Lexical stress in Polish: evidence from focus and phrase-position differentiated production data

Zofia Malisz[1], Marzena Żygis[2,3]

[1]KTH - Royal Institute of Technology, Stockholm, Sweden
[2]Leibniz Centre for General Linguistics, Berlin, Germany
[3]Humboldt Universität zu Berlin, Germany

`malisz@kth.se, zygis@zas.gwc-berlin.de`

Abstract

We examine acoustic patterns of word stress in Polish in data with carefully separated phrase- and word-level prominences. We aim to verify claims in the literature regarding the phonetic and phonological status of lexical stress (both primary and secondary) in Polish and to contribute to a better understanding of prosodic prominence and boundary interactions. Our results show significant effects of primary stress on acoustic parameters such as duration, $f0$ functionals and spectral emphasis expected for a stress language. We do not find clear and systematic acoustic evidence for secondary stress.

Index Terms: lexical stress, focus, Polish

1. Introduction

Typologically, acoustic correlates of word stress are believed to be salient in languages with unpredictable word stress (e.g. [1] for German) and considerably weaker in languages with predictable word stress (e.g. [2] for Slovak). Following this, languages like Polish, with fixed stress on the penult, are hypothesized to be marked by scarce acoustic prominence in stressed syllables [3]. This hypothesis has received indirect support in that there has been no consistent agreement upon the acoustic correlates that express primary stress in Polish so far.

Phonologically, primary stress (S1) in Polish is fixed on the penultimate syllable with few, well-defined exceptions that are rapidly undergoing regularisation to penultimate stress. Secondary stress (S2) falls on the initial syllable in words consisting of at least 3 syllables [4, 5]. Trochees are assigned from right to left with main stress falling on the penult. Rhythmic alternation is assigned from left to right with secondary stress falling on the first syllable. Thus, Polish exhibits a bidirectional stress system. Trochees are posited for quadrisyllabic words [3] with internal lapses in odd parity words.

Perceptually, S1 is robust in Polish [6]. [7] have shown that deviations from the canonical pattern trigger strong neurophysiological responses in Polish listeners, while at the same time they tend to accept these deviations on the behavioural level. As shown by [6], speakers of most languages with predictable word stress are stress "deaf". Polish speakers, however, showed weaker stress "deafness" in [6] than for example speakers of Standard French, Southeastern French, Finnish or Hungarian. This was explained by the existence of exceptions from penultimate stress (0.1% of the vocabulary) and therefore more sensitivity to stress changes.

The acoustic parameters that reflect phonological stress, focus contrasts as well as general prominence variation are fundamental frequency ($f0$), duration, intensity and spectral properties. Intensity and spectral emphasis are two operationalisations of loudness and vocal effort or voice quality. Variation in intensity, especially overall intensity, was found to correlate with stress [8] and accent distinctions. Spectral emphasis was implicated in marking lexical stress [9, 10, 11] and focal accents in e.g. Swedish [12]. Duration particularly well describes stress levels and focus in most languages [13] and contributes to marking accentual structure.

Acoustically, lexical stress in Polish has been traditionally defined as dynamic, i.e. louder, more effortful, based on intensity variation [14]. [15] however, pointed to $f0$ movements as the primary correlate, although confounds between stress and focus in his study are possible (see below). Some effects on duration were found: for example [15, 16, 17] found similar stressed to unstressed ratios for differing speech material: 1.17 in controlled, 1.2 in corpus and 1.22 in data respectively. The difference in duration between stressed and unstressed syllables is believed to be robust if small in magnitude. Crosswhite [18] used a measure of spectral emphasis and found weak evidence of a stress effect in Polish in very few minimal pairs, without differentiating focus conditions. [19] found $f0$, intensity and duration differences between primarily stressed and unstressed syllables.

[20] used perceptual annotation of three prominence levels (weak, strong, non-prominent) in a corpus of spontaneous Polish. The results showed that syllables rated as strongly prominent differed significantly from the weakly prominent ones in the duration, $f0$ and intensity dimensions. No acoustic parameter significantly distinguished between classes rated as non-prominent vs. weakly prominent. [20] suggested that acoustic correlates of prominence manifest themselves as a function of phrase accentuation; at least in a corpus analysis.

