

Squibs and replies

*Phonetic evidence for an iterative stress system: the issue of consonantal rhythm**

Beata Łukaszewicz
University of Warsaw

In her study published in this journal, Newlin-Łukowicz (2012) calls into question the existence of bidirectional stress systems. The argument hinges on the failure to detect acoustic correlates of word-internal subsidiary stress in Polish, the language hitherto considered to be a classic example of metrical bidirectionality. This paper reappraises the issue, reporting on an acoustic study of paired five- and six-syllable words in Polish (e.g. *pomido'rozwy* – *pomi'doro'wego*). The results indicate that the words differ significantly with respect to relative consonant duration (PVI values) in the onset of the third syllable, depending on whether the syllable bears subsidiary stress (as in six-syllable words) or remains unstressed (as in five-syllable words). Similar effects are reported in the initial syllable, but not in the second syllable, which remains consistently unstressed. The conclusion is that Polish has iterative stress, corroborating its traditional description as having a bidirectional stress system.

1 Introduction

In the phonological literature, Polish has been described as having a bidirectional trochaic stress system, with main stress on the penult, and iteration of subsidiary stresses from the left, as illustrated in (1).

* E-mail: B.LUKASZEWICZ@UW.EDU.PL.

I wish to thank the three anonymous reviewers and editors of *Phonology* for discussion and criticism, which were very helpful in improving the paper. I am also grateful to Janina Mołczanow for her comments on the first version of the paper. Thanks are also due to the audiences at the 23rd Manchester Phonology Meeting and *Phonetics and Phonology in Europe 2015* at the University of Cambridge, where preliminary results were presented. I would also like to thank the Polish speakers who took part in the experiment. This research was supported by the National Science Centre, Poland (grant 2015/17/B/HS2/01455). All errors are mine.

(1) *Foot parsing in Polish: standard view*

- | | | | |
|----|---|---|------------------------------------|
| a. | [(_i σσ)(ⁱ σσ)] | (_i prenu)(ⁱ merat) | ‘subscription (GEN PL)’ |
| b. | [(_i σσ)σ(ⁱ σσ)] | (_i prenu)me(ⁱ rata) | ‘subscription (NOM SG)’ |
| c. | [(_i σσ)(_i σσ)(ⁱ σσ)] | (_i prenu)(_i mero)(ⁱ wany) | ‘subscribed (NOM SG
PAST PART)’ |
| d. | [(_i σσ)(_i σσ)σ(ⁱ σσ)] | (_i prenu)(_i mero)wa(ⁱ nego) | ‘subscribed (GEN SG
PAST PART)’ |

Even-parity words are parsed exhaustively into feet, as in (a) and (c). Odd-parity words exhibit lapses adjacent to the penult, as in (b) and (d). The iterative nature of subsidiary stress can be seen in (c) and (d), and the bidirectional character of the stress system can be inferred from (b) and (d). These characteristics have been widely reported, and are explored both in traditional descriptions and theoretically oriented studies (e.g. Dłuska 1974, Rubach & Booij 1985, Halle & Vergnaud 1987, McCarthy & Prince 1993, Kraska-Szlenk 2003, McCarthy 2003). A lapse adjacent to primary stress in odd-parity words is in accordance with Kager’s (2001) claim that subsidiary stresses in bidirectional systems iterate towards the peak, not towards the opposite end.

The iterative nature of Polish stress is questioned by Newlin-Łukowicz (2012); in an acoustic study only one level of prominence – main stress on the penult – was detected in polysyllabic words containing a single lexical root, as well as in certain compounds.¹ Some prominence effects were also found in word-initial position, but are interpreted as ‘boundary effects’ caused by an amplitude-declination pattern unconnected with stress (2012: 301–302).

An important caveat is that Newlin-Łukowicz’s study of single-root words is based on four-, five- and six-syllable words, like the examples in (1a–c). Seven-syllable words such as (1d) are not included. In addition, interpreting word-initial prominence as a word-edge effect makes four- and five-syllable words uninformative with regard to the iterative (or bidirectional) characteristic of the stress pattern, because subsidiary prominence invariably coincides with word-initial position in these words. Thus the argument against the iterative (and, by implication, also bidirectional) character of the Polish stress system hinges solely on the apparent absence of acoustic markers of word-internal subsidiary (‘tertiary’) stress in six-syllable words. If the third syllable in six-syllable words is defined

¹ Newlin-Łukowicz found that compounds exhibit two stress patterns (2012: 317ff). Some are structures with two prosodic words, exhibiting two stresses corresponding to the canonical position of primary stress in Polish (e.g. *zielono-beżowy* [zɛ(ˈlɔnɔ)]_ω [bɛ(ˈʒɔvʲi)]_ω ‘green-beige (NOM SG MASC)’). Others are single prosodic words, exhibiting only one stress (e.g. *nowomodny* [nɔvɔ-(ˈmɔdnʲi)]_ω ‘fashionable (NOM SG MASC)’). The latter are claimed to essentially mirror the stress pattern of single-root words of the same length. The analysis of compounds is further complicated by such factors as lexicalisation and lexical frequency effects (2012: 321). In this study, I leave compounds aside, and focus on single-root words, which are reported to pattern uniformly, and for which only some of the methodological problems posed by compounds are relevant.

as acoustically weak, and word-initial prominence is not interpreted as metrical prominence, all the structures in (1) are left unparsed, except for the rightmost feet carrying primary stresses, as shown in (2).

(2) *Foot parsing in Polish: Newlin-Łukowicz (2012)*

- a. $[\sigma\sigma(\sigma\sigma)]$ prenu('merat)
- b. $[\sigma\sigma\sigma(\sigma\sigma)]$ prenume('rata)
- c. $[\sigma\sigma\sigma\sigma(\sigma\sigma)]$ prenumero('wany)
- d. $[\sigma\sigma\sigma\sigma\sigma(\sigma\sigma)]$ prenumerowa('nego)

This has far-reaching typological and theoretical consequences. Not only does it call into question previous phonological descriptions and analyses of the Polish metrical system, but it also casts doubt on the existence of bidirectional stress systems in general. This is because Polish is considered to be 'the only uncontested example of a bidirectional system with internal lapses' (Newlin-Łukowicz 2012: 271). If bidirectional systems with lapses are illusory, the theoretical tools developed to account for such systems must also be inadequate.

There are several reasons to doubt the validity of the acoustic results on which the above conclusion was based. The results were obtained with methods used in investigating both primary and secondary stress, on the assumption that primary and secondary prominence are two degrees of the same phenomenon, cued by the same (sub)set of acoustic parameters.² These methods, based on raw (or standardised, i.e. log-transformed) measurements of vowel parameters (F0, intensity, duration) and powerful inferential statistics (such as linear mixed-effects modelling), can produce robust and meaningful results; see for example Plag *et al.* (2011) for an acoustic study of primary and secondary stress in English. However, they are not universal, and may not be applicable to all subsidiary stress data. This becomes immediately apparent when some earlier reported characteristics of the Polish rhythm are taken into account. First, in the case of Polish, focusing on vowel parameters can hinder detection of a pattern which, according to an empirical study by Dłuska (1932), hinges mostly on the duration of onset consonants. Second, segments and syllables in longer words tend to be produced at a higher speech rate in Polish (e.g. Dłuska 1932: 9). Thus, in raw terms, the third syllable of a six-syllable word (bearing tertiary stress) will not necessarily be longer than the third (unstressed) syllable of a five-syllable word. However, we do expect the relative lengths of the third and the second syllables to be significantly different in six-syllable *vs.* five-syllable words, depending on whether the third syllable is stressed or unstressed. This suggests a different approach, based on relative (or normalised) measures,³ akin to

² Acoustic studies of secondary stress are rare. In the majority of acoustic studies of stress, the focus is on drawing a distinction between stressed and unstressed syllables, ignoring lower degrees of stress.

³ To clarify, in the case of intensity and pitch, the mere change from a linear to a logarithmic scale does not create relative measures that would allow for direct comparison

existing rhythm metrics. Third, rhythmic stress in Polish is commonly described as optional; it can be omitted in fast speech (Dłuska 1974: 27, Rubach & Booij 1985: 284). This makes its detection even more difficult, as it can reduce the average difference between stressed and unstressed positions to a negligible level. In addition to the methodological concerns expressed above, there is yet another potential factor that can make the detection of subsidiary stress in Polish impossible. Standard Polish is known to exhibit subtle phonological variation, depending on location. Dłuska's (1932) study of 'consonantal rhythm' is based on data from Warsaw and Krakow Polish (central and southern Poland); speech from other regions was deliberately excluded. In Newlin-Łukowicz's (2012) study, eight speakers were natives of Bytów, a small town near Gdańsk (northern Poland), and the remaining two were lifelong residents of Poznań (western Poland). It is impossible to assess the weight of the regional factor without prior research on potential regional variation in metrical patterns in contemporary Polish. However, such research, although interesting from the point of view of Polish dialectology, is irrelevant from the point of view of the issue at hand, i.e. whether Polish has lower degrees of stress, as claimed in traditional studies. It can be settled simply by looking at data from Warsaw and/or Krakow Polish, because these are the variants on which the traditional claims were made.

In the light of these potential problems, in this paper I reconsider the issue of the phonetic grounding of tertiary stress in Polish. I report on an acoustic study of 68 paired five- and six-syllable words, e.g. *pomidorowy* 'tomato (NOM SG ADJ)' and *pomidorowego* 'tomato (GEN SG ADJ)', collected in a word-list reading task from ten native speakers living in Warsaw. The results indicate that five- and six-syllable words differ significantly with respect to relative consonant duration (expressed in terms of the Pairwise Variability Index (PVI); e.g. Low *et al.* 2000, Ballard *et al.* 2010, Arciuli *et al.* 2014) in the onset of the third syllable, depending on whether the syllable bears tertiary stress (as in six-syllable words) or remains unstressed (as in five-syllable words). The statistical results obtained through linear mixed-effects models indicate that the consonant is significantly longer with respect to the preceding vowel in six-syllable words. Importantly, no similar effect is reported in the onset of the second syllable, which remains consistently unstressed. A comparison of the onsets of the third and second syllables (in terms of consonantal

across speakers and tokens, but only reflects the fact that perception does not work linearly (the Weber-Fechner Law). Relative measures are ratios with a raw value of the parameter in the numerator and the 'local' mean (rather than the absolute threshold value) in the denominator. The local mean can be the mean value of a given parameter within a word, a phrase or a sequence of adjacent units, depending on the domain in which prominence relations are to be captured (cf. Beckman 1986, Levi 2005, Łukaszewicz & Rozborski 2008, Malisz & Wagner 2011–12, Arciuli *et al.* 2014). Needless to say, the calculation of relative measures for intensity and F0 normally also involves standardisation of the ratios in decibels and semitones respectively.

PVIs) in five- and six-syllable words further confirms the hypothesis of relative lengthening of consonants in the stressed position. In addition, the investigation of vowel parameters – intensity, F0 and duration – indicates that the third vowel does not exhibit statistically significant differences depending on stress. Finally, overall comparisons of consonantal and vocalic PVIs in three conditions (secondary stress, tertiary stress, no stress) further corroborate the hypothesis that Polish has consonantal rhythm.

The above results have clear implications for phonological theory. Contrary to Newlin-Łukowicz's claim, the Polish stress system is iterative, which makes the argument against bidirectional systems invalid. Interestingly, primary stress and subsidiary stress need not be signalled by the same set (or hierarchy) of acoustic parameters, which may be relevant from the point of view of a metrical theory which distinguishes lexical accent from rhythmic beats (cf. van der Hulst 2014).

This paper is organised as follows. §2 discusses the methods and findings of previous acoustic studies of subsidiary stress in Polish, focusing on the choice of acoustic cues and the issue of data comparability. §3 states the hypotheses tested in this study, and describes the methodology of the experiment. §4 presents and discusses statistical results, and §5 considers theoretical implications of the acoustic analysis and briefly addresses some methodological issues which are relevant from the point of view of defining lower degrees of stress in acoustic terms. §6 summarises the conclusions and implications for future research.

