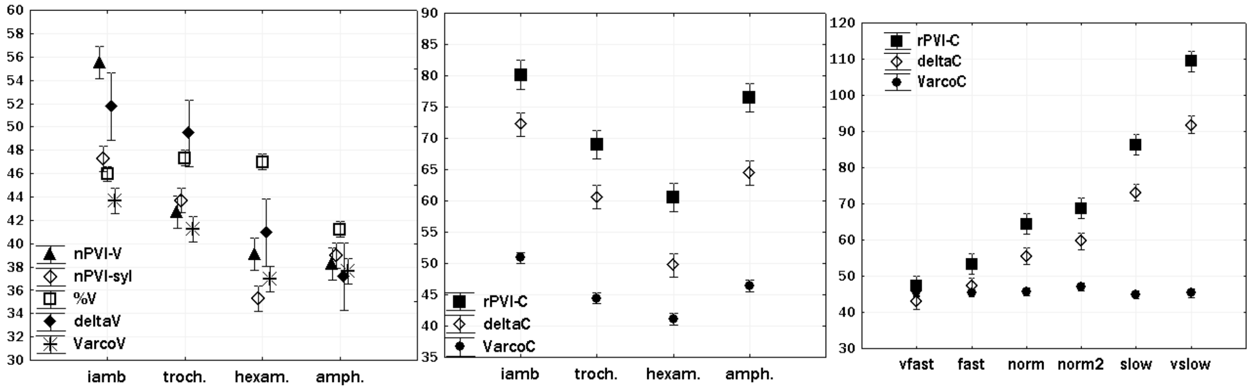

Figure 2: Means and 95% confidence intervals of vocalic and consonantal metrics in different poetic meters (iambic, trochaic, Polish hexameter and amphibrach) and at different speaking rates (only for C-metrics; see the figure to the right).

As for the effect of metrical structure, the scores of all vocalic metrics (except %V) showed the expected pattern (see: Barry, Andreeva & Koreman, 2009) of higher scores (i.e., higher durational variability) in iambic and trochaic lines than in more complex metrical forms (Figure 2). Additionally, pairwise comparisons showed that nPVI-syl distinguished among all the metrical structures and nPVI-V also between trochaic and iambic lines. The scores of consonantal metrics differed consistently among iambs, trochees and hexameters, with iambs displaying the greatest durational variability and hexameters the least variability. The grouping of amphibrachs was inconsistent across the metrics (Figure 2).

3.1.3. Tempo effects in story

Repeated-measures ANOVAs showed significant effect of tempo for deltas, rPVI-C, %V and VarcoV [for **rPVI-C**, $F(5, 90) = 98.5$; **%V**, $F(5, 90) = 16.3$; **ΔV**, $F(5, 90) = 38.9$; **ΔC**, $F(5, 90) =.117$; **VarcoV**, $F(5, 90) = 3.2$]. The size of the effect was larger for the consonantal than the vocalic metrics as indicated by η_p^2 values [**rPVI-C**, $\eta_p^2 = .85$; **%V**, $\eta_p^2 = .48$; **ΔV**, $\eta_p^2 = .67$; **ΔC**, $\eta_p^2 = .87$; **VarcoV**, $\eta_p^2 = .15$]. Pairwise comparisons using Fisher's LSD test indicated that the metrics were affected by tempo in a similar fashion to that observed in the previous analysis, i.e. the proportion of vocalic speech was significantly lower at faster and normal than at slower rates, the scores of VarcoV were significantly lower at slower than at faster rates, and scores of rPVI-C and deltas increased with tempo deceleration.

3.1.4. Elicitation method

Significant differences in mean rhythm scores across elicitation methods were found for all metrics; however, the size of the effect was marginal (i.e., below 1%). After introducing tempo as a continuous predictor into the model and adjusting the mean scores of the metrics for average syllable rate, the model gained some more explanatory power as shown by higher η_p^2 (see Table 3 in Appendix, *elicitation*syl_rate*). Nevertheless, it accounted for at most 23% of variability in the rhythm scores. This indicates that additional factors (some of which had already been identified in the previous analyses) could be included in the model to explain rhythm scores more fully.