Secondary stress is a highly contentious phenomenon in many languages from a phonetic perspective [21, 22]. Regarding Polish, traditional accounts have not provided systematic and robust evidence for S2 in the acoustic, perceptual or articulatory domain [3, 23]. [14] pointed to a lengthening of the initial consonant as evidence for the presence of S2. Similarly, Dogil found a switch of prominence from the penultimate to the initial syllable when under emphasis [24]. The former observation was not replicated by [25]. Recently, [19] speculated that S2 in Polish, phonologically located on the first syllable, might be an artifact of pitch and intensity declination in words longer than three syllables. S2 in Polish is also not clearly assigned in perception [23].

Crucially, in most of the aforementioned studies on Polish stress, the stimuli were not controlled for possible conflations of word and sentence stress. Lexical and sentence stress covary: stressed syllables align with pitch accents, differentiating these factors is important. As pointed out by [26], gleaning

word stress patterns from grammars and isolated word elicitation has likely led to the interpretation of phrasal-accent patterns as word-stress patterns.

[3] did the first study on Polish that used differentiated target words in utterances under no focus, broad and narrow focus. They found no single systematic acoustic correlate of lexical stress in Polish, independent of the intonation context. Instead, [3] proposed that stress positions in Polish are landing-sites for pitch accents with no acoustic manifestation of their own. The study also found that the only systematic parameter of stress was the relatively longer duration of syllables in S2 positions. [25] showed that the initial syllable of narrowly focused words tend to exhibit greater pitch and intensity than the initial syllable of words in a non-focus condition.

2. Hypotheses

In this paper, we analyse the imprint of several prosodic structures on the known acoustic parameters of stress and focus in a production study with 25 Polish native speakers. We examine acoustic patterns of word stress in Polish in data with carefully separated phrase- and word-level prosodic conditions. We hypothesise that:

1) primary stress will have a robust effect on at least one acoustic parameter both under- and in no focus conditions qualifying Polish primary stress as "true" lexical stress i.e. not contingent on phrase accentuation. We also study the interrelationships between word and sentence stress, as well as phrase boundary by studying target syllables and nuclei in different phrase positions.

In addition to primary stress, we investigate secondary stress in light of contradictory evidence concerning its acoustic realisation [19, 27] and perceptual robustness [23]. Given the duration effect of S2 found by [3], we hypothesise that:

2) a clear effect of S2 will be manifested in the syllable onset. If found, the question whether such syllable onset strengthening must be interpreted as word-initial strengthening [28, 29] rather than S2-related phenomenon, will be left for future study.

3. Methods

3.1. Experimental design and procedure

We designed our stimuli with the aim to differentiate between lexical stress and focus conditions as well as between a partial set of possible phrase positions, namely: large intonational phrase medial vs. final positions. Additionally, we control for syllable shape and onset voicing in the choice of test words. The test words occur in each sentence accent and phrase position condition. Table 1 presents the study design and some stimuli examples.

We designed test words with target syllables of the shape CV. The target syllables contain counterbalanced $-/+$ voice stop onsets: /b, d, g, p, t, k/. As vowel quality co-varies with stress and has an effect on spectral emphasis, which itself is affected by stress [17, 30, 31], we study only syllables of the form C and /a/ (dispensing with, e.g., round or palatalising vowels).

Next, we place these syllables in real Polish words (lemmata N=42, no compounds). We have two sets of words: four- and five-syllables long, eliciting S2 that falls on the initial in Polish (in words longer than three syllables). The eventual word set had each target syllable (e.g. /pa/) occurring once in the penultimate position for *primary* stress, in the initial position for *secondary* and in other, non-final position for *unstressed*. As an

example, the target syllable /pa/ was differentiated between lexical stress positions in the word *tuliPAny* for S1, in *PApierosy* for S2 and in *kopaliny* for an unstressed position.