2 Previous instrumental studies

The trochaic rhythm of Polish polysyllabic words, with a lapse immediately preceding the penult (primary stress) in odd-parity words, has been extensively explored in the phonological literature (e.g. Dłuska 1974, Rubach & Booij 1985, Kraska-Szlenk 2003), but there have been few instrumental studies directly addressing its phonetic underpinnings. The few sources available include Dłuska (1932), Dogil (1999) and Newlin-Łukowicz (2012). A striking observation about the previous research is that, despite the obvious advance in instrumental techniques and statistical methods from which more recent studies benefit, the only successful detection of acoustic markers of subsidiary prominence is reported in Dłuska (1932), an early empirical study using a crude kymograph and simple manual calculations. Importantly, independent of the instrumental apparatus, there are two prime methodological issues that differentiate Dłuska (1932) from the more recent studies: (i) the choice of potential acoustic cue(s) to secondary stress, and (ii) the issue of data comparability (i.e. making allowance for the relative rather than the absolute nature of stress). Because there is no doubt that the success or failure of a particular acoustic analysis may be determined by such initial methodological assumptions, the aim of this section is not only to present previous

findings on Polish subsidiary stress, but also to pinpoint potentially relevant methodological differences among the studies. A detailed comparison of different approaches provides the background for the methodological choices in the present study.

Dłuska (1932) reports on an optional pattern of alternating stresses in polysyllabic words, with syllable prominence marked by greater duration of the consonant in onset position, determined by a kymographic analysis. Duration measurements were based on more than 500 words of different length, occurring in different sentences, and read by five native speakers of the Warsaw or Krakow standard. No information is given about the exact number of tokens, but the majority of items were produced by two informants, who read more than 100 sentences, repeating each sentence at least twice. The sentences were recorded on a kymograph, with the time scale established by a tuning fork vibrating at 100 Hz. Measurements were accurate to 5 ms (corresponding to a half-cycle of the tuning fork). On average, the differences between the stressed and unstressed onset consonants were reported to be between 10 and 40 ms (1932: 71). Raw duration values were compared directly within a single token. On the basis of these comparisons, general observations were made about the relationship between consonant duration and prosodic position. The results identified two positions to which prosodic prominence was invariably anchored in Polish: initial position (secondary stress) and penultimate position (primary stress) (1932: 21). In addition, lengthening of the onset of the third syllable was observed, especially when the syllable was not immediately followed by main stress. We may thus infer that lengthening in the third syllable occurred in six-syllable and longer words, but not in five-syllable words.

What seems to be of vital importance is not only the choice of consonantal duration as a potential acoustic cue to stress, but also the type of consonantal material selected for comparison. Dłuska's study compared identical consonants occurring within the same prosodic word, e.g. *perorować* [pɛrɔ'ɔvɔtɛ] 'to sermonise', *wyładowują* [vɨ'vadɔ'vujɔɯ̃] '(they are) unloading', *zakomunikuje* [zakɔ'mupɨ'kuje] '(I will) announce', *niepowetowana* [ɲepɔ'vɛtɔ'vana] 'irreparable (NOM SG FEM)'. The method reflects Dłuska's awareness of the fact that syllables and segments in longer words in Polish tend to be produced at a higher rate than corresponding elements in shorter words (1932: 9). The phenomenon has become known as 'polysyllabic shortening', and has been reported and measured for English, as well as for other languages, such as Swedish (e.g. Lehiste 1972, Lindblom & Rapp 1972, Turk & Shattuck-Hufnagel 2000). Direct comparison of identical consonants within the same token also has another obvious advantage: it circumvents the problem of cross-category variation (cf. §3.1.2 below), as well as inter- and intraspeaker variation. It also has limitations, some of which were pointed out by Dłuska. The consonants compared were phonemically identical, but did not occur in the same vocalic context, which could have some impact on the differences reported. Pairwise comparison, as in the examples

mentioned above, allows us to trace potential differences in duration between any two prosodic positions. However, some positions were under-represented in the study, or sometimes entirely absent, because the dataset did not include examples containing identical consonants in these positions. This does not allow to verify the validity of some of the reported patterns using a modern paradigm. Specifically, trochaic rhythm in six-syllable or longer words, which was claimed to be signalled by onset lengthening in the third syllable, was documented by examples such as *niepowetowana* [ɲepɔvɛtɔ'vɔna] 'irreparable (NOM SG FEM)', *karykaturalne* [kari'katu'ralnɛ] 'caricature (NOM SG NEUT ADJ)' and *nie powytrzępywał* [ɲepɔvɪtʂɛ'pɪvɔw] 'he has not dusted', which allow direct comparison of the onset of the third syllable with the onsets of the penultimate, the initial and final syllables respectively (see (3)). However, there were no examples such as *ananasowego* [ana'nasɔ'vegɔ] 'pineapple (GEN SG ADJ)', where the third syllable (a potentially strong position) could be compared directly with the preceding syllable (a weak position) to verify the hypothesis of alternating stresses. The examples in (3) serve as an illustration of the limited scope of Dłuska's pairwise comparisons and the fact that the relative lengthening of the onset of the third syllable was observed especially in comparison with the final syllable (3c). It is interesting to note, however, that in some tokens the third syllable exhibited a longer onset than the penult; e.g. example (3a) for speaker A.

(3) *Onset consonant duration in different positions in six-syllable words*

a. *Third syllable vs. penultimate syllable*: [v]

niepowetowana [ɲepɔvɛtɔ'vɔna] 'irreparable (NOM SG FEM)'

speaker A: 70 ms vs. 60 ms	speaker B: 50 ms vs. 60 ms
60 ms vs. 55 ms	55 ms vs. 60 ms

b. *Initial syllable vs. third syllable*: [k]⁴

karykaturalne [kari'katu'ralnɛ] 'caricature (NOM SG NEUT ADJ)'

speaker A: 110 ms vs. 100 ms	speaker B: 100 ms vs. 65 ms
110 ms vs. 90 ms	95 ms vs. 80 ms

c. *Third syllable vs. final syllable*: [v]

nie powytrzępywał [ɲepɔvɪtʂɛ'pɪvɔw] 'he has not dusted'

speaker A: 60 ms vs. 55 ms	speaker B: 65 ms vs. 50 ms
70 ms vs. 45 ms	55 ms vs. 40 ms

Dłuska's (1932) findings on rhythmic stress were not confirmed by two later acoustic studies using more advanced computer-based tools, Dogil (1999) and Newlin-Łukowicz (2012), which differ considerably with respect to the amount of data analysed and the scope of the investigated patterns.

⁴ As mentioned earlier, the words in Dłuska's experiment were embedded in sentences, which made measuring the closure phase of a word-initial voiceless stop feasible. The initial *k* of the word *karykaturalne* appeared in the sentence *To było karykaturalne* 'It was like a caricature' (1932: 74).

Dogil's study investigated acoustic correlates of secondary stress in two separate small-scale experiments. The first experiment addressed the issue of the phonetic underpinnings of degrees of stress, and was based on a single six-syllable word, *marmoladowymi* [ˌmarmɔˈladɔˈvimi] 'marmalade (INSTR PL ADJ)', produced by three male speakers in three different contexts: no focus, broad focus and narrow focus. (No information was given about the area of Poland from which the speakers came.) Measurements were conducted for nine tokens (3 speakers × 3 positions), which were elicited using a constructed dialogue paradigm.⁵ The acoustic parameters under investigation were F0 (average, minimum, maximum, variability), intensity (average, minimum, maximum, variability), syllable duration and vowel quality (presence *vs.* absence of vowel reduction). The duration of onset consonants was not taken into consideration as a potential stress correlate. Measurements were conducted for all six syllables of the word, with parameter values averaged over the three speakers. In no-focus position, the only acoustic parameters which correlated with the penultimate syllable carrying main stress were maximum F0 and a steep F0 slope. The initial syllable was reported to be characterised by increased length and absence of vowel reduction. These characteristics of the initial syllable were found regardless of the presence or absence of focus. No prominence effects were found in the syllable *la*, traditionally associated with word-internal subsidiary (tertiary) stress. Under broad focus, no position was characterised by pitch accent. Under narrow focus, the positions of primary and secondary stresses were reported to be switched, i.e. it was the initial syllable that had higher F0 values. In sum, no context-invariant correlates of word stress were discovered, leading to the conclusion that the initial and penultimate positions were merely landing sites for the intonational pitch accent (Dogil 1999: 291). With respect to rhythmic stress, no acoustic correlates of word-internal prominence were identified.

The second experiment focused on two identical syllables, *-po.po-*, in three different grammatical forms of the same lexical item, 'hippopotamus': *hipopotam* (NOM SG), *hipopotama* (GEN SG) and *hipopotamami* (INSTR PL). The addition of inflectional endings makes the word longer, and changes the rhythmic relations within the *-po.po-* sequence.⁶ The study was based on 27 tokens (3 words × 3 speakers × 3 repetitions), produced in the frame 'I said ... twice'. The parameters and the averaging methods were the same as in the first experiment. The results indicated

⁵ The target sentence containing the word *marmoladowymi* (*Taca z marmoladowymi ciastkami leży na stole* 'The tray with marmalade cookies is on the table') was elicited in response to the investigator's questions focusing on one of the constituents of the sentence.

⁶ Dogil's transcription of the word *hipopotama* (*hi.popo'tama*, rather than *hipopo'tama*) is not consistent with the Polish bidirectional stress pattern. According to the traditional literature (e.g. Dłuska 1974), stressing the second rather than the initial syllable can only result from splitting the prosodic word into two separate units, a rare phenomenon enforced by poetic metre. It is not expected to occur in ordinary speech under normal conditions; see Dłuska (1974: 18–20) on prosodic splits (*zestroje rozpadowe*) and prosodic contractions (*zestroje ściągnięte*) in speech *vs.* poetry.

that secondary stress was absent: it was not implemented by any of the phonetic cues under investigation. Additionally, the same three words were recorded using an electromagnetic articulographic technique. In this part of the experiment, eight repetitions of the words were produced by a single speaker (giving a total of 24 tokens). No significant differences were discovered in the articulatory trajectories. Although the two studies were admitted to be 'limited both with respect to number of speakers and number of tokens', the result was claimed to be 'rather peculiar given the clear intuition in favour of trochaic rhythm expressed by many Polish phonologists' (1999: 304).

Newlin-Łukowicz's (2012) acoustic study of secondary stress was based on a vast amount of data, analysed using powerful inferential statistics. The data were collected in a word-list reading task. The word-list contained 100 single-root words and 38 compounds. From a methodological point of view, compounds exhibit a superset of the problems relating to single-root words (see note 1); these will be ignored in what follows. Single-root words were divided into four-, five- and six-syllable words. Six-syllable words are of particular interest for the present study, because iteration of subsidiary stress can only take place in these words. These were the smallest subset, containing only ten items (see the supplementary materials to Newlin-Łukowicz 2012).⁷ The word-list was read three times by ten speakers. (As already mentioned in §1, none of the speakers came from central or southern Poland, the only areas for which rhythmic stress has been successfully detected thus far.) The acoustic analysis focused mostly on vowel parameters: vowel duration, maximum intensity and F0 (split into maximum, change and slope). The expectation was that subsidiary stress would be cued acoustically by the same parameters as primary stress, or at least by a subset of these parameters. The effect of syllable position on potential stress parameters was tested statistically with a series of mixed-effects regression models, fitted separately for each parameter.

In general, the results of the study confirmed the status of the penultimate syllable as the most prominent prosodic position. The penult was characterised by increased values for most parameters under investigation: it had longer duration than other positions (except for the pretonic position in the subset of high vowels) and the highest maximum intensity (a property it shared with the initial syllable), as well as the largest pitch slope and pitch change (a property it shared with the final syllable).⁸ In contrast, the syllable positions traditionally associated with lower levels of stress, the initial syllable in four-, five-, and six-syllable words and the third syllable in six-syllable words, were reported not to show much acoustic prominence. Indeed, the third syllable in six-syllable words had lower values for all the parameters than the second syllable. Interestingly, initial

⁷ Available at <https://doi.org/10.1017/S0952675712000139>.