Planned comparisons using adjusted mean scores of all vocalic metrics except %V showed significantly greater variability in timing patterns in spontaneous monologue than in other speaking styles (Figure 3). In addition, nPVI-V and nPVI-syl showed significantly lower durational variability in vowels and syllables in the sentences compared to other elicitation methods. These results corroborate the findings presented in Arvaniti (2012). Post-hoc tests using adjusted mean scores of consonantal metrics showed significantly lower variability in duration of consonantal intervals in dialogue and poetry than in the other styles.

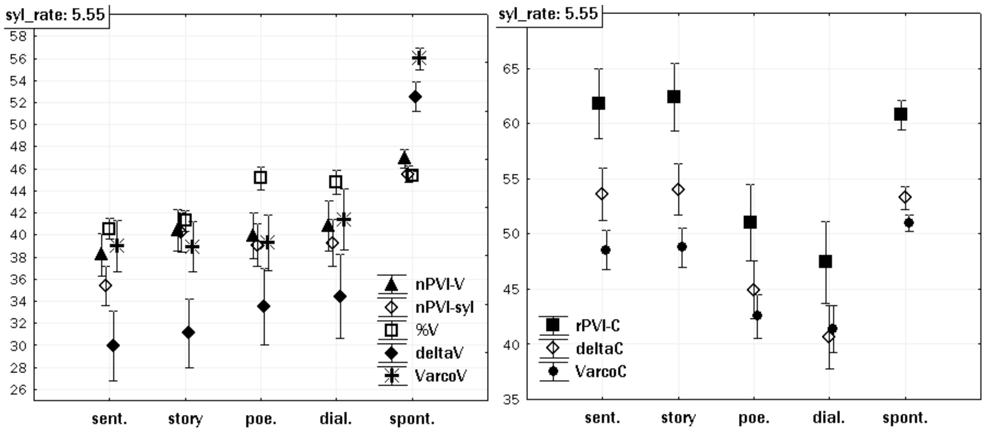

Figure 3: Means (adjusted for average syllable rate) and 95% confidence intervals of vocalic and consonantal metrics in different elicitation methods: sentences (pooled data), story, poems (pooled data), dialogues and spontaneous monologue.

3.2. Rhythmic classification of Polish

A cross-linguistic comparison of timing variability in Polish and prototypical stress- and syllable-timed languages (based on data from Arvaniti, 2012) indicated that in a two-dimensional "rhythm space" determined by the scores of the rhythm metrics (nPVI-V & rPVI-C, VarcoV & VarcoC, %V & ΔC) Polish is localized in regions other than those associated with stress- or syllable-timing (Figure 4). Polish seems to be a "rhythmic outlier", with the lowest variability in vowel duration as indicated by nPVI-V and VarcoV, relatively little vocalic speech as indicated by %V and moderate variability in consonantal interval durations. The plots in Figure 4 also indicate that grouping of languages on the basis of their timing patterns is inconsistent across the metrics. While PVIs and %V and ΔC provide the expected separation between English and German (stress-timing) on the one hand and Spanish and Italian (syllable-timing) on the other, Varcos do not. Further inconsistencies are revealed by the Euclidean distances: According to PVIs Polish is closer to the syllable-timed than stress-timed languages (SP: 11.6; IT: 13.4, DE: 17.8, EN: 24.3), but this does not hold for %V & ΔC, which place Polish close to Spanish, but also to English (SP: 8.5; IT: 11.4, DE: 12, EN: 10.7). Varcos indicate that Polish is equidistant from the two rhythm classes (SP: 16; IT: 18, DE: 15.6, EN: 19).