To elicit focus and phrase differentiation, the words were embedded in carrier sentences ($N_{sentence}$=2611) in a simulated question-answer task (Discourse Completion Task). Examples of carrier sentences are shown in Table 1. To illustrate, to elicit the placement of focus on the word *tulipany* (*tulips*) in a phrase-final position, the reading of the carrier sentence *Kamila bought tulips, not roses* was prompted by a question: *Did Kamila buy roses?* To elicit focus on the target word in a phrase-medial position, the sentence *Kamila bought tulips in the supermarket, not roses* was prompted by a question: *Did Kamila buy roses in the supermarket?* and so on.

During the experiment the participants were asked to listen to the pre-recorded context question in Polish and to read the stimulus sentence in response as naturally as possible, as if "in an everyday conversation with a friend". One repetition per factor level combination was recorded per speaker, yielding a repeated measures design. No instructions as to speech tempo were given. If a self-observed mispronunciation occurred, the speaker had the choice to repeat the sentence immediately.

3.2. Speakers

Speakers of standard Polish from the city (or region) of Wrocław (Lower Silesia) (N=7) and from the city (or region) of Poznań (Wielkopolska) (N=18) were recorded. Speakers were students and university clerks between 19 to 43 years old. They had not lived abroad or outside their native region in Poland for more than a year and were monolingual. They also had not received or entered an intensive foreign language programme and/or pronunciation courses at the time of the recording.

3.3. Acoustic parameter extraction

The following acoustic parameters were extracted: duration, *f0*, intensity and spectral emphasis for the target segments in different focus, stress and boundary positions.

We measured the duration of the target syllable itself and of its vocalic nucleus: /a/. For *f0* extraction in Praat, we set the pitch floor and ceiling using a procedure described in [32] for a better estimation of pitch extrema. Root Mean Square intensity was extracted with default settings. Max. *dB* and *Hz* values for all vowels in the carrier sentence were z-score normalised by the carrier sentence mean and standard deviation. Pitch slope was calculated in Hz/sec over the target vowels.

Our material contains optimal target syllable forms to measure spectral emphasis: the same vowel in the nucleus, ensuring no vowel quality confounds. A Fast Fourier Transform spectrum was obtained from the vocalic nuclei of the syllables in question and split into two frequency bands [33]: B1 = 0-1kHz, B2 = 1-4kHz. Energy sum was taken for each of the bands and the difference between: E = B2-B1 was calculated in dB Sound Pressure Level. Thus, the whole mid-frequency range in one band B2 was captured. More energy in the mid/high band is generally perceived as louder [9].

3.4. Statistical analysis

We analysed the data using linear mixed models (using the lme4 [34] and lmerTest [35] packages in R). Separate models for each acoustic parameter as a dependent variable were formulated. Fixed effects entered were: syllable count, word stress, focus, boundary, location (Wrocław, Poznań), gender and voicing

Table 1: *A sample of study design: lexical stress position (Stress: primary = 1, secondary = 2, unstressed = 0) is varied for target syllables (voiceless onsets /pa/, /ta/, /ka/; voiced onsets /ba/, /da/, /ga/) in real Polish words. Focus (under focus = 1) and IPhrase position (IP medial = 0, IP final = 1) conditions are elicited by embedding in meaningful sentences in a Discourse Completion Task (context sentences not presented; target syllables were not boldfaced in the actual experiment).*