⁸ The largest pitch change was seen in the penultimate but not in the final syllable only in the case of the low vowel /a/ (Newlin-Łukowicz 2012: 297). The two syllables also differed in the preferred direction of pitch excursion.

position did show some prominence effects: most conspicuously, it exhibited the largest maximum intensity of all syllable positions, except the penult. However, Newlin-Łukowicz (2012: 301) tentatively discards the possibility that the heightened intensity could underpin secondary prominence in favour of the interpretation that it is merely a ‘word-boundary effect’ based on amplitude declination unconnected to stress. In sum, the presence of secondary stress on the initial syllable is rejected not because of the lack of certain stress-related acoustic effects in this syllable, but because no similar effects were discovered word-internally. Thus the argument against rhythmic stress in Polish, and, by implication, bidirectional systems in general, hinges on the purported absence of stress in one position, i.e. the third syllable of six-syllable words.

Although Newlin-Łukowicz’s study is based predominantly on vowel parameters, she also considers consonantal length to be a potential correlate of subsidiary stress. Measurements of consonantal duration were based on 18 single-root words, repeated three times by ten speakers (a total of 481 tokens). The sound measured was [b] in intervocalic context. Thus, similarly to Dłuska (1932), the tokens compared belonged to a single consonantal category. However, the comparison of raw duration values was conducted across word tokens, and not within a single prosodic word. This potentially creates a bias against six-syllable words, in which syllables and segments tend to be shorter regardless of prosodic position, because of the faster speech rate in longer words. As in all other analyses, random intercepts of Speaker and Item were included. The latter was supposed to help control the variation caused by differences in word length. However, an additional observation that can be made on the basis of the word-list provided is that six-syllable words in this analysis were under-represented in comparison to four- and five-syllable words. The dataset contained only three six-syllable words with the sound [b], representing only two syllable positions out of five for which the measurements were conducted. Among them, there was only one six-syllable word with [b] potentially bearing tertiary stress: *nagabywanego* [naga₁b₁iva'nɛgɔ] ‘chat up (GEN SG PAST PART)’. The remaining two six-syllable examples were *banalizowali* ‘trivialise (3PL PAST IMPERF)’ and *banalizowanie* ‘trivialising (N)’, both with [b] in the onset of the initial syllable. There were no tokens with [b] in the onset of the second syllable to control for the polysyllabic shortening effect. The major findings obtained through a mixed-effects model were as follows (2012: 290–291). Onsets of initial syllables were on average 8.5 ms longer than those of the baseline (i.e. the penultimate syllable; $M = 76.6$ ms). Onsets in the third syllable of six syllable words, traditionally associated with iterative rhythmic stress, were reported to be shorter than those in other syllable positions. However, the differences were not statistically significant. Also, they were claimed to be smaller than the smallest ‘just noticeable difference’ posited in the literature, i.e. 10 ms (Klatt 1976). In short, no prominence effect of consonant duration was detected, and Newlin-Łukowicz interprets the lengthening of word-initial onsets as yet another word-boundary effect.

Summing up, the two recent studies employing relatively advanced instrumental techniques and/or statistics shed some light on the role of vowel parameters in defining stress levels in Polish. Despite the methodological differences, for the vowel in the word-initial syllable traditionally associated with secondary stress they both find heightened parameter values, although the reported parameters are different for each study (duration and lack of reduction *vs.* maximum intensity). This suggests that Polish has a ‘hammock’ (Elenbaas & Kager 1999: 309) or ‘dual’ (Gordon 2002: 495ff) stress system, a claim further supported by the fact that word-initial prominence – expressed in terms of yet another parameter, consonant duration – is also reported by Dłuska (1932). They also seem to confirm that the acoustic underpinnings of tertiary stress are not likely to be found in vowels. Because there has been as yet no attempt to replicate Dłuska’s successful detection of subsidiary stress in a contemporary framework, the issue of whether this type of stress is illusory or physically real has not been resolved.

Intuitively, given the regular rhythmic structure of six-syllable words, it would not be at all surprising if the beat perceived on the third syllable were purely epiphenomenal, an image of a trochee built from the properties of the context rather than a real object itself (cf. Hawkins 2010 on illusory ‘auditory objects’). But it is also possible that word-internal subsidiary prominence is a real effect, but too subtle to emerge in calculations based on the across-token baseline to which estimates for the means corresponding to particular positions are compared. Above all, comparisons based on raw duration values carry a risk of the prominence of the third syllable in six-syllable words being masked by the polysyllabic shortening effect.

Consider the example in Fig. 1 as an illustration. The data come from recordings conducted for the purpose of the present study, and represent two word tokens produced by a single speaker (male; aged 23): *ananasowego* [ʌnʌnʌsɔvʲɛgɔ] ‘pineapple (GEN SG ADJ)’ and *ananasowy* [ʌnʌnʌsɔvʲ] ‘pineapple (NOM SG ADJ)’. The two words differ in the number of syllables (due to the different inflectional endings: *-ego vs. -y*), and hence also in terms of their rhythmic structure. If iterative stress is not illusory, we expect the onset of the third syllable to be longer than the onset of the second syllable in *ananasowego*, but not in *ananasowy*. Because the two words contain two identical consonants (dental nasals), Dłuska’s method of direct comparison of consonants in different positions within the same prosodic word can be applied. In *ananasowego* (Fig. 1a), the duration of [n] is 57 ms in the unstressed second syllable, and 83 ms in the stressed third syllable, a difference of 26 ms. In *ananasowy* (Fig. 1b), both *n*’s occur in prosodically weak positions, and have approximately the same length, 77 ms *vs.* 75 ms. In short, the prediction about the length difference in the two words is borne out. Applying the across-token analysis to the same data, we obtain a result that underestimates the real difference: the average duration of *n* in the second syllable is 67 ms, so the stressed *n* of *ananasowego* is only 16 ms longer than the baseline. Suppose the real

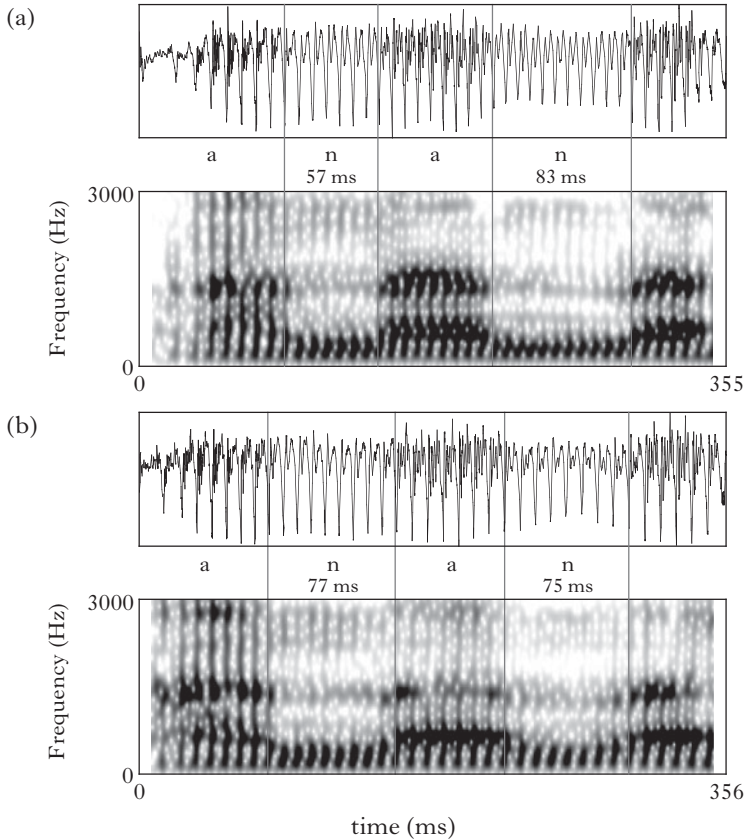


Figure 1

Duration measurements of word-internal tokens of (a) unstressed *vs.* stressed [n] in *anasowego* [ana.nasɔ'vɛgɔ] 'pineapple (GEN SG ADJ)' and (b) unstressed [n]'s in *anasowy* [anana'sɔvi] 'pineapple (NOM SG ADJ)'.

difference between the unstressed and stressed *n* in *anasowego* was 20 ms, i.e. the duration value was 57 ms for the unstressed *n* and 77 ms for the stressed *n* (not 83 ms as in Fig. 1). The stressed *n* would be 'on average' only 10 ms longer than the unstressed *n* of the second syllable. This shows that comparison against the global baseline is not necessarily a good method for detecting subtle local effects of a relative kind. In such cases, relative measures might be a good solution, since they ensure that the local differences are not irretrievably lost in the averaging procedures. Notice also that the lengthening effect in Fig. 1 is not so obvious if judged solely on the basis of the third syllable: in *anasowego*, the stressed *n* is only 8 ms longer than its unstressed counterpart in *anasowy*. The difference in duration is much better seen in the preceding syllable,

where the unstressed *n* is 20 ms shorter in the six-syllable word, i.e. in the context of a following stressed syllable. This indicates that a syllable is stressed not simply because its acoustic parameters oscillate around some threshold values, but primarily because it has higher values for these parameters in comparison to neighbouring syllables within a single prosodic domain. Such an interpretation is compatible with the standard definition of 'stress' in the phonological literature as relative prominence (e.g. Liberman & Prince 1977).

The issue of relative measures may be particularly important from the point of view of investigating the stress pattern of Polish, a language in which even primary stress has been acknowledged to be perceptually robust but acoustically weak. On the one hand, native speakers of Polish do not exhibit the strong 'stress deafness' characteristic of speakers of other languages with predictable stress (Peperkamp & Dupoux 2002, Peperkamp *et al.* 2010). Traditional descriptions of word stress in Polish are unanimous in pointing out its dynamic character, and in defining the difference between stressed and unstressed syllables mostly in terms of 'strength' or 'power' (e.g. Benni 1923, Szober 1923, Doroszewski 1952, Dłuska 1974). This suggests that intensity is the main acoustic correlate. On the other hand, acoustic studies of primary stress in Polish are often inconclusive with regard to the hierarchy of the parameters underpinning stress, or arrive at findings which completely contradict earlier impressionistic descriptions, and are typologically bizarre. While Newlin-Łukowicz (2012) reports that all the classical stress parameters (F0, intensity, duration) are involved in marking primary prominence at the word level (albeit non-uniquely), Dogil (1999) reports none, associating the penultimate position with an intonational pitch accent. In an acoustic study based on word-list data as well as spontaneous utterances, Jassem (1962) conducts measurements of F0, intensity and duration in syllables, and concludes that Polish is a tonal language. In contradistinction, Łukaszewicz & Rozborski (2008), on the basis of naturalistic data and focusing on standard vowel parameters, establish a hierarchy of acoustic correlates (via a discriminant analysis) which shows more compatibility with impressionistic descriptions: mean intensity > mean F0 > duration.⁹ These discrepancies indicate the rather ephemeral and evasive nature of primary stress in Polish. This characteristic leaves little space for distinguishing several intermediate prominence levels, especially if we expect the differences to be expressed along a single acoustic dimension. In the light of the above, relative measures seem indispensable to ensure that subtle local prominence relations are reflected most faithfully. In the present study, a normalisation procedure known as the Pairwise Variability Index is chosen as a convenient expression of local durational differences (see §3.1.5 below). In addition, the focus is on consonant

⁹ To my knowledge, this is the only acoustic study of primary stress in Polish based exclusively on spontaneous speech tokens (1044 vowels) and measurements of vowel parameters (F0, intensity, duration) expressed in proportion to the average value of a given parameter within a given word token (cf. Beckman 1986, Levi 2005).

duration rather than vowel parameters, which are considered only to (dis-) confirm previous results using more detailed methods. Contrary to common belief, primary and subsidiary prominence may not be signalled by the same cues.

3 The current study

3.1 Method

3.1.1 *Participants.* Ten native speakers of Polish (five men and five women; mean age = 33.5, SD = 10.1) participated in the experiment. None of the speakers reported having speech or hearing disorders. All of the participants spoke standard Polish, had been born in Warsaw and had lived there continuously or for most of their lives. All had university degrees, and had some familiarity with foreign languages (usually more than one).

