Speech rhythm in native and non-native Polish

Agnieszka Wagner

Institute of Linguistics, Adam Mickiewicz University in Poznań, Poland

https://doi.org/10.36505/ExLing-2012/05/0030/000236

Abstract

This paper presents a study on the realization of speech rhythm in native and nonnative Polish. The paper examines the usefulness of $\% V-\Delta C$ and nPVI-rPVI metrics for rhythm analysis in native and non-native speech and their efficiency in capturing speaker idiosyncrasies in the realization of rhythm. The metric scores are also used to investigate the effect of speakers' L1 (German or Korean) on the realization of speech rhythm in Polish.

Key words: rhythm, metrics, native & non-native speech

Introduction

Traditionally, rhythm was defined as a production-based effect (Abercrombie 1965), but observations of overall durational variability of vocalic and consonantal intervals resulting from differences in the phonotactic structure supported by findings of empirical research led to a redefinition of this concept. Nowadays rhythm is considered as the result of interaction of a number of components: phonetic and phonological on the one hand and segmental and prosodic on the other (Dauer 1987). In the measurement of rhythm three sets of metrics are generally used: $\% V-\Delta C$ (Ramus et al. 1999), PVIs (Grabe & Low 2002) and Varcos (Dellwo 2006).

It can be expected that as in case of segmental errors (e.g. substitutions, insertions or deletions) the realization of rhythm in nonnative speech will also be influenced by the native language (L1) and transfers of timing patterns to the foreign language will occur. This kind of information can be used in pronunciation and prosody training in the foreign language e.g. with respect to curriculum preparation and foreign accent assessment. Another area where such knowledge might be valuable is speaker classification (e.g. for forensic purposes): here the specific realization of rhythm, along with other features, can be used to distinguish non-native from native speakers, to classify foreign accents or to identify particular speakers.

In the current study we try to answer the following questions: Do the %V- ΔC and PVI metrics provide a reliable basis for distinction between Koreanaccented, German-accented and native Polish? To which extent do the metrics capture speaker idiosyncrasies in the realization of rhythm? And finally, do the metric scores show the effect of speaker's L1 on the realization of rhythm in Polish?

ExLing 2012: Proceedings of 5th Tutorial and Research Workshop on Experimental Linguistics, 27-29 August 2012, Athens, Greece

Methodology

The speech material comes from speakers whose native language was Polish, German or Korean. German is a typical stress-timed language, whereas the status of Polish and Korean is unclear. Polish is sometimes classified as stress-timed (Ramus et. al 1999), but most often it is considered as rhythmically mixed: like syllable-timed Spanish or French it has no vowel reduction, but on the other hand, like stress-timed English or German, it can have very complex syllable structure. Rhythm in Korean is more syllable-timed than stress-timed (Arvaniti 2009, Mok & Lee 2008), but there is also evidence that Korean is one of the mora-timed languages (Cho Moon-Hwan 2004).

Speech material

The speech material includes recordings of a short literary story read by 7 German and 5 Korean learners of Polish and 5 Polish native speakers. The corpus (19 sentences x 17 speakers) was segmented into vocalic and consonantal intervals on the basis of automatically obtained and manually verified phonetic transcription and segmentation. Vocalic intervals were defined as stretches of signal between vowel onset and offset and consonantal intervals as stretches of signal between vowel offset and vowel onset. Segmentation, annotation and measurements were carried out in Wavesurfer and Praat.

Acoustic measurements

For each sentence 4 acoustic parameters were extracted from the duration measurements of the vocalic and consonantal intervals: %V – proportion of vocalic intervals, ΔC – standard deviation of consonantal intervals (Ramus et. al 1999), nPVI and rPVI – vocalic and consonantal pairwise variability indices (Grabe & Low 2002).

Results

The distribution of %V- Δ C and nPVI-rPVI metric scores calculated on the basis of native Polish (pl), German-accented (de) and Korean-accented Polish (kor) speech data is illustrated in Figure 1.

The vocalic nPVI is the same in German-accented and native Polish (46), which indicates that German speakers avoided vowel reduction (present in German, but absent in Polish). The rPVI value is higher in German-accented than in native Polish (75 vs. 70.6), but at the same time it is significantly lower than in German (55.3 in Grabe & Low 2002). This effect, also observed in Korean-accented Polish, can be attributed to complex structure of Polish syllables: difficulty in the realization of long consonant clusters causes greater variability in the timing of the consonantal intervals.