| Elicited sentence	Stress	Focus	Int.Phrase
Kamila kupiła tuli**pa**ny, a nie róże.	1	1	1
Kamila has bought tulips, not roses.			
Nie, to Kamila kupiła or**ga**nizator w supermarkecie, a nie Małgosia.	0	0	0
No, it was Kamila who bought a folder in the supermarket, not Malgosia.			
Kamila kupiła **ta**picerki w supermarkecie, a nie ogumienie.	2	1	0
Kamila has bought upholstering in the supermarket, not tyres.			
Nie, to Kamila kupiła me**da**liki, a nie Małgosia.	0	0	1
No, it was Kamila who bought the medallions, not Malgosia.			
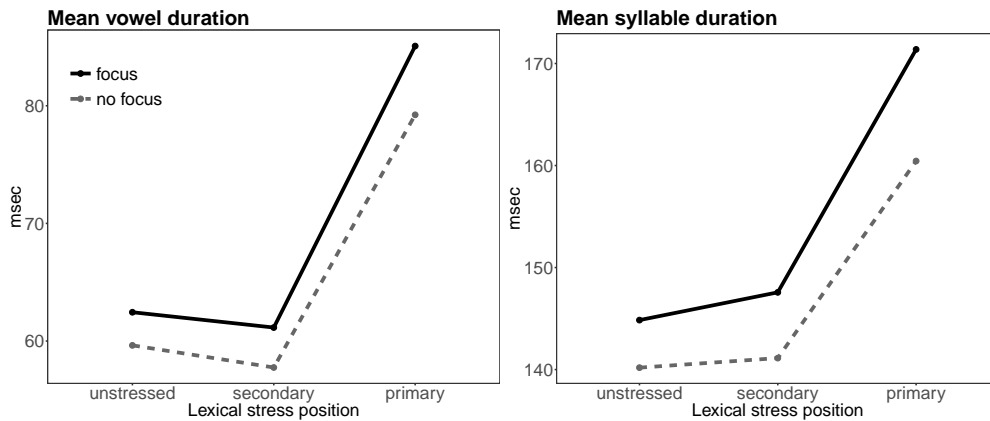
Figure 1: *Duration means for a) the nucleus and b) the whole syllable in three lexical stress positions and two focus conditions. Solid line: stimuli under no focus, dashed line: stimuli under focus.*

status of the onset consonant. We also entered several hypothetically relevant interaction terms (e.g. focus*boundary, focus*word stress, gender*focus). We used backward model selection by first formulating a maximum model with interaction terms for each prosodic fixed factor and a random structure involving random slopes and intercepts for all the interactions. There were convergence errors that led to simplifications of the random structure: we removed higher terms yielding random slopes for word stress, focus and boundary per speaker and per word. We henceforth report only on statistically significant fixed effects of interest.

4. Results

We present mean duration values for the whole syllable and the nucleus in different stress and focus positions in Figure 1.

The results of the linear mixed modeling confirm the effects suggested in the plot: we obtain a robust main effect of S1 on nucleus duration (b= 17.5, p<.001) with a significant positive interaction of focus and S1 (b= 2.7, p<.05). We do not find a statistically significant effect of S2 on the duration of the target vowel relatively to the same vowel in an unstressed position. We also evidence a positive final phrase position and S1 interaction effect on vowel duration (b = 11.7, p<.001).

The entire syllable duration, on the other hand, is first of all affected by S1: the models show a main effect with syllables in S1 positions longer by 19 msec (p<.001) than unstressed

syllables. Again, this effect is stronger, when the primarily stressed syllable is under focus (b = 5.2, p<.001). Importantly, in this case, S2 also affects overall syllable duration by 11 msec (p<.01). We also find a final boundary effect when in the primarily stressed, penultimate position (b = 13.6, p<.001).

Regarding *f0* and intensity parameters and spectral emphasis, the variation of focus, boundary and stress positions results in mean values shown in Figure 2.

Similarly to the results of the nucleus duration model, analysis shows a positive effect of S1 on vocalic spectral emphasis (b = 1.3, p<.01). The LMM did not evidence any statistically significant interactions or other positional effects apart from the positive main effect of focus (b=0.8, p<.001).

The modeling shows a positive main effect of focus on pitch peaks (b = 8.5, p<.001) and an interaction of focus and S1 (b= 5.4, p<.001) but no main effect of S1. Additionally, there is an independent, statistically significant *negative* effect of S2 (b = - 2.7, p<.0001). The presence of an IP final boundary also depresses pitch peaks (b = - 9.3, p<.001).

Models reveal that syllables under focus have a higher intensity (b=2.21, p<.001), especially when the focused syllable is in the S1 position (b= 9.8, p<.001). In the S2 position, syllables have a lower intensity (b=-4.63, p<.001). Overall intensity also falls with the IP-final boundary present (b= -8.6, p<.001).