3.1.2 *Stimuli: a paired design.* The list contained 68 words, making up 34 pairs (listed in Appendix A in the online supplementary materials).¹⁰ Each pair consisted of a five-syllable word (the unstressed condition) and a corresponding six-syllable word (the tertiary stress condition): e.g. *ɸomido'row + y* – *ɸomi'doro'w + ego*. Apart from the suffix, the segmental content of the two words was identical.¹¹ For all speakers, the list of 68 words was randomised, to avoid order effects.

The paired design allowed us to circumvent the problem of variation in length caused by the intrinsic durational properties of different vowels and consonants, as well as shifts in length caused by differences in segmental contexts (for discussion of the impact of these factors on segment duration in English, see Peterson & Lehiste 1960 and Lehiste 1970, for example). In this study, the expected variation within the group (consonant duration in a particular position) is larger than the expected difference between the two groups (unstressed *vs.* tertiary stress). The two sets of duration measurements (unstressed *vs.* tertiary stress) are also expected to correlate strongly with each other.

For example, in the recorded tokens of *recytatorski* – *recytatorskiego*, the average length of the affricate [ts] (orthographic *c*) in the onset of the second syllable in the intervocalic context is 113.9 ms (SD = 12.3), while the average length of the alveolar nasal [n] in exactly the same position in *konusowaty* – *konusowatego* and *ananasowy* – *ananasowego* is only 65.9 ms (SD = 12.2). Thus, on average, the affricate is 48 ms longer than the nasal in the same position. Direct comparisons across different segmental categories and/or different segmental contexts are thus irrelevant for detecting a difference between the groups (stressed and unstressed positions) that could be as little as 10 ms, the lowest just noticeable

¹⁰ Available at <https://doi.org/10.1017/S0952675717000392>.

¹¹ There was one pair of words that differed in the initial onset: *powinowactwo* and *spowinowacony*.

difference (JND) usually assumed in the literature.¹² It is clear that if tertiary stress is manifested by enhanced consonant duration, this effect is more likely to be detected by comparison within a single consonant category than across different categories. Needless to say, normalisation across categories, contexts and speakers would not be feasible.

Another important issue is the well-known tendency of segment length to vary in inverse proportion to word length in Polish (Dłuska 1932: 9). Syllables and segments in six-syllable words tend to be shorter than their equivalents in five-syllable words. Thus, when juxtaposing the duration of a given segment in a five-syllable token and the corresponding segment in the paired six-syllable token, it is necessary to use relative rather than raw duration measures. As already mentioned, I use the Pairwise Variability Index (PVI) as a convenient expression of relative length: the length of a segment is expressed in relation to the length of the preceding segment (for details see §3.1.5 below), in order to account for potential polysyllabic shortening effects.

Care was taken to select words containing CV syllables rather than syllables with codas or complex margins, as the presence of both singletons and clusters would increase variation in length in consonantal and vocalic intervals. However, both codas and complex syllable margins are very frequent in Polish, and can scarcely be avoided. Because the consequences of including both singletons and clusters in a paired design did not seem very severe, some examples with clusters were included. In 61% of the words (21 out of 34 pairs), the first three syllables exhibit the canonical (C)V structure (e.g. *ko.nu.so.wa.ty* – *ko.nu.so.wate.go*, *o.pa.tu.lo.ny* – *o.pa.tu.lo.ne.go*). In the remaining examples, consonantal intervals are complex, due to the presence of a coda or a complex onset in the word-initial syllable, e.g. *am.pu.to.wa.ny* – *am.pu.to.wa.ne.go* and *pro.du.ko.wa.ny* – *pro.du.ko.wa.ne.go*. The study focused on the second and third syllables, as it was designed primarily to investigate the underpinnings of word-internal stress. 97% of these syllables (66 out of 68 pairs) are of the CV type.

3.1.3 Experimental procedure and apparatus. All recordings were made in Warsaw. The speakers were recorded individually, reading the list of 68 words from two sheets of paper, in a random order. The carrier phrase was *Powiedziała __ po raz drugi* ('She said __ for the second time'). The recordings were made in a quiet room, with an H4 Zoom portable recorder set at a sampling rate of 44.1 kHz and an AT897 microphone positioned approximately eight inches from the speaker's mouth. Participants were

¹² There is no literature on JNDs specific to Polish, so the threshold of 10 ms can be assumed only tentatively. The results are not subject to straightforward interpretation, as they depend on many different factors. However, it is generally agreed that, in the range of durations between 30 and 300 ms, JNDs are between 10 and 40 ms, although the limit of perceptibility under optimal conditions may be much smaller (Lehiste 1970: 13). The same duration difference limens are assumed in more recent work (Fletcher 2010: 526; see also the literature cited therein).

given time to familiarise themselves with the stimuli before the experiment started. They were encouraged to speak naturally, at their normal speaking rate and volume, and, in case of any mispronunciation or hesitation, to repeat the whole phrase before moving on to the next one. They were allowed to take a break if needed. They were also instructed not to lean towards or move away from the microphone during the recording. The task was described as ‘conveying short messages in a neutral unemotional fashion’. The purpose of the study was not revealed prior to the experiment.

3.1.4 Segmentation. A high-resolution waveform editor (Sound Forge) was used to mark boundaries between the consonants and vowels in the first three syllables of each word. The segmentation was carried out on the basis of expanded waveforms, and was based on three criteria: (i) visual examination of the waveform, i.e. detection of abrupt changes in the amplitude and shape of successive glottal pulses, (ii) visual inspection of the spectrogram in Praat (Boersma & Weenink 2015) and (iii) auditory perception. The boundaries of vowels were aligned with glottal periods, i.e. they were marked at zero crossings, so that each excised segment contained a complete number of cycles. Care was taken to apply uniform segmentation criteria across all tokens. The outcome of the segmentation procedure is exemplified in Appendix B. Of the 3900 segment tokens (10 speakers \times 68 words \times 6 segments minus 10 speakers \times 18 words containing word-initial onsetless syllables), a handful (less than 1%) had to be rejected because the amount of coarticulation between adjacent segments made segmentation unfeasible. In the analysis including initial onsets, a further two tokens were discarded because they contained a short pause before the voiceless stops *k* and *p*.¹³ In addition, tokens containing word-initial onsetless syllables, as well as those with onset clusters, were systematically omitted, and syllables with complex onsets were also omitted in the overall analysis of vowels.

3.1.5 *Measurements and the PVI calculation*

3.1.5.1 Duration. Six duration measurements were taken in each token (the measurements were generated automatically, using a Praat script). The segments measured were the consonants in the onsets of the first, second and third syllables and the following vowels. Raw duration values were then used to calculate PVI values, using the formula in (4).

$$(4) \text{ PVI} = 100 \times ((d_k - d_{k+1}) / ((d_k + d_{k+1}) / 2))$$

where *d* is the duration of the *k*th segment in a word

¹³ Other tokens with word-initial voiceless stops were included because the speakers smoothly incorporated the tokens in the frame, and no disfluency of speech of any kind was observed.

The PVI is the difference in duration for any two adjacent segments (or syllables), divided by the average duration for those segments (or syllables), multiplied by 100 (Low *et al.* 2000, Ballard *et al.* 2010, Ballard *et al.* 2012, Arciuli *et al.* 2014).¹⁴ It is one of several normalisation procedures which have been proposed as rhythm metrics in the literature (cf. Ramus *et al.* 1999, Low *et al.* 2000), and is generally utilised to analyse the degree and direction of stress across syllables of a word. It is usually calculated for vowels and consonants separately, with individual results (obtained via the formula in (4)) summed up and averaged, giving a global rhythm metric (e.g. Low *et al.* 2000: 383). The PVI in (4) is a local measure, as used in Arciuli *et al.* (2014). Values approaching zero indicate nearly equal prominence, as in (5a), while values larger than zero signify differences in prominence. The greatest difference between the two elements is when one of them approaches zero. In such cases, depending on the direction of contrast, peak PVI values would approach -200 or 200 , as in (5b).

(5) PVI values in relation to the smallest/greatest contrast

a. *smallest contrast*: $d_k = d_{k+1}$

$$\text{PVI}_{\min} = 100 \times ((d_k - d_{k+1}) / ((d_k + d_{k+1}) / 2)) = 0$$

b. *greatest contrast (negative or positive)*: $d_k = 0$ or $d_{k+1} = 0$

$$\text{absolute PVI}_{\max} = |\pm 200|$$

In this study, PVI-duration values were used to measure both the magnitude and direction of contrast. The \pm sign is therefore preserved in all calculations (cf. Arciuli *et al.* 2014, where absolute $|\text{PVI}|$ values are used to express the magnitude of contrast only).

Three kinds of PVIs were calculated for the $\#C_1V_1.C_2V_2.C_3V_3$ string: (i) PVI at a V.C juncture ($\text{PVI}_{dur}(V.C)$) – the duration of the onset consonant in relation to the immediately preceding heterosyllabic vowel, (ii) consonantal PVI ($\text{PVI}_{dur}(C.C)$) – the duration of the onset consonant in relation to the onset of the immediately preceding syllable, and (iii) vocalic PVI ($\text{PVI}_{dur}(V.V)$) – the duration of the vowel in relation to the vowel of the immediately preceding syllable. The full details of the PVI formulas used to express normalised duration within the $\#C_1V_1.C_2V_2.C_3V_3$ string are given in (6). In order to capture all potentially relevant local temporal relations across the entire string, two PVI values of the three types were obtained for each token: one calculated on the basis of the second and third syllables (6a, c, e), the other on the basis of the initial and second syllables (6b, d, f).

¹⁴ For discussion on the origin of the PVI measure, see Low *et al.* (2000: 382, n. 2).

The multiplication of the output by 100 is arbitrary, and independent of the normalisation procedure. It is employed only because the normalisation produces fractional values (see Low *et al.* 2000: 383).

(6) *PVI formulas for #C₁V₁.C₂V₂.C₃V₃**PVIs at V.C junctures*

a. $PVI_{dur}(V_2.C_3) = 100 \times ((V_2 - C_3) / ((V_2 + C_3) / 2))$

b. $PVI_{dur}(V_1.C_2) = 100 \times ((V_1 - C_2) / ((V_1 + C_2) / 2))$

Consonantal PVIs

c. $PVI_{dur}(C_2.C_3) = 100 \times ((C_2 - C_3) / ((C_2 + C_3) / 2))$

d. $PVI_{dur}(C_1.C_2) = 100 \times ((C_1 - C_2) / ((C_1 + C_2) / 2))$

Vocalic PVIs

e. $PVI_{dur}(V_2.V_3) = 100 \times ((V_2 - V_3) / ((V_2 + V_3) / 2))$

f. $PVI_{dur}(V_1.V_2) = 100 \times ((V_1 - V_2) / ((V_1 + V_2) / 2))$

In general, if Polish has tertiary stress, six-syllable words are expected to exhibit significantly lower PVI values for the second and third syllables than the corresponding five-syllable words (a, c, e). The numerator part of the PVI formula (i.e. $d_k - d_{k+1}$) is hypothesised to be consistently smaller in six-syllable words than in corresponding five-syllable words, because we expect a larger d_{k+1} value (the length of the onset consonant or the vowel of the third syllable) and, possibly, also a smaller d_k value (the length of the preceding unstressed onset or the length of the preceding vowel). Simultaneously, no such decrease should be observed for PVI values based on the first and second syllables which are defined as prosodically the same in both types of words (b, d, f).

The experiment was designed to test the hypothesis of tertiary stress by comparing segmentally identical material from five- and six-syllable words (i.e. the unstressed and the tertiary stress condition respectively). Only formulas (6a, c, e) are directly relevant to this hypothesis, because they refer to segmentally identical but prosodically different conditions. However, for PVIs at a V.C juncture, both $PVI_{dur}(V_2.C_3)$ (6a) and $PVI_{dur}(V_1.C_2)$ (6b) were calculated. This was done to separate local tertiary stress effects from general contraction effects in six-syllable words, which potentially could have more impact on vowels, and thus produce an artificial decrease in $PVI_{dur}(V_2.C_3)$ (a rise in relative consonant length in a six-syllable word unrelated to stress). If the hypothesis of tertiary stress is borne out, a decrease in $PVI_{dur}(V_2.C_3)$ should be observed, but not in $PVI_{dur}(V_1.C_2)$. As $PVI_{dur}(V_2.C_3)$ expresses duration of the onset consonant with respect to the preceding vowel and in proportion to the whole V.C string, a decrease in its value, indicating tertiary stress, subsumes three scenarios: (i) consonant lengthening, (ii) shortening of the preceding vowel and (iii) consonant lengthening combined with shortening of the preceding vowel. All three cases are attested in the data (see Appendix B).