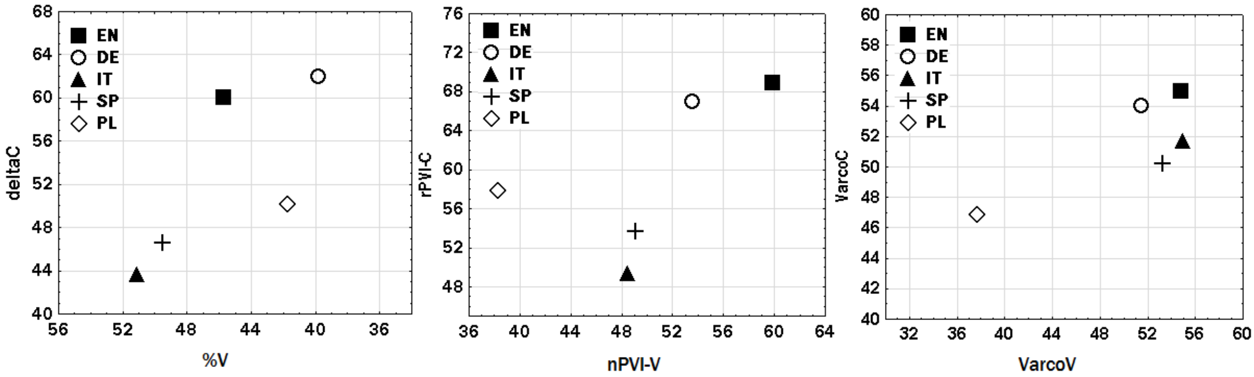

Figure 4: Distribution of Polish (PL), German (DE), English (EN), Spanish (SP) and Italian (IT) in the rhythm space. The scores of the three pairs of rhythm metrics were calculated for data at normal tempo and pooled across all elicitation methods.

What is even more problematic for the rhythmic classification of Polish is the significant amount of within-language durational variability that sometimes exceeds the between-language variability, e.g. the manipulation of the phonotactic complexity introduced higher durational variability in the Polish sentence materials than the variability between pooled Polish and English data as indicated by higher Euclidean distances between Polish uncontrolled and stress-timed sentences (PVIs: 29.9; %V & ΔC: 27.9; Varcos: 21.4) than between Polish and English (PVIs: 24.3; %V & ΔC: 10.7; Varcos: 19.0). Moreover, if we consider within-language durational variability induced by different elicitation methods, a substantial cross-linguistic overlap in timing patterns can be observed, e.g. the similar position in the %V & ΔC "rhythm space" of Polish spontaneous data and the English story (Euclidean dist. = 2.1) and English sentences (1.7), or the Polish story and the German sentences (1.7). This overlap is even more pronounced in case of the tempo-normalized Varcos. This shows that the amount of durational variability, which can vary greatly within a language, should not be treated as an inherent rhythmic property of a language, as it is in the rhythm metric-based approach to speech rhythm. Similarly to the previous analysis the metrics appear to be inconsistent, because they indicate a different amount of the within- and between-language variability in duration.

4. Discussion & Conclusions

The analyses concerning the stability of timing patterns (sec. 3.1) showed that none of the metrics appears to be "immune" to speaking rate variation. At the same time it could be seen that irrespective of the properties of the linguistic material (sentences, poems, etc.), rate-normalization indeed reduces (to some extent) the effect of tempo on the rhythm scores as indicated by lower η^2 for nPVI-V, nPVI-syl and Varcos. Generally, an *increase* in speaking rate was associated with a *decrease* in the overall duration variability indicated by lower rhythm scores at slower than at normal and faster rates. In particular, deltas and rPVI-C showed a consistent effect of tempo across different settings. The analysis with separate slopes model (which included speech only at a normal rate) indicated that to a different degree, part of the variability in the scores of a particular rhythm metric can be explained by variation in the speaking rate. The current results clearly demonstrate that any analysis which applies rhythm metrics to investigate contrastive timing patterns of a language needs to *include speaking rate as a covariate*. On the other hand, there is evidence suggesting that tempo constitutes *an inherent component* of speech rhythm (e.g., Dellwo & Wagner, 2003). In particular the results of Malisz (2013), who applied the *coupled oscillator framework* (O'Dell and Nieminen 1999, 2009) to analyze rhythmic patterning in Polish spontaneous dialogues (with no induced speaking rate variation), showed that at faster (natural) speaking rates the higher-level inter-stress oscillator dominates the lower-level syllabic oscillator resulting in a greater overall timing variability, which can be interpreted as a tendency towards stress-timing. At slower rates the situation is the reverse (i.e., the syllabic oscillator dominates). These results support the concept of *coexisting rhythms* in language (Nolan & Asu, 2009) and challenge the rationale for seeking inherent differences in rhythm class timing.