Finally, we find steeply rising pitch slopes under focus (b = 411.9, p<.001). We also obtain a main negative effect of S1 on the pitch slope (b = -174, p = .048) that enters a significant
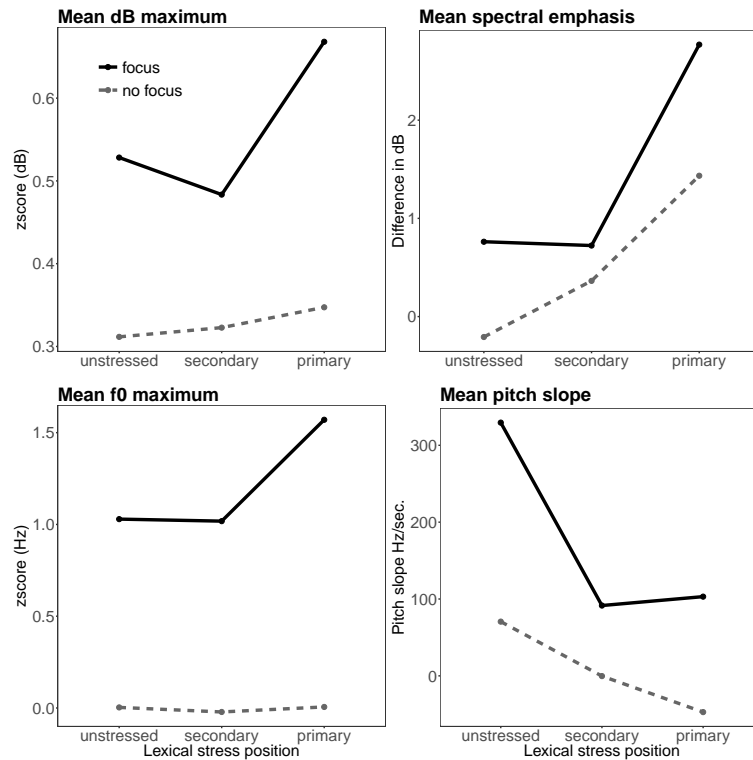
Figure 2: *Top left: mean vocalic intensity maxima. Bottom left: mean f0 maxima; both z-score normalised based on carrier phrase. Top right: mean spectral emphasis for the target vowel. Bottom right: mean vocalic pitch slope. All parameters measured in three lexical stress positions (unstressed, secondary and primary stress) and two focus conditions (solid line: no focus, dashed line: under focus).*

interaction with focus (b = -144.7, p<.046).

5. Discussion

We hypothesised that primary stress will have an effect on at least one acoustic parameter both under- and in no focus conditions, qualifying Polish primary stress as "true" lexical stress i.e. not contingent on phrase accentuation.

Our results show that, primary stress robustly affects syllabic duration (+19 msec), primarily manifesting in the nucleus (+17.5 msec). Spectral emphasis (the difference between the energy sum in band 1-4 kHz and band 0-1kHz) in the /a/ vowel portions of the target syllables also encodes S1 (+1.3 dB).

In terms of parameters encoding primary lexical stress, the acoustic patterns in Polish follow those established in typical stress-based languages: duration and spectral emphasis [9]. However, the acoustic magnitude of the prominence is not as great as it would be in languages with movable stress.

We also find that fundamental frequency variation is a robust correlate of focus but *f0* peaks do not differ across word stress conditions in the no focus condition. That is, only if pitch accent is present, S1 syllables significantly depart in pitch from all others in our carefully controlled data. The findings regarding pitch support the notion that *f0* encodes accentual structure in Polish with S1 syllables serving as landing sites for pitch accents, first put forward by Grzegorz Dogil and colleagues [3] (later also [20]). Intensity peaks behave in a parallel fashion to pitch peaks. The correlated intensity and pitch effects stem from a well-known, physiologically conditioned co-variation [12] and therefore do not constitute a strong independent evidence for stress effects.

We also posited that the most clear effect of secondary stress in Polish will be manifested in the syllable onset. Our analysis shows that the secondary stress position causes syllable onset consonant lengthening (+11 msec).