As pointed out by anonymous reviewers, it would also be interesting to see if both levels of subsidiary stress in Polish are cued by consonantal or vocalic duration (or other vowel parameters; see §3.1.5.2). In order to test a more general hypothesis of consonantal *vs.* vocalic rhythm, consonantal and vocalic PVIs were also calculated for the first and second syllables to include the secondary stress condition, using formulas (6d) and (f). PVI

values obtained via these formulas were further multiplied by -1 to obtain the additive inverse of these values. This is equivalent to changing the order of elements in the denominator, resulting in $(C_2 - C_1)$ and $(V_2 - V_1)$ respectively. As a result of this adjustment, the consonantal and vocalic PVIs for all three conditions (secondary stress, tertiary stress, unstressed) could be expressed in relation to the same element (C_2 or V_2), ensuring minimal comparability among the conditions. However, the secondary stress position remained segmentally unrelated to the other two positions. To resolve a potential segmental confound, an additional analysis based on a single consonant category was conducted; see §4.1.5 for clarification.

3.1.5.2 *F0 and intensity*. Mean F0 and intensity measurements were obtained automatically, using Praat scripts. F0 was measured in terms of the autocorrelation method, with different settings for the pitch floor and ceiling for male and female speakers: 75–300 Hz and 100–500 Hz respectively. The outcome was inspected for possible octave-jump errors. Intensity was measured with the same pitch floor as used in pitch measurements.

Standardised PVIs for F0 and intensity were calculated in accordance with the formulas in (7), in order to express the difference between vowels in the adjacent syllables in semitones (ST) and decibels (dB) respectively.

(7) *Formulas for relative standardised F0 and intensity*

- a. $PVI_{F0} = 12 \log_2(F0_k / F0_{k+1}) = ST_k - ST_{k+1}$
where $F0_k$ is the mean F0 of the k th vowel in a word, expressed in Hz
- b. $PVI_{int} = 20 \log(p_k / p_{k+1}) = dB_k - dB_{k+1}$
where p_k is the mean pressure of the k th vowel in a word, expressed in Pa

3.1.6 *Hypotheses of the current study*. The first subset of hypotheses to be tested is given in (8).

- (8) a. There is tertiary stress in six-syllable words marked by a decrease in $PVI_{dur}(V_2.C_3)$ relative to five-syllable words.
- b. There are contraction effects in six-syllable words leading to reduced length in vowels marked by a decrease in $PVI_{dur}(V_1.C_2)$ relative to five-syllable words.
- c. There is tertiary stress in six-syllable words marked by a decrease in $PVI_{dur}(C_2.C_3)$ relative to five-syllable words.
- d. There is tertiary stress in six-syllable words marked by a decrease in $PVI_{dur|F0|int}(V_2.V_3)$ relative to five-syllable words.

The null hypotheses are that the PVI difference is 0 ($\alpha = 0.05$), while the alternative hypotheses are that it is not.

In addition, the two general hypotheses mentioned in §3.1.5.1 concerning consonantal *vs.* vocalic rhythm were tested. The prediction is

that syllables bearing secondary and tertiary stress will exhibit onset consonant lengthening and/or heightened vowel-parameter values relative to unstressed syllables.

3.1.7 Statistical analyses. All data were analysed using SPSS (version 23). The statistical analyses were divided into two parts. First, a series of linear mixed effects (lme) models were fitted, to compare the corresponding PVI values in two conditions: five- *vs.* six-syllable words. This was done to test hypotheses (8a–d) above. Second, separate lme models were used to test the more general hypotheses concerning consonantal *vs.* vocalic rhythm.

In the first set of analyses (§4.1.1–§4.1.4), the words in each pair were coded as a single ‘item’. This was done to account for different combinations of segments yielding disparate PVI values, dependent on intrinsic differences in length and independent of the stress pattern. Thus ‘item’ was understood as a particular segmental sequence, recurring in two conditions, the five- *vs.* six-syllable word, rather than as ‘an item in a word-list’. Since $PVI_{dur}(V_2.C_3)$, $PVI_{dur}(V_1.C_2)$, $PVI_{dur}(C_2.C_3)$ and $PVI_{dur}(V_2.V_3)$ values were considered as dependent variables in separate analyses, the two values of the fixed-effect term ‘condition’ were interpreted as prosodically different (tertiary stress *vs.* unstressed) or prosodically identical (secondary stress) conditions, depending on position.

In the second set of analyses, overall PVI values were compared (§4.1.5.1). The paired sequences within a given pair of words were coded as single ‘items’, and the fixed-effect term ‘condition’ was assigned one of three values: secondary stress, tertiary stress, unstressed. The corresponding sequences for ‘tertiary stress’ and ‘unstressed’ were coded as single items; the sequences for ‘secondary stress’ were coded as separate items, because they were segmentally unrelated to the other two conditions. In cases reaching significance, pairwise comparisons based on estimated marginal means, with Šidák adjustment for multiple comparisons, were further applied, to determine which stress conditions were significantly different from each other.

In all analyses, except for those involving PVI_{F0} , ‘condition’ was the only fixed-effect term, because using relative measures as well as pairing identical sequences obviated the need for including such effects as Vowel height and Gender.¹⁵ In the analysis with PVI_{F0} as the dependent variable, the interaction term Gender \times Condition was included as a statistically significant predictor (see §4.2).

To account for non-independence (e.g. Pinheiro & Bates 2000, Baayen *et al.* 2008), three random effects were considered, encompassing all the three types of ‘repeated’ or ‘replicated’ measures occurring in the experiment: Speaker, Item and Item \times Speaker. The competing models were compared in terms of likelihood ratio tests, using the restricted maximum likelihood (REML) estimation method. Speaker-specific

¹⁵ We cannot *a priori* exclude the possibility that male and female speakers realise the stress categories differently. However, in the initial lme models neither Gender nor Gender \times Condition turned out to be significant factors.

slopes for the fixed effect 'condition' proved unnecessary, hence random intercept models were constructed. In all the analyses, except for those involving consonantal PVIs, the best fit was achieved by lme models with random intercepts for Speaker, Item and Item \times Speaker. In the case of consonantal PVIs as dependent variables, only the random intercepts for Item and Item \times Speaker were used. The random intercept for Speaker was discarded, as it did not contribute significantly to the model's goodness of fit (cf. for $PVI_{dur}(C_2.C_3)$: $\chi^2(1) = 0.472$, $p = 0.492$; for the overall $PVI_{dur}(C)$: $\chi^2(2) = 1.58$, $p = 0.455$).

For all the models, p -values were established in two ways. First, the models were evaluated with likelihood ratio tests, using the maximum likelihood estimation method. Two nested models differing only in their specification of the fixed-effect term 'condition', a full model (with 'condition') and a reduced model (without 'condition'), were compared using the standard χ^2 distribution reference. Second, p -values were obtained from the SPSS tests of fixed effects using the Satterthwaite approximation to degrees of freedom.

4 Experiment results

4.1 Duration

4.1.1 *Results for the second juncture ($PVI_{dur}(V_2.C_3)$).* The lme model for $PVI_{dur}(V_2.C_3)$ indicated a highly significant effect of stress ($\chi^2(1) = 163.221$, $p < 0.0001$). The PVI difference between five-syllable words (the unstressed condition) and six-syllable words (the tertiary stress condition) amounted to 18.39; see Table I. (The reference category in Table I is the six-syllable word, i.e. the tertiary stress condition.)

	estimate	SE	df	t	p	CI	
						lower	upper
Intercept	-51.65	6.600	40.867	-7.824	<0.0001	-64.98	-38.31
5 syllables (unstressed)	18.39	1.265	333.895	14.529	<0.0001	15.90	20.88
6 syllables (tertiary)	0	0					

Table I

Estimates of fixed effects for the dependent variable $PVI_{dur}(V_2.C_3)$. Lower and upper bounds for confidence intervals are given at the 95% level in this and all subsequent tables. The six-syllable parameter is set to zero because it is redundant.

The positive value of the difference results from the fact that the $PVI_{dur}(V_2.C_3)$ values are negative in both five- and six-syllable words,

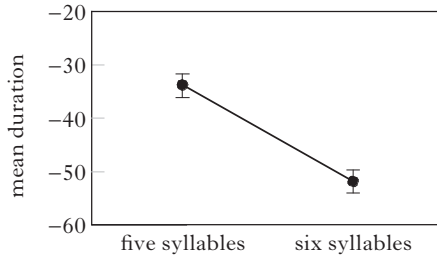


Figure 2

Mean duration for $PVI_{dur}(V_2.C_3)$ (± 1 SE) in five- and six-syllable words: the unstressed *vs.* tertiary stress scenarios.

and smaller in the latter group; cf. the estimated marginal means in Table VI in Appendix C. The observed PVI decrease in six-syllable words shows that the onset of the third syllable is considerably longer with respect to the preceding vowel in six-syllable words than in five-syllable words, as shown in Fig. 2. With the mean raw duration of segments across the second juncture being 67.9 ms, the PVI difference of 18.39 amounts to a 12.5 ms difference in raw duration terms. This result is discussed with respect to optionality of the rhythmic pattern in §5 below.

4.1.2 *Results for the first juncture ($PVI_{dur}(V_1.C_2)$)*. For the $V_1.C_2$ juncture between the first syllable (with secondary stress in both types of words) and the second syllable (unstressed in both types of words), a small but highly significant effect in the opposite direction was found ($\chi^2(1) = 32.063$, $p < 0.0001$). The predicted mean PVI difference was -7.4 , as in Table II.

	estimate	SE	df	t	p	CI	
						lower	upper
Intercept	-13.56	8.151	40.025	-1.663	0.104	-30.03	2.91
5 syllables (unstressed)	-7.40	1.275	337.437	-5.803	<0.0001	-9.91	-4.89
6 syllables (unstressed)	0	0					

Table II

Estimates of fixed effects for the dependent variable $PVI_{dur}(V_1.C_2)$. The six-syllable parameter is set to zero because it is redundant.

The negative difference value signifies a PVI increase in six-syllable words. This reveals that consonants in the onset of the second syllable of six-syllable words tend to be slightly shorter with respect to the preceding vowels than their equivalents in five-syllable words (estimated

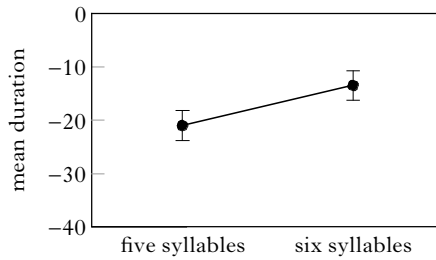


Figure 3

Mean duration for $PVI_{dur}(V_1.C_2)$ (± 1 SE) in five- and six-syllable words: the unstressed scenarios in both conditions.

mean = -13.56 vs. -20.95); see Fig. 3, and Table VII in Appendix C. Although this result was quite unexpected – it was initially anticipated that there would be no difference between the pair of conditions, rather than a PVI increase in six-syllable words – the discovery of the shortening pattern by no means undermines the hypothesis of iterative stress. On the contrary, it offers further support in favour of the alternating stress pattern in six-syllable words. First, we are dealing with a PVI increase, rather than a decrease. We can therefore be sure that the decrease reported for the following juncture earlier in §4.1.1 does not result from an overall tendency to reduce vowel duration rather than consonant duration in six-syllable words, due to their higher speech rate. Second, the relative shortening of the onset of the unstressed syllable in six-syllable words may occur in anticipation of the following stressed syllable, to enhance contrastivity across the two. As the mean raw duration of segments across the first juncture was 79.4 ms, the PVI difference of -7.40 between five- and six-syllable words amounts to a 5.9 ms difference in segmental length.

	estimate	SE	df	t	p	CI	
						lower	upper
Intercept	-2.67	7.196	34.459	-0.371	0.713	-17.29	11.95
5 syllables (unstressed)	8.07	1.157	334.688	6.972	< 0.0001	5.79	10.35
6 syllables (tertiary)	0	0					

Table III

Estimates of fixed effects for the dependent variable $PVI_{dur}(C_2.C_3)$. The six-syllable parameter is set to zero because it is redundant.

