THE ROLE OF RHYTHMIC STRESS IN THE PERCEPTION OF WORDS

Anna Łukaszewicz[1], Beata Łukaszewicz[1], Janina Mołczanow[1]

[1]University of Warsaw
a.b.lukaszewicz@uw.edu.pl, b.lukaszewicz@uw.edu.pl, jmolczanow@uw.edu.pl

ABSTRACT

Polish has a grammatical stress system with two degrees of stress. Apart from the main stress falling regularly on the penult, it has iterative rhythmic stresses applying from the left word edge. Previous studies show that rhythmic stresses are optional and acoustically weak. The functionality of rhythmic stress is unclear. In this paper, we test the hypothesis that rhythmic stress facilitates word segmentation: despite being phonologically redundant, it may have a differential effect on making lexical decisions. We report a perceptual experiment, in which we investigated whether rhythmically congruent and incongruent lexical items vs. segmentally matched nonce words would trigger different response times (RTs) in a word/non-word decision task. The results point to a statistically significant interaction between wordhood and rhythm. Faster RTs were obtained for rhythmically congruent lexical items compared to rhythmically incongruent ones. For nonwords, the pattern was reversed. Rhythmic structure is thus associated with word-likeness.

Keywords: perception, Polish, rhythm, RTs, stress

1. INTRODUCTION

This paper investigates the relationship between phonologically redundant rhythmic (secondary) stress and lexical processing, on the basis of behavioural data (response times; RTs) obtained in a perception experiment involving a word/non-word decision task. While it is clear that prosody plays an important role in speech recognition, facilitating utterance parsing and word comprehension (e.g. see [1] and the literature cited therein), it is not obvious whether the optional alternation of rhythmic beats and metrically weak syllables, as the one occurring in iterative secondary stress systems, plays a similar role. We focus here on rhythmic stress in Polish (the so-called Warsaw and Cracow standard varieties), which is a well-known example of an iterative stress system in the phonological literature (e.g. [2], [3], [4], [5], [6]). Rhythmic stress in Polish is fully predictable and hence phonologically redundant. It is also known to be optional, which prompts questions concerning its functionality.

Little is known about the perception of rhythmic stress in Polish. To our knowledge, the only perception experiment conducted thus far was the one reported in [7]; it revealed no sensitivity of native speakers of Polish to the presence of rhythmic stress. On the one hand, that negative result could be associated with rhythmic stress being acoustically weak. (The lack of acoustic robustness of subsidiary degrees of stress has often been raised as an issue cross-linguistically, putting into question the physical reality of the phenomenon in general [8].) On the other hand, the method used in that experiment was based on the participants' transcription of stress, and thus it relied heavily on introspection. Here we report an experiment based on a standard behavioural (RT) paradigm, which circumvents those methodological issues. The hypothesis that we put forward is that, though acoustically weak and phonologically redundant, rhythmic stress affects word segmentation and lexical processing, facilitating word recognition.

The Polish rhythmic stress pattern is illustrated below. It is a well-documented phenomenon. Basic generalizations are found, among others, in an early kymographic study by Dłuska, published in 1932 ([2]), traditional descriptions of the Polish prosody (such as [9]), phonological analyses couched in various theoretical frameworks (the metrical grid in [3]; optimality theory in [4]), as well as in recent studies of acoustic correlates of rhythmic stress ([5], [6], [10], [11]). As can be seen in the examples adduced below, Polish has primary stress on the penult and an iteration of secondary (rhythmic) beats on every odd-numbered syllable. Important from the point of view of the current study is that the iterative pattern, which begins at the left edge of the word, is [ˌσσ] (stressed – unstressed), and not (σˌσ) (unstressed – stressed). As clarified below, this observation is the basis for the rhythmically congruent and rhythmically incongruent stimuli in the experiment we report here.

- *poˈmidor* ('tomato' N., nom. sg.)
- *ˌpomiˈdora* ('tomato' N., gen. sg.)
- *ˌpomidoˈrowy* ('tomato' Adj., nom. sg.)
- *ˌpomiˌdoroˈwego* ('tomato' Adj., gen. sg.)