Both primary and secondary stress significantly affect overall syllable duration, relative to unstressed syllables. These effects are independent of the focus condition [3]. We suspect, however, that the syllable onset lengthening found in the S2 position is due to word-initial strengthening rather than stress, as the lengthening manifests itself mainly in the consonantal onset [28]. Consequently, due to the structural overlap between head and edge effects, it is necessary to empirically differentiate these positions next, in order to answer the question whether there is acoustic evidence for secondary stress in Polish.

Assuming the above interpretation of S2 induced lengthening (cf. [19, 36]), we were not able to show systematic acoustic evidence for secondary stress in our study. However emphatic, post-lexical stress might still be evidenced in Polish: we propose that the secondary stress position is an affordance for acoustic prominence realised under emphasis in specific discourse styles such as media speech [3, 24, 22].

6. Acknowledgements

7. References

[1] M. Jessen, K. Marasek, K. Schneider, and K. Claßen, "Acoustic correlates of word stress and the tense/lax opposition in the vowel system of German," in *Proceedings of the International Congress of Phonetic Sciences*, vol. 13, no. 4, 1995, pp. 428–431.

[2] Š. Beňuš and K. Mády, "Effects of lexical stress and speech rate on the quantity and quality of Slovak vowels," in *Proceedings of the 5th Speech Prosody Conference*, Chicago, 2010.

[3] G. Dogil, "The phonetic manifestation of word stress in Lithuanian, Polish and German and Spanish," in *Word prosodic systems in the languages of Europe*, H. van der Hulst, Ed. Berlin: Mouton de Gruyter, 1999, pp. 273–311.

[4] J. Rubach and G. Booij, "A grid theory of stress in Polish," *Lingua*, vol. 66, pp. 281–319, 1985.

[5] I. Kraska-Szlenk, *The phonology of stress in Polish*. LINCOM, 1995.

[6] S. Peperkamp, I. Vendelin, and E. Dupoux, "Perception of predictable stress: A cross-linguistic investigation," *Journal of Phonetics*, vol. 38, no. 3, pp. 422–430, 2010.

[7] U. Domahs, J. Knaus, P. Orzechowska, and R. Wiese, "Stress deafness in a language with fixed word stress: An ERP study on Polish," *Frontiers in Psychology*, vol. 3, no. 439, pp. 1–15, 2012.

[8] G. Kochanski, E. Grabe, J. Coleman, and B. Rosner, "Loudness predicts prominence: Fundamental frequency lends little," *The Journal of the Acoustical Society of America*, vol. 118, no. 2, pp. 1038–1054, 2005.

[9] A. M. Sluijter and V. J. Van Heuven, "Spectral balance as an acoustic correlate of linguistic stress," *The Journal of the Acoustical society of America*, vol. 100, no. 4, pp. 2471–2485, 1996.

[10] A. M. Sluijter and V. J. van Heuven, "Effects of focus distribution, pitch accent and lexical stress on the temporal organization of syllables in Dutch," *Phonetica*, vol. 52, no. 2, pp. 71–89, 1995.

[11] J. Volín and J. Zimmermann, "Spectral slope parameters and detection of word stress," *Technical Computing Prague*, pp. 125–5, 2011.

[12] M. Heldner, "On the reliability of overall intensity and spectral emphasis as acoustic correlates of focal accents in Swedish," *Journal of Phonetics*, vol. 31, no. 1, pp. 39–62, 2003.

[13] M. E. Beckman and J. Edwards, "Articulatory evidence for differentiating stress categories," in *Papers in laboratory phonology III: Phonological structure and phonetic form*, P. Keating, Ed. Cambridge: Cambridge University Press, 1994, pp. 7–33.

[14] M. Dłuska, *Fonetyka polska* [The phonetics of Polish]. Warszawa: PWN Polskie Wydawnictwo Naukowe, 1950.

[15] W. Jassem, *Akcent języka polskiego* [The accent in Polish]. Wrocław: Ossolineum, 1962.

[16] K. Klessa, "Modelowanie iloczasu głoskowego na potrzeby syntezy mowy polskiej *[Modelling segmental duration for Polish speech synthesis purposes]*," Ph.D. dissertation, Adam Mickiewicz University, Poznań, 2006.

[17] P. Nowak, "Vowel reduction in Polish," Ph.D. dissertation, University of California, Berkeley, CA, 2006.