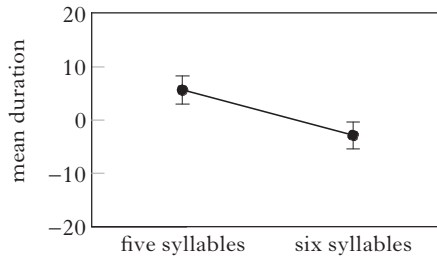


Figure 4

Mean duration for $PVI_{dur}(C_2.C_3)$ (± 1 SE) in five- and six-syllable words: the unstressed *vs.* tertiary stress scenarios.

4.1.3 *Results for onset consonants* ($PVI_{dur}(C_2.C_3)$). The outcome of the lme model testing the effect of stress on consonantal PVIs allows us to reject the null hypothesis that no difference between five- and six-syllable words exists ($\chi^2(1) = 45.41, p < 0.0001$; see Table III).

The mean $PVI_{dur}(C_2.C_3)$ difference between five- and six-syllable words was 8.07, meaning that there was a PVI decrease in six-syllable words, due to the relative lengthening of the consonant bearing tertiary stress. Fig. 4 gives the predicted means for consonantal PVIs in five- and six-syllable words, assuming positive *vs.* negative values respectively (estimated means 5.4 *vs.* -2.67; see also Table VIII in Appendix C). The predicted PVI difference of 8.07 translates into the difference in raw duration amounting to 6.8 ms. (In raw terms, the mean duration of consonants in the second and third syllable’s onsets was 84.8 ms.) This result is further considered in the light of the optionality of rhythmic stress in Polish in §5 below.

4.1.4 *Results for vowels* ($PVI_{dur}(V_2.V_3)$). The fourth lme model in the first series was built to test the relationship between tertiary stress and

	estimate	SE	df	t	p	CI	
						lower	upper
Intercept	-15.85	4.896	35.000	-3.238	0.003	-25.79	-5.91
5 syllables (unstressed)	0.58	1.554	334.308	0.370	0.711	-2.48	3.63
6 syllables (tertiary)	0	0					

Table IV

Estimates of fixed effects for the dependent variable $PVI_{dur}(V_2.V_3)$. The six-syllable parameter is set to zero because it is redundant.

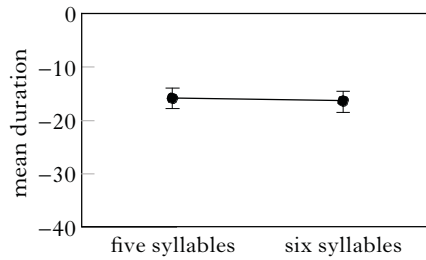


Figure 5

Mean duration for $PVI_{dur}(V_2.V_3)$ (± 1 SE) in five- and six-syllable words: the unstressed *vs.* tertiary stress scenarios.

vowel duration. The PVI decrease in six-syllable words by 0.58 was only minimal, and did not reach significance ($\chi^2(1) = 0.137$, $p = 0.711$); see also Table IV and Fig. 5 (as well as Table IX in Appendix C).

4.1.5 Consonantal *vs.* vocalic rhythm

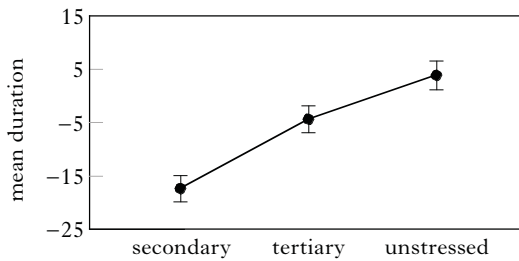
4.1.5.1 *Overall analyses.* The results of two separate lme models testing the effect of stress on consonantal and vocalic PVIs are presented in this section. As mentioned in §3.1.5.1, the PVIs based on the first and second syllables were adjusted: changing the order of arguments to the PVI function in (6d) and (f) had the desired effect of making the second syllable the reference category in all the PVI calculations, so that the first and third syllables could be compared to the second syllable. With the first element now being the same across the PVI formulas, grouping the data into the three categories – secondary stress, tertiary stress, unstressed – could be done on the basis of the prosodic status of the second element featuring in the PVI calculations (i.e. the first and third syllables). For the secondary stress category it was C_1 or V_1 in both five- and six-syllable words, for the tertiary stress category C_3 or V_3 in six-syllable words and for the unstressed category C_3 or V_3 in five-syllable words.

In the lme model for consonantal PVIs, the effect of stress was highly significant ($\chi^2(2) = 53.502$, $p < 0.0001$). The estimates of fixed effects showed a highly significant difference between the unstressed condition and the tertiary stress condition ($p < 0.0001$); the difference between secondary and tertiary stress also reached significance ($p = 0.039$); cf. Table V. In multiple comparisons using Šidák adjustment (Table XI in Appendix C), the difference between ‘secondary stress’ and ‘unstressed’ was significant ($p < 0.005$); however, the difference between the two subsidiary stress positions turned out to be non-significant ($p = 0.113$), although this difference was bigger than the highly significant difference between the unstressed and tertiary stress conditions (cf. Fig. 6, and Table X in Appendix C). The secondary stress condition (based on the initial syllable) was not segmentally paired with the other two conditions

	estimate	SE	df	t	p	CI	
						lower	upper
Intercept	-4.27	6.947	42.006	-0.615	0.542	-18.29	9.74
unstressed	8.01	1.139	503.477	7.032	< 0.0001	5.77	10.24
secondary	-15.79	7.241	24.071	-2.180	0.039	-30.73	-0.85
tertiary	0	0					

Table V

Estimates of fixed effects for the dependent variable $PVI_{dur}(C)$. The tertiary parameter is set to zero because it is redundant.

*Figure 6*

Predicted mean duration for $PVI_{dur}(C)$ (± 1 SE) in five- and six-syllable words, depending on stress.

(both based on the third syllable). Thus, unlike the other two conditions, segmental variability could not be accounted for in terms of the intercept model with Item as the random-effect term. (As explained in §3.1.2, comparison of consonants depending on position in Polish demands paired statistics, because the expected differences within a single position are much bigger than the expected differences between positions.) To control for the potential segmental effect in the secondary stress position, an additional analysis was run, using a small subset of data containing the consonant *m* in the three conditions; see §4.1.5.2.

In the lme model for vocalic PVIs, the effect of stress was non-significant ($\chi^2(2) = 0.448$, $p = 0.799$); see also Tables XII and XIII in Appendix C. The estimates of fixed effects point to very small and non-significant differences between the unstressed and tertiary stress conditions, as well as between the secondary and tertiary stress conditions. This corroborates the findings presented earlier in §4.1.4 that vowel duration is not an acoustic cue to subsidiary stress in Polish (see Fig. 7).

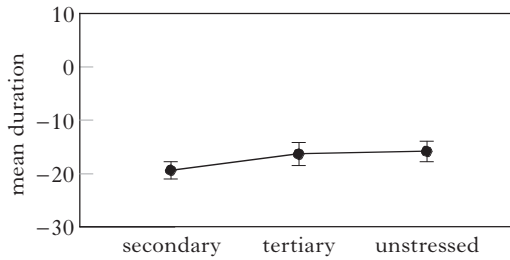
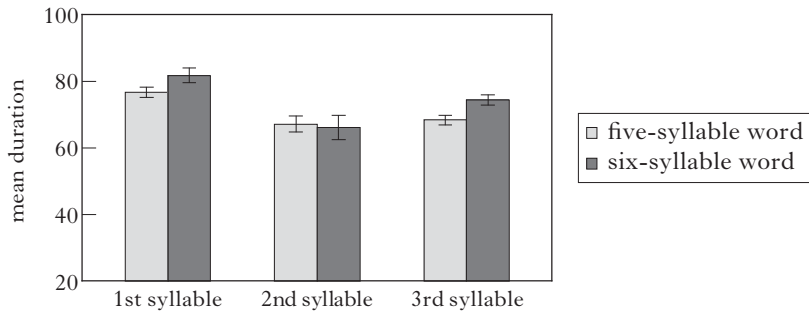


Figure 7

Predicted mean duration for $PVI_{dur}(V)$ (± 1 SE) in five- and six-syllable words, depending on stress.

4.1.5.2 *Effect of syllable position on the duration of [m]*. I report here the results of a small-scale analysis targeting the consonant [m] occurring in the first three syllables of the word, conducted to disentangle the potential segmental confound described earlier in §4.1.5.1. The analysis is based on 200 tokens of [m], all appearing in intervocalic contexts. Separate analyses were run for five- and six-syllable words (recall the discussion of *anasowoy* – *anasowego* in §2). Prominence effects were expected in the first syllable (secondary stress) for both types of words and in the third syllable (tertiary stress) for six-syllable words. The second syllable in both types of words and the third syllable in five-syllable words were predicted to exhibit relatively smaller duration values. An lme model with a random intercept for Speaker was employed to test the hypothesis of the relationship between consonant duration and syllable position. Raw duration values were used.

For both types of words, the effect of stress was highly significant (five-syllable words: $\chi^2(2) = 19.253$, $p < 0.0001$; six-syllable words: $\chi^2(2) = 17.639$, $p < 0.0001$), and the initial position was significantly different from all the remaining positions. In multiple comparisons, the difference between the initial (secondary stress) and second (unstressed) syllables was 10 ms in five-syllable words ($p < 0.005$) and 16 ms in six-syllable words ($p < 0.001$). (Temporal enhancement rather than a polysyllabic shortening effect is seen in the secondary stress position in six-syllable words.) The difference of 8 ms between the second (unstressed) and third (tertiary stressed) syllables in six-syllable words also reaches statistical significance ($p = 0.046$). Additionally, a difference of 7 ms was found between the initial (secondary stress) and third (tertiary stress) positions in six-syllable words ($p < 0.005$). The difference between the second (unstressed) and third (unstressed) positions in five-syllable words was non-significant ($p = 0.711$). These results further corroborate the hypothesis of consonantal rhythm, and tentatively indicate that the prominence effect in word-initial position might be slightly larger than in the syllable

*Figure 8*

Mean duration (ms) for the consonant [m] (± 1 SE) according to syllable position and word-type.

carrying tertiary stress. The mean duration scores in five- *vs.* six-syllable words according to syllable position are shown in [Fig. 8](#).

4.2 F0

Two separate lme models were built to compare mean F0 in different stress environments. F0 was expressed as a local difference in semitones (PVI_{F0}). The dependent variable in the first model was $PVI_{F0}(V_2, V_3)$. A PVI decrease was anticipated in six-syllable words as compared to five-syllable words. The second model was fitted to test the overall effect of stress on F0. In this analysis, analogously to §4.1.5.1, the fixed-effect term Condition had three values – secondary stress, tertiary stress and unstressed. The initial lme models also included Gender as the main factor and the interaction term Gender \times Condition.

In the first analysis, no effect of tertiary stress on F0 was detected. The analysis established a significant interaction of Gender and Condition ($p < 0.001$), due to a steeper F0 decline (PVI increase) in the third syllable of six-syllable words in male speech. In the overall analysis, the interaction term Gender \times Condition was also a significant predictor ($p < 0.0001$). In multiple comparisons, the secondary stress condition turned out to be significantly different from the other two conditions ($p < 0.0001$). The degree of contrast between the first and the second syllables was smaller than between the second and the third syllables: the predicted mean PVI_{F0} for the secondary stress condition was close to zero ($PVI_{F0}(V_1, V_2) = 0.01$ ST); the predicted overall mean difference (0.50 ST) between the second and the third syllable indicated a slight F0 lowering in the third syllable both in five- and six-syllable words. For male speakers, the F0 lowering in the third syllable was robust only in six-syllable words. The relatively stable F0 pattern across the first two syllables followed by a decrease in the third syllable was also reported in Newlin-Łukowicz's (2012) study. These effects are unconnected with subsidiary stress.