In what follows, we describe the details of a rhythmic stress perception experiment involving RTs, which we conducted for the purpose of the present study. We test whether native speakers of Polish are sensitive to a change in the temporal structure across the first two syllables in a word, resulting in prominence reversal, [ˌσσ] (rhythmically congruent stimuli) > [σˌσ] (rhythmically incongruent stimuli), and whether this has a differential effect on the word recognition process.

2. METHOD

Overall, the design of our RT experiment was based on a classic stress perception study [1], the main difference being that the semantically congruent vs. incongruent condition in [1] is replaced by the word vs. nonword condition in our experiment. Analogously to [1], we aimed to determine whether the metrical and lexical aspects of spoken language are processed independently or in interaction. Those aspects were thus manipulated orthogonally, so as to create four conditions. Words and nonwords appeared in the rhythmically congruent vs. rhythmically incongruent versions, forming segmentally matched quadruplets, described in more detail in 2.1 below.

2.1. Stimuli

Overall, the stimuli consisted of 128 tokens (32 lexical items and 32 nonce words, both in the rhythmically congruent and rhythmically incongruent versions. The stimuli can thus be described as having the following structure:

- W+R+, a rhythmically congruent word
- W+R– , a rhythmically incongruent word
- W-R+, a rhythmically congruent nonword
- W–R–, a rhythmically incongruent nonword

2.1.1. Lexical vs. nonce stimuli

All target stimuli consisted of five syllables: the first three syllables constituted the (lexical or nonce) ‘stem’; the remaining two included an inflectional ending. To ensure maximal segmental and prosodic comparability between lexical and nonce items, in both the last two syllables contained the same inflectional ending *-ami* (instr. pl.). The ending contains the penultimate syllable of the word, which bears primary stress.

The lexical items and the nonce words were segmentally matched: the nonce word differed from its lexical counterpart only in the consonant occurring at the beginning of the third syllable; cf. *czekoladkami* [t͡ʂɛkɔlatkami] ‘chocolates’ (instr. pl.) – *czekoszadkami* [t͡ʂɛkɔʂatkami] (nonce word). Another such pair (out of the total of 32 paired words and nonwords) is illustrated in Table 1. It is well-known that phonotactic probability influences on-line processes and has an impact on how quickly words are recognised on the basis of fluent speech (e.g. [12], [13], [14]). The lexical items were thus frequency matched on the basis of the balanced subsections of the National Corpus of Polish (Narodowy Korpus Języka Polskiego). The corresponding nonce words were created in such a way as to avoid an overlap between their stems and existing lexical stems. (Because Polish is a highly inflectional language, for all items lemmatic searches were conducted, i.e. all possible grammatical forms were taken into account.)

Word	Nonce
samochod+ami	samoziod+ami
‘car’ (instr.pl)	
/samɔxɔdami/	/samɔʑɔdami/

Table 1: An example of a lexical item - nonce word pair in the stimuli set.

2.1.2. Rhythmically congruent/incongruent stimuli

All word and nonword stimuli were prepared in two versions: rhythmically congruent and rhythmically incongruent. The rhythmically congruent version was the one agreeing with the secondary stress pattern of Polish, i.e. with the initial syllable having enhanced duration relative to the second syllable. For the purpose of obtaining such stimuli, target words and nonwords were recorded by a phonetically trained native speaker of the standard variety of Polish (the so-called Warsaw Polish). This variety of Polish is known to exhibit the required pattern (e.g. [5], [6]). The target stimuli were incorporated in a frame, which was the same for all items. (*Powiedziała __ po raz pierwszy*, ‘She said __ for the first time’.) The recordings were conducted in a sound-treated booth.