Apart from tempo, rhythm metrics are also sensitive to the phonotactic complexity of utterances, poetic meter and elicitation method. The relationship with phonotactic structure was consistent across the metrics: their scores increased, indicating greater durational variability, with increasing phonotactic complexity. The observed inconsistencies between vocalic and consonantal metrics in their relationship with poetic metrical structure and elicitation method may indicate that they capture different aspects of timing. The scores of vocalic metrics (nPVI-V, VarcoV) and nPVI-syl, but not of the consonantal measures of rhythm, showed the expected patterns of greater durational variability in simple than complex metrical structures and in more natural than controlled speaking styles (see also Barry et al., 2009; Arvaniti, 2012). One possible interpretation of these results is that the consonantal metrics reflect rather the phonotactic complexity of utterances, whereas the vocalic metrics (and nPVI-syl) are sensitive to differences in the underlying prosodic/rhythmic structure such as, for instance, variable rhythmical groupings and distribution of stresses/prominences in different poetic meters, or more variability in prosodic structures in spontaneous than in controlled speech/isolated sentences. But even if we take these considerations into account, it would still be unjustified to claim that the rate-normalized vocalic metrics (and nPVI-syl) provide a stable and comprehensive description of speech rhythm. The fact that listeners perceive differences between languages' rhythms cannot, and should not, be explained by referring to a single aspect of a spoken utterance, such as durational variability, given numerous cognitive, phonetic, phonological and linguistic

constraints on both rhythmic speech production and perception (Wagner, 2008).
The comparison of Polish with prototypical stress- and syllable-timed languages in a two-dimensional “rhythm space” (sec. 3.2) showed no evidence of an *intermediate* or *mixed rhythm* in Polish (Nespor, 1990; quoted after Grabe & Low, 2002). Instead, if we rely solely on the distribution of the metric scores for the Polish data, then Polish would appear to be a ”rhythmic outlier”. The comparison of within-language and cross-linguistic variation in timing provides more support for the view of a *rhythmic continuum* rather than a few discrete categories (e.g. White & Mattys, 2007b).
Altogether, the lack of a clear rhythm class affiliation of the Polish language, the significant amount of within-language variability in the timing patterns and inconsistencies between consonantal and vocalic metrics as well as between different pairs of metrics reported in the current study, call into question the status of rhythm metrics as acoustic correlates of linguistic rhythm and provide no support for the rhythm class hypothesis.

5. References

Abercrombie, D. (1967). *Elements of General Phonetics.* Edinburgh University Press.
Arvaniti, A. (2012). The usefulness of metrics in the quantification of speech rhythm. *Journal of Phonetics 40*(3), 351-373.
Arvaniti, A., & Rodriquez, T. (2013). The role of rhythm class, speaking rate, and F0 in language discrimination. *Laboratory Phonology 4*(1), 7-38.
Barry, W., Andreeva, B., & Koreman, J. (2009). Do rhythm measures reflect perceived rhythm?. *Phonetica 66*(1-2), 78-94.
Bradlow, A. R., Baker, R. E., Choi, A., Kim, M., Van Engen, K. J. (2007). The Wildcat corpus of native- and foreign-accented English. *Journal of the Acoustical Society of America*, *121*(5), pp. 3072-3072.
Clopper, C. G., & Smiljanic, R. (2015). Regional variation in temporal organization in American English. *Journal of Phonetics 49,* 1-15.
Dauer, R. (1983). Stress-timing and syllable-timing reanalyzed. *Journal of Phonetics*, *11*, 51–62.
Dauer, R. (1987). Phonetic and phonological components of language rhythm. In U. Viks (Ed.), *Proceedings of the 11th international congress of phonetic sciences* (pp. 447–450). Tallinn, Estonia.
Dellwo, V. (2010). *Influences of speech rate on the acoustic correlates of speech rhythm*. Unpublished PhD thesis, Universität Bonn, Germany.
Dellwo, V., & Wagner, P. (2003). Relations between language rhythm and speech rate. In D. Recasens, M. J. Solé, & J. Romero (Eds.), *Proceedings of the 15th International Congress of*