4.3 Intensity

No significant results were obtained from analogous analyses of the role of intensity. There was no tertiary stress effect on $PVI_{int}(V_2.V_3)$ ($\chi^2(1) = 0.308$, $p = 0.579$). (The predicted mean $PVI_{int}(V_2.V_3)$ for the tertiary stress condition was -0.48 dB, and for the unstressed condition -0.42 dB.) In the overall analysis, there was no significant difference among the secondary, tertiary and unstressed conditions ($\chi^2(2) = 0.424$, $p = 0.809$). The estimated mean for $PVI_{int}(V_1.V_2)$ (the secondary stress condition) was -0.64 dB, indicating heightened amplitude values in the first vowel of a word relative to the vowel in the second syllable. No further declination of amplitude was observed across the second and third syllables, as shown by $PVI_{int}(V_2.V_3)$ estimates. This departs from the findings reported in Newlin-Łukowicz (2012), and at first glance supports a stress-related interpretation of the raised intensity values in the first syllable. However, the second syllable had a lower amplitude not only relative to the first syllable, but also compared to the third syllable, regardless of stress. In short, heightened intensity values do not lend themselves to a straightforward interpretation, and this matter must be left for future research.

5 Discussion: theoretical implications

The comparison of PVIs in five- and six-syllable words supports word-internal prominence effects – tertiary stress is connected with relative lengthening of the onset consonant in the third syllable of six-syllable words. An interesting finding is that the effect in the consonants alone is quite subtle, but is reinforced by changes in the neighbourhood. A sharper decrease in $PVI_{dur}(V_2.C_3)$ than in $PVI_{dur}(C_2.C_3)$ is achieved by onset-consonant lengthening, accompanied by shortening of the immediately preceding vowel. The study also reveals a slight temporal enhancement in the environment preceding the third syllable in six-syllable words, as compared to their five-syllable counterparts, showing that tertiary stress is expressed not only by higher parameter values in one position, but also by subtle rescaling of the temporal relations in the syllables before it. (The relative shortening of the onset of the second syllable with respect to the vowel across the first juncture can be seen as an anticipatory change to enhance contrastivity across the first three syllables of the six-syllable word.) However, the vowels of the second and third syllables do not display a difference in relative length depending on stress. Also, higher F0 and intensity values have not been detected for the vowel bearing tertiary stress. In consequence, we can conclude that vowel parameters do not contribute directly to the prominence of the third syllable. In the analysis involving three levels of stress, the unstressed position appears to be different from both the secondary and the tertiary stress positions in consonantal length, but not in vowel-based cues. Rhythmic

stress in Polish thus seems unusual not only because it represents a rare bidirectional pattern with internal lapses, but also because it is based predominantly on consonantal rather than vocalic parameters.

The p -values obtained for the tertiary stress effect (the $PVI_{dur}(V_2.C_3)$ decrease and the $PVI_{dur}(C_2.C_3)$ decrease) are very low ($p < 10^{-36}$ and $p < 10^{-10}$ respectively). This makes the effect extremely significant statistically. However, especially in the case of onset consonant duration relative to the preceding onset ($PVI_{dur}(C_2.C_3)$), the mean difference between the tertiary stress condition and the unstressed condition seems to be small when expressed in raw terms: the 6.8 ms difference is below the generally assumed smallest JND (10 ms). When onset duration is expressed relative to the length of the vowel across a syllable juncture ($PVI_{dur}(V_2.C_3)$), the mean difference turns out to be somewhat larger: 12.5 ms is above the JND threshold. The problem of small mean estimates can be considered from the point of view of the optionality of the Polish rhythmic stress pattern, which has been acknowledged in the literature (see e.g. Rubach & Booij 1985: 284), but has not been analysed quantitatively. Figure 9 shows the frequencies of differences within a range of intervals. As can

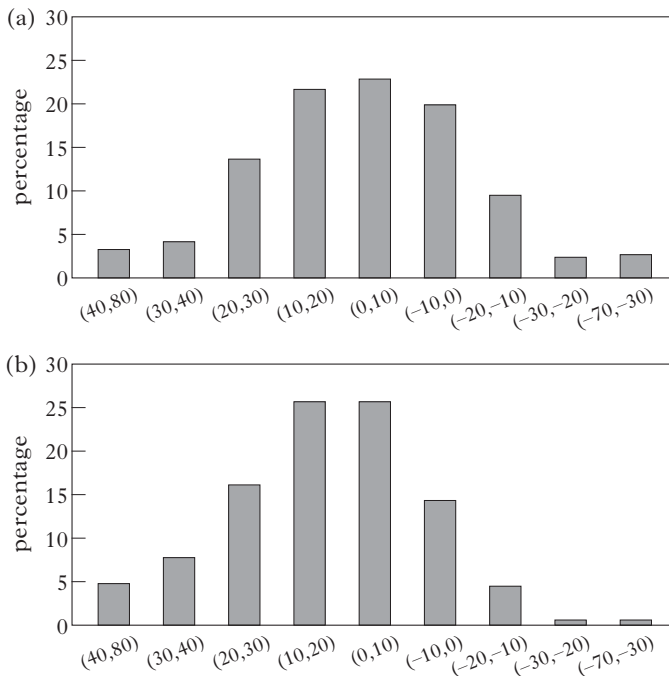


Figure 9

Frequencies of the difference in duration (in ms) between the tertiary stress condition and the unstressed condition for (a) $PVI_{dur}(C_2.C_3)$ and (b) $PVI_{dur}(V_2.C_3)$.

be seen, for both types of relative duration a relatively high percentage of differences fall within the 10–20 ms interval (22% in the case of $PVI_{dur}(C_2.C_3)$ and 26% in the case of $PVI_{dur}(V_2.C_3)$). Differences above 20 ms have a frequency of 21% for $PVI_{dur}(C_2.C_3)$ and 29% for $PVI_{dur}(V_2.C_3)$. In sum, differences above 10 ms occur in 43% of tokens in the case of $PVI_{dur}(C_2.C_3)$ and 55% of tokens in the case of $PVI_{dur}(V_2.C_3)$. Differences in the opposite direction (below –10 ms) are relatively infrequent: 14% for $PVI_{dur}(C_2.C_3)$ and 5% for $PVI_{dur}(V_2.C_3)$. At the same time, both types also show a large proportion of smaller differences around zero, having the effect of lowering the overall mean. If we assume that differences up to 10 ms in either direction are below the perception threshold, the overall proportion of such unnoticeable differences is 43% in the case of $PVI_{dur}(C_2.C_3)$ and 40% in the case of $PVI_{dur}(V_2.C_3)$. Although there are no clear criteria for how optional a process can be in order to count as a phonological pattern rather than random noise, some insight into minimal optionality conditions is offered by studies dealing with statistical learning. Peperkamp *et al.* (2006) implement a statistical algorithm to detect allophonic processes based on complementary distribution, testing its resistance to noise and its capacity to detect optional patterns. In pseudo-language corpora containing 5000 utterances, separation of the real allophonic process from random distributions takes place in the presence of up to 40% of noise, starting from 20% of rule application. If we tentatively adopt these minimal conditions for any kind of phonological process in natural language, the representation of the Polish rhythmic pattern in the phonetic output data seems to be well above the required minimum. Given that the pattern was first observed in 1932, and was considered an optional process even then, the fact that it has survived into modern Polish gives further support to the hypothesis that, at least thus far, it has been sufficiently represented in the speech signal. As the results of this study show, its most conspicuous and stable expression is decrease in $PVI_{dur}(V_2.C_3)$.

Subsidiary stress in Polish is based on consonantal length. Accordingly, the methods used in the research reported here were partially different from those most commonly used in investigating stress patterns based on vowel parameters. A paired design was used to control for differences in intrinsic length (seemingly more disparate in consonants than in vowels); PVI was applied to additionally control for polysyllabic shortening. Among the relative measures posited in the literature (see note 3), PVI, which is a rhythm metric, seems to ensure the most faithful reflection of local temporal relations that lie at the heart of an alternating subsidiary stress pattern. Whether it is necessary to use relative measures in investigating lower degrees of stress, and which measures are most suitable, will need to be established in future research. Still, it seems uncontroversial that if the acoustic robustness of the pattern is uncertain, more elaborate methods may need to be applied in order to conclude that a given pattern does or does not exist.

As mentioned at the beginning of the paper, acoustic studies of lower degrees of stress are much rarer than studies of lexical stress. The available research on secondary stress often points to a lack of acoustic evidence for the stress patterns described in phonological literature. Successful detection of secondary stress has been reported only in a handful of studies (e.g. Rietveld *et al.* 2004, Plag *et al.* 2011). A commonly expressed concern is that phonological analyses of subsidiary stress are mostly based on impressionistic descriptions rather than solid evidence. For example, Tabain *et al.* (2014) argue that, *contra* previous observations by non-native language researchers, there are no acoustic markers of secondary stress in Pitjantjatjara. The original misanalysis is explained in terms of ‘stress ghosting’, an illusion of stress based on non-native speaker’s perception of rhythm. As this is just one of a number of similar patterns which have been disconfirmed by acoustic analysis, the existence of secondary stress itself has been questioned (Tabain *et al.* 2014: 63). However, it is not only the ‘impressionistic descriptions’ but also the ‘solid evidence’ that can be a problem. In the case of the Polish rhythmic pattern, the impressionistic descriptions could not be an effect of stress ghosting, because they were based on intuitions of phonologists who were native speakers of Polish. Dłuska’s (1932) kymographic study of consonant duration in more than 500 word tokens was preceded by her initial impressionistic assessment that the cues for Polish subsidiary stress were located in consonants rather than vowels. Thus the lack of vowel-based phonetic evidence of lower degrees of stress in Polish does not point to an illusion of rhythm, but to a misapprehension that all degrees of stress are expressed by higher values of vowel parameters.

The empirical results confirm the iterative character of Polish subsidiary stress. The theoretical implications of these findings are at least twofold. First, the argument against bidirectional stress systems with internal lapses, hinging on the purported non-iterativeness of the Polish stress, has been refuted. The results thus support metrical theories which predict the existence of such systems (e.g. Rubach & Booij 1985, Kager 2001, 2005, Gordon 2002, Kraska-Szlenk 2003, Alber 2005, Hyde 2008), however rare such systems appear to be. Apart from Polish, which represents the best-documented case, several other languages have been reported to have similar stress patterns. For example, Piro (Matteson 1965) and Lenakel (Lynch 1978), with primary stress on the penult and iteration of secondary stresses from the left edge, exemplify the same rhythmic characteristics as Polish. Garawa (Furby 1974) represents the mirror-image pattern, with primary stress on the initial syllable and subsidiary stresses applied from the right edge. In Gordon’s (2002: §3.2.2) factorial typology of stress, Polish, Piro and Garawa are all classified as ‘binary + internal lapse’ systems, and, *mutatis mutandis*, are generated in terms of the same constraint set. ‘Binary + internal lapse’ systems differ typologically from ‘binary + clash systems’ (Gordon 2002: §3.3). Interestingly, the latter have also been shown to be very rare, but their existence has been confirmed acoustically (Gordon & Rose 2006, Hintz 2006;

see the discussion in Newlin-Łukowicz 2012: 324). In optimality-theoretic terms (Prince & Smolensky 1993), the difference between the two types of bidirectional systems can be expressed in terms of different rankings of the general *CLASH and *LAPSE constraints, banning sequences of stressed and unstressed syllables respectively (Gordon 2002: 529 and the literature cited therein). If bidirectional systems with clashes are predicted and those with internal lapses non-existent (Newlin-Łukowicz 2012: 324), it is difficult to see ‘the inadequacy of the constraint sets that have been used to generate [the latter] systems’ (2012: 323–324), because both types of systems involve nearly the same rankings, and are only minimally different. The results of the present study thus allow preservation of a symmetric stress inventory containing both bidirectional systems with stress lapses and those with stress clashes.