The rhythmically incongruent counterparts were obtained by changing the relative duration of the first and second syllables. Care was taken to change the lengths within the ecological bounds. To this end, the original recordings were time-compressed by the 0.7 factor to obtain the initial syllable and time-stretched by the 1.4 factor to obtain the second syllable, using the ‘overlap-add’ algorithm in *Praat* (v. 6.2.09; [15]). The relevant syllables were then extracted from the continuous acoustic signal using a high-resolution waveform editor (Sound Forge). This allowed to align the boundaries of vowels with zero crossings, so that each excised segment contained a complete number of cycles (e.g. [5], [6], [16], [17]). (This was necessary to offset a potential auditory effect of the resulting speech signal being artificially distorted.) The manipulated syllables were then spliced to replace the two initial syllables in the original sound files, using a *Praat* script. (We did not use a single manipulated version as the lexical-nonce pairs differed in the third syllable's onset, which has an effect on the second syllable's vowel formants.) All items were placed within the same frame – a single token of the frame was selected, such that the formant structure of the final vowel of *Powiedziała* was least affected by the following sound. Fig. 1 shows an example of the original version, a rhythmically congruent pattern, juxtaposed against the manipulated version, a rhythmically incongruent pattern of the word *balonikami* /balɔɲiˈkami/ 'balloon' (Dimin. instr. pl.) The first syllable (*ba*) is longer, and the second syllable (*lo*) is shorter in the rhythmically congruent item (top panel); the pattern is reversed in the rhythmically incongruent item (bottom panel).

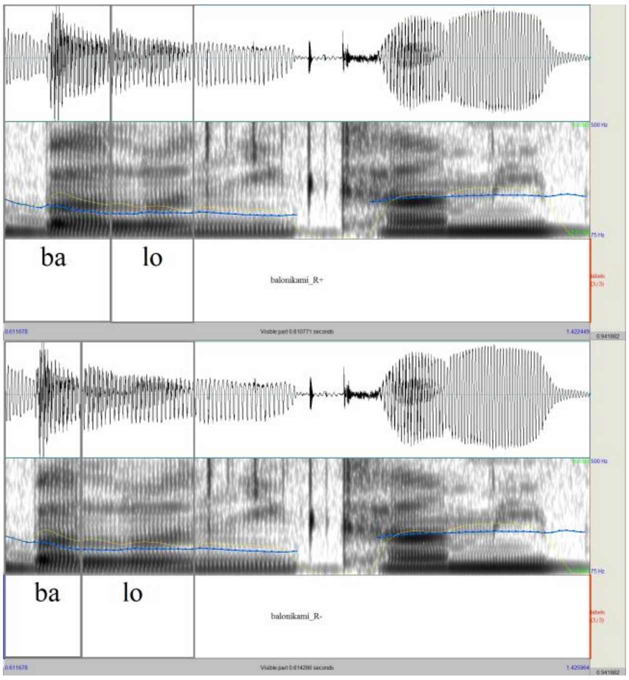

Figure 1: The rhythmically congruent (original) version (top panel) and the rhythmically incongruent version (bottom panel) of the same word.

2.2. Participants

24 (16F) participants took part in the pilot study that we report here. All of them were native speakers of Polish, living in Warsaw, with no reported history of speech or hearing impairment. The mean age was 25.6 (age range 19-41). Most speakers were right-handed (20 out of 24). As pointed out e.g. in [5] and [6], apart from being optional and weakly expressed in acoustic terms, rhythmic stress may also be subject to regional variation. It was thus necessary to take the dialectal factor into account and to base the experiment on the dialects known to exhibit the rhythmic pattern (the so-called Warsaw or Cracow Polish). As required for research involving human subjects, the protocol of the experiment was approved by the Committee for the Ethics of Research Involving Human Participants at University of Warsaw (131/2022).

2.3. Technical apparatus

The experiment was programmed and presented using *E-Prime 3.0*. RTs were measured with the Chronos device. The stimuli were recorded in a sound-treated booth at the sampling frequency of 44.1 kHz, 32-bit depth, using an AT897 microphone. They were manipulated using *Praat* (v. 6.2.09; [15]) and Sound Forge (a high-resolution waveform editor). The participants were wearing closed headphones (ATH-M50x) throughout the experiment. The statistics were conducted in *SPSS* (v. 28).

2.4. Procedure

The participants' task was to press, as quickly as possible, the extreme right or extreme left button on the Chronos panel, depending on whether the stimulus they heard was a word or a nonword. For half of the participants, the extreme right and left buttons corresponded to the choice "word" and "nonword", respectively. For the other half, the buttons were reversed. The experiment was preceded by a short training session, during which the participants could practice pressing the buttons in response to a set of stimuli designed analogously to those appearing later in the experiment. As in [1], each participant heard only one item from a given word – nonword – rhythmic – nonrhythmic quadruplet during the experiment. In this way, we avoided a potential bias created by the participant's being repeatedly exposed to the same initial sequences. The stimuli were balanced across the sessions and were randomized within each session, using the "random order" function in *E-Prime*; all stimuli were equally represented in the experiment.