Phonetic Sciences (pp. 471-474). Barcelona, Spain.
Grabe, E., & Low, E. L. (2002). Durational variability in speech and the rhythm class hypothesis. *Laboratory Phonology, 7,* 515-546.
Horton, R., & Arvaniti, A. (2013). Clusters and Classes in the Rhythm Metrics. *San Diego Linguistics Papers 4,* 28-52.
Klessa, K. (2016). *Annotation Pro. Enhancing analyses of linguistic and paralinguistic features in speech.* The Faculty of Modern Languages and Literature, Adam Mickiewicz University in Poznan. ISBN 978-83-946017-0-6. Loukina, A., Kochanski, G., Rosner, B., Keane, E., & Shih, C. (2011). Rhythm measures and dimensions of durational variation in speech. *Journal of the Acoustical Society of America 129*(5), 3258-3270.
Malisz, Z. (2013). *Speech rhythm variability in Polish and English: A study of interaction between rhythmic levels.* Unpublished PhD dissertation, Adam Mickiewicz University in Poznan.
Nolan, F., & Asu, E. L. (2009). The pairwise variability index and coexisting rhythms in language. *Phonetica*, *66*(1-2), 64-77.
O'Dell, M., & Nieminen, T. (1999). Coupled oscillator model of speech rhythm. *Proceedings of the XIVth international congress of phonetic sciences,* Vol. 2, pp. 1075-1078.
O'dell, M. L., & Nieminen, T. (2009). Coupled oscillator model for speech timing: Overview and examples. *Proceedings of the Xth conference of Nordic Prosody,* Helsinki (pp. 179-190).
Pike, K. (1945). *The Intonation of American English.* Ann-Arbor: University of Michigan Press.
Ramus, F., Nespor, M., & Mehler, J. (1999). Correlates of linguistic rhythm in the speech signal. *Cognition 73,* 265-292.
Rathcke, T. V., & Smith, R. H. (2015). Speech timing and linguistic rhythm: On the acoustic bases of rhythm typologies. *Journal of the Acoustical Society of America, 137*(5), 2834-2845.
Wagner, P. (2008). *The Rhythm of Language and Speech: Constraining Factors, Models, Metrics and Applications.* Unpublished Habilitation Monograph, Universität Bonn, Germany.
Wagner, A., Klessa, K., & Bachan, J. (2016). Polish rhythmic database – new resources for speech timing and rhythm analysis. *Proceedings of 10th International Conference on Language Resources and Evaluation* (pp. 4678-4683). Paris, France: ELRA. ISBN 978-2-9517408-9-1.
White, L., & Mattys, S. L. (2007a). Calibrating rhythm: First language and second language studies. *Journal of Phonetics 3*(5), 501-522.
White, L., & Mattys, S. L. (2007b). Rhythmic typology and variation in first and second languages. *Amsterdam Studies in the Theory and History of Linguistic Science 4*, 237-282.
Wiget, L., White, L., Schuppler, B., Grenon, I., Rauch, O., & Mattys, S. L. (2010). How stable are acoustic metrics of contrastive speech rhythm? *Journal of the Acoustical Society of America*

127, 1559-1569.

6. Acknowledgements

This research has been carried out in the scope of the project "Rhythmic structure of utterances in the Polish language: A corpus analysis" supported by National Science Centre (NCN) grant no. 2013/11/D/HS2/04486.

7. Appendix

effect	**statistics**	**nPVI-V**	**rPVI-C**	**nPVI-syl**	**%V**	**deltaV**	**deltaC**	**VarcoV**	**VarcoC**
structure	η_p^2	.28	.92	.80	.86	*.042*	.91	.50	.94
	F(2, 53)	10.6	319	110	169	*1.16*	259	27.4	402
tempo	η_p^2	.11	.91	*.039*	.46	.76	.91	*.031*	*.038*
	F(5, 265)	6.6	518	*2.2*	46.7	171	536	*1.75*	*2.13*
tempo* structure	η_p^2	*.046*	.55	.12	.088	*.038*	.53	*.056*	*.062*
	F(10, 265)	*1.31*	32.3	3.82	2.63	*1.06*	30.9	*1.62*	*1.77*

Table 2: Summary of repeated measures ANOVA: phonotactic structure, tempo & their interaction. η_p^2 stands for partial ETA squared and expresses the size of the effect. Results at p > .05 are in italics.

effect	**statistics**	**nPVI-V**	**rPVI-C**	**nPVI-syl**	**%V**	**deltaV**	**deltaC**	**VarcoV**	**VarcoC**
elicitation	η_p^2	.021	.021	.0037	.045	.078	0.027	.027	0.019
	F(4, 3028)	16.8	16.3	2.87	36.8	64.5	21.6	21.2	15.4
elicitation* syl_rate	η_p^2	.074	.21	.030	.036	.23	.20	.028	.024
	F(5, 3028)	48.6	163	18.8	23.3	187	156	17.9	15.3

Table 3: The effect of elicitation method on rhythm metrics. The row "elicitation" shows results of ANOVA; the row "elicitation syl_rate" shows results of ANCOVA separate slopes model.*