Figure 1. Distribution of metric scores in native and non-native Polish.

High nPVI value in Korean-accented Polish (about 60, i.e. the same as in native Korean, Mok & Lee 2008) suggests that the speakers transferred the timing of vocalic intervals from their L1 to Polish. Since vowels are more significant to the perception of rhythm than consonants, it can be concluded that concerning rhythm, Korean-accented Polish differs more from native Polish than German-accented Polish and that German speakers were more successful in the realization of rhythm in Polish then the Korean ones. This is in line with informal perceptual evaluation of rhythm in non-native Polish.

The distribution of %V- Δ C metrics reveals more variability in vocalic intervals' duration in Korean-accented than in native and German-accented Polish, but comparing to nPVI values the effect of speakers' L1 is smaller (see also ANOVA results, Table 1). The timing of consonantal intervals of German and Korean speakers is distinct (higher Δ C values) from that of Polish speakers, but the differences do not pattern in the expected direction (i.e. that of speaker's L1) – the explanation is the same as in case of rPVI values. ANOVA results (Table 1) show significant effect of *speaker's L1* on all the metric scores except for rPVI. %V and nPVI are affected by speaker's L1 to a greater extent than Δ C and rPVI. Pairwise comparisons based on Tukey HSD post-hoc test revealed that only %V is statistically different among the three speaker groups, therefore it can be considered the best predictor of speaker's L1.

metric score	%V	ΔC	nPVI	rPVI
(ANOVA)	(F=49.9,	(F=3.44,	(F=56.25,	(F=0.92,
	p<0.01)	p<0.01)	p<0.01)	p=0.4)
pairwise	all	pl x de	kor x pl	
comparison			kor x de	

Table 1. The effect of speaker's L1 on the metric scores.

124 A. Wagner

In order to test the robustness of the metrics in capturing speech rhythm differences between individual speakers ANOVA and Tukey HSD test were carried out using *speaker* as the predictor variable. The results showed statistically significant differences among speakers in %V and nPVI. Pairwise comparisons indicated that these differences are mainly due to speaker's L1 and reflect rhythmic variability among individual speakers to a limited extent. The scores form distinct groups according to speaker's L1, but in each group some statistically significant deviations in %V can also be found.

Discussion

The metrics capture differences in the realization of speech rhythm in German-accented, Korean-accented and native Polish and it is mainly vocalic interval duration variability that conveys these differences. $\% V-\Delta C$ are generally more robust in this task than PVIs and % V can also be used to analyze inter-speaker variability in rhythm. The metrics show transfer of some timing patterns from L1, but more importantly, they reflect differences in speech rhythm resulting from distinct phonotactic structure of the languages.

Acknowledgements

This work is supported from the financial resources for science in the years 2010-2012 as a development project (National Centre for Research and Development).

References

- Abercrombie, D. 1965. Studies in Phonetics and Linguistics. London: Oxford University Press.
- Arvaniti, A. 2009. Rhythm, timing and the timing of rhythm. Phonetica 66, 46-63.
- Dauer, R. 1987. Phonetic and phonological components of language rhythm. Proc. of the XIth Intern. Congress of Phonetic Sciences, 447-450, Tallinn, Estonia.
- Dellwo, V. 2006. Rhythm and speech rate: A variation coefficient for deltaC. In Karnowski, Szigeti (Eds.) Language and language processing: Proc. of the 38th linguistic colloquium, 231–241, Peter Lang, Frankfurt.
- Grabe, E. & Low, E.L. 2002. Durational variability in speech and the rhythm class hypothesis. In Gussenhoven, C. & Warner, N. (Eds.) Laboratory Phonology 7, 515-546, Mouton de Gruyter, Berlin.
- Mok, P. & Lee, S.I. 2008. Korean speech rhythm using rhythmic measures. Proc. of the 18th Intern. Congress of Linguists (CIL18), Seoul, Korea.
- Moon-Hwan, Ch. 2004. Rhythm typology of Korean speech. Cognitive Processing 5(4), 249-253.
- Ramus, F., Nespor, M. & Mehler, J. 1999. Correlates of linguistic rhythm in the speech signal. Cognition 72, 1-28.