3. RESULTS

RTs were measured from the onset of the auditory stimulus. The target item began at 0.61 s (which was the fixed length of the initial part of the frame). Only accurate responses were submitted to the statistical analysis. The accuracy was 95.7 %. The statistics were obtained through a fitted generalized mixed effects model with the Gamma probability distribution and the log link function. The model included 'participant' as the random effect. The fixed effects were: wordhood (word/nonword), rhythmicity (congruent/incongruent), the interaction term rhythmicity*wordhood, and the trial number. The dependent variable was RT. We also included the 'word length' as a covariate. Most factors were significant (we report the F-statistics for brevity); wordhood: $F(1, 698) = 15.8$, $p < 0.001$; rhythmicity*wordhood: $F(1, 698) = 26.5$, $p < 0.001$; trial: $F(31, 698) = 3.9$, $p < 0.001$; word length $F(1, 698) = 4.2$, $p = 0.041$). The only effect that turned out insignificant was 'rhythm'. The estimated RT means (in [ms]) depending on rhythmic congruence and wordhood are graphed in Fig. 2. The interaction between the two factors is clear: rhythmicity has the opposite effect for words and nonwords.

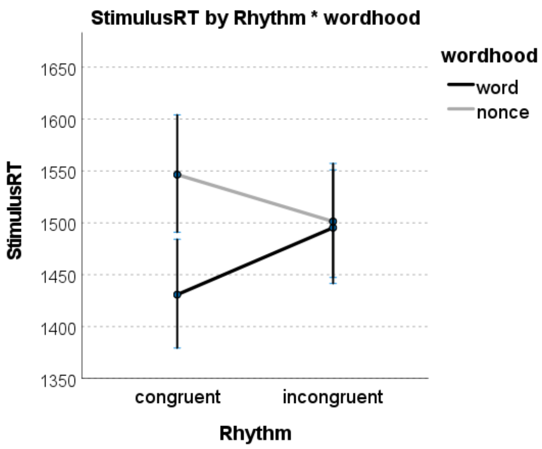

Figure 2: Estimated marginal means for RTs depending on rhythm and wordhood.

4. DISCUSSION

Our results corroborate the hypothesis that rhythmic stress in Polish, though phonologically redundant and acoustically weak, plays an essential role in lexical access and word recognition. The rhythmic structure has a crucial effect on the word-likeness. Rhythmically congruent words are recognised faster than rhythmically incongruent words. The word-likeness judgements of nonwords also depend on the rhythmic factor: rhythmically incongruent nonwords trigger faster decisions. In sum, wordhood and rhythmicity seem to go hand in hand: the recognition of a speech chunk as a word is facilitated if it is well-formed from the point of view of the rhythmic pattern.

This result is important because it is based on rhythmic structure which is entirely redundant from a phonological point of view. Interestingly, while wordhood and the interaction between rhythmicity and wordhood turned out to be significant factors in our analysis, rhythmicity appeared non-significant as the main effect. However, it is by no means a negligible factor in language processing. Our results concerning the role of rhythmic structure corroborate its influence on on-line processing and word recognition, parallel to the one exhibited by sublexical segmental structures (phonotactic probability; e.g. [13]).

Finally, it is important to add that the rhythmic stress in Polish has sometimes been reported to be difficult to detect in the acoustic speech signal ([18]; see also [6] and the literature cited therein), which likely depended on the dialect choice and methodological questions (as discussed thoroughly in [6]). Nevertheless, its being phonologically optional and acoustically weak poses an intriguing question about whether it is perceptible and what its functionality can be. In some previous work (e.g. [3]), the presence of rhythmic stress was associated with slow careful speech. A recent experimental study (11) has demonstrated no such tendency in hyperarticulated speech produced at slower rates (Lombard speech). This would suggest that the presence of rhythmic stress is more important from the point of view of perception than production. The only previous perception study that we know of (i.e. [7]) pointed to the lack of awareness of the presence of rhythmic stress in native speakers of Polish, based on introspection and explicit judgement. Our behavioural experiment, based on a standard RT paradigm, points to the association between rhythmic stress and word-likeness. Native speakers of Polish may not be aware of the rhythmic stress pattern, but its presence has an effect on word segmentation and recognition.