[18] K. Crosswhite, "Spectral tilt as a cue to stress in Polish, Macedonian and Bulgarian," in *Proceedings of the 15th International Congress of Phonetic Sciences*, Barcelona, 2003, pp. 767–770.

[19] L. Newlin-Łukowicz, "Polish stress: looking for phonetic evidence of a bidirectional system," *Phonology*, vol. 29, no. 02, pp. 271–329, 2012.

[20] Z. Malisz and P. Wagner, "Acoustic-phonetic realisation of Polish syllable prominence: A corpus study," *Speech and Language Technology*, vol. 14/15, pp. 105–114, 2012.

[21] A. Arvaniti, "Secondary stress: evidence from Modern Greek," in *Papers in laboratory phonology II: gesture, segment, prosody*, G. Docherty and D. Ladd, Eds. Cambridge: Cambridge University Press, 1992, pp. 398–423.

[22] J. I. Hualde and M. Nadeu, "Rhetorical stress in Spanish," in *Word stress: Theoretical and typological issues*, H. van der Hulst, Ed. Cambridge: Cambridge University Press, 2014, p. 228.

[23] M. Steffen-Batogowa, *Struktura akcentowa języka polskiego*[The accentual structure of Polish]. Warszawa: PWN Polskie Wydawnictwo Naukowe, 2000.

[24] G. Dogil, *Autosegmental account of phonological emphasis*. Linguistic Research, 1979, vol. 25.

[25] F. Hamlaoui, M. Żygis, J. Engelmann, and M. Wagner, "Acoustic correlates of focus marking in Polish," in *Proceedings of the 18th International Congress of Phonetic Sciences*, Glasgow, UK, 2015.

[26] M. Gordon, "Disentangling stress and pitch-accent: a typology of prominence at different prosodic levels," in *Word stress: Theoretical and typological issues*, H. van der Hulst, Ed. Cambridge: Cambridge University Press, 2014, pp. 83–118.

[27] B. Łukaszewicz, "Polish rhythmic stress revisited: phonetic evidence of an iterative system," *Phonetics and Phonology in Europe (PAPE 2015), Cambridge, UK*, 2015.

[28] D. K. Oller, "The effect of position in utterance on speech segment duration in English," *The Journal of the Acoustical Society of America*, vol. 54, no. 5, pp. 1235–1247, 1973.

[29] L. White, "Communicative function and prosodic form in speech timing," *Speech Communication*, vol. 63, pp. 38–54, 2014.

[30] M. Ortega-Llebaria and P. Prieto, "Disentangling stress from accent in Spanish: Production patterns of the stress contrast in deaccented syllables," *Amsterdam Studies in the Theory and History of Linguistic Science Series 4*, vol. 282, p. 155, 2007.

[31] N. Campbell and M. Beckman, "Stress, prominence, and spectral tilt," in *ESCA Workshop on Intonation: Theory, Models and Applications*. Athens: University of Athens, 1997, pp. 67–70.

[32] C. De Looze and D. Hirst, "Detecting changes in key and range for the automatic modelling and coding of intonation," in *Proceedings of the 4th Speech Prosody Conference*. Campinas: Editora RG/CNPq, 2008, pp. 135–138.

[33] I. Plag, G. Kunter, and M. Schramm, "Acoustic correlates of primary and secondary stress in North American English," *Journal of Phonetics*, vol. 39, no. 3, pp. 362–374, 2011.

[34] D. Bates, M. Mächler, B. Bolker, and S. Walker, "Fitting linear mixed-effects models using lme4," *Journal of Statistical Software*, vol. 67, no. 1, pp. 1–48, 2015.

[35] A. Kuznetsova, P. Bruun Brockhoff, and R. Haubo Bojesen Christensen, *lmerTest: Tests in Linear Mixed Effects Models*, 2016, r package version 2.0-33. [Online]. Available: https://CRAN.R-project.org/package=lmerTest

[36] B. Łukaszewicz, "Phonetic evidence for an iterative stress system: The issue of consonantal rhythm," *Phonology*, vol. 35, no. 1, pp. 115–150, 2018.