

How far can phonological properties explain rhythm measures?

Elinor Keane, Anastassia Loukina, Greg Kochanski, Burton Rosner

Oxford University Phonetics Laboratory, UK

Chilin Shih

EALC/Linguistics, University of Illinois, Urbana-Champaign, USA

This is the abstract for a talk (presented by Elinor Keane) at the 2010 meeting of the British Association of Academic Phoneticians, BAAP 2010, London, UK. 28-31 March 2010. It can be downloaded from <http://kochanski.org/gpk/papers/2010/BAAP/Keane>.

Speech rhythm has long been thought to reflect the phonological structure of a language (e.g., Roach 1982; Dauer 1983, 1987). Syllable structure is a key example: languages that allow complex consonant clusters would have a rhythm characterized by much more variability in consonant length than a language like Mandarin where consonant clusters are rare. We explored this experimentally by seeing how well a range of popular rhythm measures were predicted by the phonological properties of the text.

The results are based on 3059 paragraphs read by 62 native speakers of English, Greek, French, Russian and Mandarin. The paragraphs were selected from the novel *Harry Potter and the Chamber of Secrets*, to represent the full range of phonological variation existing in each language. They included pairs of paragraphs chosen for particularly high and particularly low values of eleven different phonological properties. These were calculated from the expected transcription and included the average complexity of consonant clusters, percentage of diphthongs in the text and average sonority (assigning a sonority level of 0 to obstruents, 1 to sonorants and 2 to vowels).

First, we confirmed that languages indeed have different phonotactics, based on the expected transcription. For example, the complexity of consonant clusters in the English data was significantly greater than in the Mandarin data. A classifier based on a pair of averaged phonological properties (e.g. mean consonant cluster length and mean sonority) would correctly identify the language of 70% to 87% of the paragraphs (1Q-3Q range, depending on the pair of properties, chance=20%).

The recorded speech was divided into vowel-like and consonant-like segments using a language-independent automatic segmenter, trained on all five languages. From this, we computed 15 statistical indices proposed as rhythm measures in the literature, e.g. %V, VnPVI (references in Loukina et al. 2009): all were devised to capture durational variability between languages. In contrast to the classifiers based on phonological properties, we found large overlap between languages.

Phonological properties were found to predict paragraph-to-paragraph differences in rhythm measures rather poorly. The largest correlations involved the percentage of vowel-like segments in speech vs. the percentage of voiced segments in text, but these only explained 9% of the variance in Russian and 18% in Mandarin. Instead, interspeaker differences accounted for much more of the variation in the rhythm measures in a linear regression analysis. For example, for Russian, the average adjusted r^2 across different rhythm measures was .112 for regressions against phonological properties, but .295 for regressions against speakers. The corresponding values for English were .139 and .335.

These results indicate that differences in timing strategies between speakers, even within the same language, are at least twice as important as the average phonological properties of the paragraph. It suggests that rhythm, in the sense of durational variability, is determined more by performance differences between individuals than differences in the phonological structure of languages.