Second, Newlin-Łukowicz’s optimality-theoretic reanalysis of the Polish stress system in terms of undominated *FTFT (‘no adjacent feet’; Kager 1994) cannot be correct. It is self-evident that in a stress system that requires maximal binary foot parsing, the mechanism of clash avoidance expressed by *FTFT cannot be utilised. The parsing of five- and six-syllable words established through acoustic analysis in the present study indicates that Polish is a language which avoids stress clashes, not foot clashes, and thus confirms the metrical analyses of Rubach & Booij (1985) and Kraska-Szlenk (2003). The foot-parsing supported by the phonetic evidence adduced in this paper is as given in (1), and repeated in (9).

- (9) a. [(σ σ)(σ σ)]
 b. [(σ σ)σ(σ σ)]
 c. [(σ σ)(σ σ)(σ σ)]
 d. [(σ σ)(σ σ)σ(σ σ)]

6 Summary of conclusions

This paper has adduced phonetic evidence for consonantal rhythm in Polish, supporting traditional descriptions of the Polish stress system and, in more general terms, questioning the validity of the argument against theories of bidirectional foot parsing generating word-internal lapses. It has shown that tertiary stress in Polish is expressed most conspicuously as a $PVI_{dur}(V_2, C_3)$ decrease – relative lengthening of the onset consonant with respect to the preceding vowel across a syllable juncture. The fact that both vowel shortening and consonant lengthening contribute to the word-internal subsidiary prominence effect indicates that a subtle effect in one position is enhanced by changes occurring in the neighbourhood. Prominence of the third syllable of six-syllable words is also confirmed by a $PVI_{dur}(C_2, C_3)$ decrease – relative lengthening of the onset consonant with respect to the onset of the preceding syllable. Thus Polish subsidiary stress is not a phonologist’s illusion, but a physical entity.

Needless to say, a more complete acoustic description of different degrees of stress in Polish and other bidirectional systems is still required. For Polish, several important issues remain unresolved. With regard to consonantal duration, it would be interesting to see whether (and to what extent) this parameter contributes to the acoustic grounding of primary stress. This would shed light on whether different levels of prominence can be expressed in terms of non-overlapping subsets of acoustic parameters (vocalic *vs.* consonantal), or rather are more uniformly cued in a multidimensional acoustic space. Left for future research is also the acoustic description of prominence relations in seven-syllable words, which potentially combine iterative and bidirectional characteristics. Relevant to subsidiary stress is also the issue of temporal enhancement *vis-à-vis* polysyllabic shortening occurring in longer words. Finally, because Polish has been claimed to be the only uncontested example of a bidirectional stress system with internal lapses, more fieldwork is needed on languages whose prosodic systems potentially show similar features, and which, as of yet, remain *terra incognita* in the phonological literature. For example, according to traditional descriptions, Ukrainian may exhibit an intricate interplay between free lexical stress and rhythmic beats assigned from the opposite word edges, an intuition verified in two instrumental studies by Łukaszewicz & Mołczanow (2018, to appear).

REFERENCES

- Alber, Birgit (2005). Clash, lapse and directionality. *NLLT* 23. 485–542.
- Arciuli, Joanne, Briony S. Simpson, Adam P. Vogel & Kirrie J. Ballard (2014). Acoustic changes in the production of lexical stress during Lombard speech. *Language and Speech* 57. 149–162.
- Baayen, R. H., D. J. Davidson & D. M. Bates (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language* 59. 390–412.
- Ballard, Kirrie J., Danica Djaja, Joanne Arciuli, Deborah G. H. James & Jan van Doorn (2012). Developmental trajectory for production of prosody: lexical stress contrastivity in children ages 3 to 7 years and in adults. *Journal of Speech, Language, and Hearing Research* 55. 1822–1835.
- Ballard, Kirrie J., Donald A. Robin, Patricia McCabe & Jeannie McDonald (2010). A treatment for dysprosody in childhood apraxia of speech. *Journal of Speech, Language, and Hearing Research* 53. 1227–1245.
- Beckman, Mary E. (1986). *Stress and non-stress accent*. Dordrecht: Foris.
- Benni, Tytus (1923). Fonetyka opisowa. In T. Benni, J. Łoś, K. Nitsch, J. Rozwadowski & H. Ułaszyn (eds.) *Gramatyka języka polskiego*. Krakow: Polska Akademia Umiejętności. 1–55.
- Boersma, Paul & David Weenink (2015). *Praat: doing phonetics by computer* (version 6.0.08). <http://www.praat.org/>.
- Dłuska, Maria (1932). *Rytm spółgłoskowy polskich grup akcentowych*. Krakow: Polska Akademia Umiejętności.
- Dłuska, Maria (1974). *Prozodia języka polskiego*. Warsaw: Państwowe Wydawnictwo Naukowe.

- Dogil, Grzegorz (1999). The phonetic manifestation of word stress in Lithuanian, Polish, German and Spanish. In Harry van der Hulst (ed.) *Word prosodic systems of the languages of Europe*. Berlin & New York: Mouton de Gruyter. 273–311.
- Doroszewski, Witold (1952). *Podstawy gramatyki polskiej*. Warsaw: Państwowe Wydawnictwo Naukowe.
- Elenbaas, Nine & René Kager (1999). Ternary rhythm and the lapse constraint. *Phonology* **16**. 273–329.
- Fletcher, Janet (2010). The prosody of speech: timing and rhythm. In William J. Hardcastle, John Laver & Fiona E. Gibbon (eds.) *The handbook of phonetic sciences*. 2nd edn. Malden, Mass.: Wiley-Blackwell. 523–602.
- Furby, Christine (1974). *Garawa phonology*. Canberra: Australian National University.
- Gordon, Matthew (2002). A factorial typology of quantity-insensitive stress. *NLLT* **20**. 491–552.
- Gordon, Matthew & Françoise Rose (2006). Émérillon stress: a phonetic and phonological study. *Anthropological Linguistics* **48**. 132–168.
- Halle, Morris & Jean-Roger Vergnaud (1987). *An essay on stress*. Cambridge, Mass.: MIT Press.
- Hawkins, Sarah (2010). Phonological features, auditory objects, and illusions. *ʒPh* **38**. 60–89.
- Hintz, Diane M. (2006). Stress in South Conchucos Quechua: a phonetic and phonological study. *International Journal of American Linguistics* **72**. 477–521.
- Hulst, Harry van der (2014). Representing rhythm. In Harry van der Hulst (ed.) *Word stress: theoretical and typological issues*. Cambridge: Cambridge University Press. 325–365.
- Hyde, Brett (2008). Bidirectional stress systems. *WCCFL* **26**. 270–278.
- Jassem, Wiktor (1962). *Akcent języka polskiego*. Wrocław: Ossolineum.
- Kager, René (1994). Ternary rhythm in alignment theory. Ms, Utrecht University. Available as ROA-35 from the Rutgers Optimality Archive.
- Kager, René (2001). Rhythmic directionality by positional licensing. Handout of paper presented at the 5th Holland Institute of Linguistics Phonology Conference, Potsdam. Available as ROA-514 from the Rutgers Optimality Archive.
- Kager, René (2005). Rhythmic licensing theory: an extended typology. *Proceedings of the 3rd Seoul International Conference on Linguistics (SICOL)*. Seoul: Linguistic Society of Korea. 5–31.
- Klatt, Dennis H. (1976). Linguistic uses of segmental duration in English: acoustic and perceptual evidence. *ʒASA* **59**. 1208–1221.
- Kraska-Szlenk, Iwona (2003). *The phonology of stress in Polish*. Munich: Lincom.
- Lehiste, Ilse (1970). *Suprasegmentals*. Cambridge, Mass.: MIT Press.
- Lehiste, Ilse (1972). The timing of utterances and linguistic boundaries. *ʒASA* **51**. 2018–2024.
- Levi, Susannah V. (2005). Acoustic correlates of lexical accent in Turkish. *Journal of the International Phonetic Association* **35**. 73–97.
- Liberman, Mark & Alan Prince (1977). On stress and linguistic rhythm. *LI* **8**. 249–336.
- Lindblom, B. & K. Rapp (1972). Reexamination of the compensatory adjustment of vowel duration in Swedish words. *Occasional Papers, University of Essex* **13**. 204–224.
- Low, Ee Ling, Esther Grabe & Francis Nolan (2000). Quantitative characterizations of speech rhythm: syllable-timing in Singapore English. *Language and Speech* **43**. 377–401.
- Łukaszewicz, Beata & Janina Mołczanow (2018). Rhythmic stress in Ukrainian: acoustic evidence of a bidirectional system. *ʒL* **54**. <https://doi.org/10.1017/S0022226717000305>.

- Łukaszewicz, Beata & Janina Mołczanow (to appear). Leftward and rightward stress iteration in Ukrainian: acoustic evidence and theoretical implications. In Bartłomiej Czaplicki, Beata Łukaszewicz & Monika Opalińska (eds.) *Phonology, fieldwork and generalisations*. Frankfurt am Main: Lang.
- Łukaszewicz, Beata & Bogdan Rozborski (2008). Korelaty akustyczne akcentu wyrazowego w języku polskim dorosłych i dzieci. *Prace Filologiczne* **54**. 265–283.
- Lynch, John (1978). *A grammar of Lenakel*. Canberra: Australian National University.
- McCarthy, John J. (2003). OT constraints are categorical. *Phonology* **20**. 75–138.
- McCarthy, John J. & Alan Prince (1993). Generalized alignment. *Yearbook of Morphology 1993*. 79–153.
- Malisz, Zofia & Petra Wagner (2011–12). Acoustic-phonetic realisation of Polish syllable prominence: a corpus study. *Speech and Language Technology* **14/15**. 105–114.
- Matteson, Esther (1965). *The Piro (Arawakan) language*. Berkeley: University of California Press.
- Newlin-Łukowicz, Luiza (2012). Polish stress: looking for phonetic evidence of a bidirectional system. *Phonology* **29**. 271–329.
- Peperkamp, Sharon & Emmanuel Dupoux (2002). A typological study of stress ‘deafness’. In Carlos Gussenhoven & Natasha Warner (eds.) *Laboratory phonology 7*. Berlin & New York: Mouton de Gruyter. 203–240.
- Peperkamp, Sharon, Rozenn Le Calvez, Jean-Pierre Nadal & Emmanuel Dupoux (2006). The acquisition of allophonic rules: statistical learning with linguistic constraints. *Cognition* **101**. B31–B41.
- Peperkamp, Sharon, Inga Vendelin & Emmanuel Dupoux (2010). Perception of predictable stress: a cross-linguistic investigation. *JPh* **38**. 422–430.
- Peterson, Gordon E. & Ilse Lehiste (1960). Duration of syllable nuclei in English. *JASA* **32**. 693–703.
- Pinheiro, José C. & Douglas M. Bates (2000). *Mixed-effects models in S and S-PLUS*. New York: Springer.
- Plag, Ingo, Gero Kunter & Mareile Schramm (2011). Acoustic correlates of primary and secondary stress in North American English. *JPh* **39**. 362–374.
- Prince, Alan & Paul Smolensky (1993). *Optimality Theory: constraint interaction in generative grammar*. Ms, Rutgers University & University of Colorado, Boulder. Published 2004, Malden, Mass. & Oxford: Blackwell.
- Ramus, Franck, Marina Nespor & Jacques Mehler (1999). Correlates of linguistic rhythm in the speech signal. *Cognition* **73**. 265–292.
- Rietveld, Toni, Joop Kerkhoff & Carlos Gussenhoven (2004). Word prosodic structure and vowel duration in Dutch. *JPh* **32**. 349–371.
- Rubach, Jerzy & Geert Booij (1985). A grid theory of stress in Polish. *Lingua* **66**. 281–319.
- Szober, Stanisław (1923). *Gramatyka języka polskiego*. Lwów & Warsaw: Książnica Polska.
- Tabain, Marija, Janet Fletcher & Andrew Butcher (2014). Lexical stress in Pitjantjatjara. *JPh* **42**. 52–66.
- Turk, Alice E. & Stefanie Shattuck-Hufnagel (2000). Word-boundary-related duration patterns in English. *JPh* **28**. 397–440.