5. CONTRIBUTIONS

Initial conceptualization: AŁ, BŁ, JM; experiment design and preparation of stimuli: AŁ, BŁ; E-prime programming: AŁ; participant recruitment: JM; conducting the experiment: AŁ, BŁ, JM; statistical analyses: BŁ; writing the paper: AŁ, BŁ, JM; revisions: AŁ, BŁ.

6. ACKNOWLEDGEMENTS

We would like to thank two anonymous reviewers for their questions and comments. Research was funded by the National Science Centre (Poland), grant 2017/27/B/HS2/00780.

7. COMPETING INTERESTS

The authors declare no conflict of interest.

8. REFERENCES

[1] Magne, C., Astésano, C., Aramaki, M., Ystad, S., Kronland-Martinet, R., Besson, M. 2007. Influence of

Syllabic Lengthening on Semantic Processing in Spoken French: Behavioral and Electrophysiological Evidence. *Cerebral Cortex* 17, 2659-2668.

[2] Dłuska, M. 1932. *Rytm spółgłoskowy polskich grup akcentowych.* Polska Akademia Umiejętności.

[3] Rubach, J., Booij, G. E. 1985. A grid theory of stress in Polish. *Lingua* 66, 281–319.

[4] Kraska-Szlenk, I. 2003. *The phonology of stress in Polish.* Lincom.

[5] Łukaszewicz, B. 2015. Polish stress revisited: Phonetic evidence of an iterative system. Paper presented at 23rd Manchester Phonology Meeting.

[6] Łukaszewicz, B. 2018. Phonetic evidence for an iterative stress system: The issue of consonantal rhythm. *Phonology* 35, 115–150.

[7] Steffen-Batogowa, M. 2000. *Struktura akcentowa języka polskiego*. Wydawnictwo Naukowe PWN.

[8] Tabain, M., Fletcher, J., Butcher, A. 2014. Lexical stress in Pitjantjatjara. *Journal of Phonetics* 42, 52–66.

[9] Dłuska, M. 1974. *Prozodia języka polskiego.* Państwowe Wydawnictwo Naukowe.

[10] Łukaszewicz, B., Zajbt, E., Krawczyk, U. 2018. The rhythm of heptasyllabic words: Evidence for metrical bidirectionality. *Speech Prosody* 9, 676‑679.

[11] Łukaszewicz, B., Zajbt, E., Mołczanow, J. 2020. Polish iterative stress and its phonetic parameters in quiet vs. noisy environments. *Lingua* 240, 102835.

[12] Pitt, M. A., Samuel, A. G. 1995. Lexical and sublexical feedback in auditory word recognition. *Cognitive Psychology* 29, 149–188.

[13] Vitevitch, M. S., Luce, P. A. 1999. Probabilistic phonotactics and spoken word recognition. *Journal of Memory and Language* 40, 374–408.

[14] Vitevitch, M. S., Luce, P. A. 2003. The influence of sublexical and lexical representations on the processing of spoken words in English. *Clinical Linguistic Phonetics* 17, 487–499.

[15] Boersma, P., Weenink, D. 2022. Praat: doing phonetics by computer. http://www.praat.org/. (v. 6.2.09)

[16] Łukaszewicz, B., Rozborski, B. 2008. Korelaty akustyczne akcentu wyrazowego w języku polskim dorosłych i dzieci. *Prace Filologiczne* 54, 265–283.

[17] Łukaszewicz, B., Rozborski, B. 2008. Acoustic correlates of word stress in child and adult Polish and the acquisition of phonological rhythm. Paper presented at Generative Linguistics in Poland (GLiP-6), Warsaw.

[18] Dogil, G. 1999. The phonetic manifestation of word stress in Lithuanian, Polish, German and Spanish. In: van der Hulst, H. (ed.), *Word prosodic systems of the languages of Europe.* Mouton de Gruyter, 273–311.