

The Acoustic Realization of Prosodic Prominence in Polish: Word-level Stress and Phrase-level Accent

Aleksandra Ćwiek, Petra Wagner

Bielefeld University, Germany

aleksandra.cwiek@uni-bielefeld.de, petra.wagner@uni-bielefeld.de

Abstract

The current study addresses the question of how word-level (“stress”) and phrase- or sentence-level prominence (“accent”) is realized in Polish. For this purpose, a production experiment eliciting semi-spontaneous utterances was conducted, closely following the methodological approach introduced in [1]. Our acoustic analyses are based on identical target syllables which are embedded in sentences under conditions that allow to disentangle word-level and phrase-level prominence. The acoustic realizations of these target syllables are then subject to linear mixed-effect models fitted for various acoustic parameters: duration, fundamental frequency maximum, intensity, and spectral balance. The models indicate that prominence marking in Polish is realized acoustically in a stable fashion on phrase-level only. Word stress marking occurs only in cases where a lexically stressed syllable simultaneously realizes a phrase-level accent.

Index Terms: stress, accent, acoustic correlates, prosody, prominence, Polish

1. Introduction

Prosodic prominence has been found to be realized by various acoustic correlates such as fundamental frequency, intensity, duration, and spectral tilt (or spectral balance) (e.g. [2, 3, 4, 5]), which are manifested language-dependently (e.g. [6, 7, 1]). In a recent cross-linguistic study on the acoustic correlates of German and Hungarian, [1] found no clear-cut acoustic realization of word-level prominence in Hungarian, while phrase-level prominence is realized using the full range of examined acoustic correlates. In Hungarian, word stress is fixed on the word-initial syllable, thus fully predictable, and consequently has no distinctive function. For German, a language with variable word stress position and a (marginally) distinctive function of word stress, all examined acoustic parameters were employed by speakers to mark both word- and phrase-level prominence, albeit to different degrees.

Whether this systematic difference in employing acoustic cues to prominence marking on different levels of the prosodic hierarchy is indeed caused by the typological differences — a distinctive vs. a predominantly delimitative function of word level prominence — in the prosodic systems, requires further evidence from other languages. We therefore aim to establish a set of acoustic correlates used for word-level (henceforth “stress”) and phrase-level prominence (henceforth “accent”) marking in Polish. Polish presents an interesting test case, as it has fixed word stress like Hungarian, but also knows systematic deviations and realizes stress in a different position than Hungarian. To ensure the best possible comparability with results in [1], their methodology to disentangle word-level stress and phrase-level accent was reimplemented as closely as possible.

1.1. Polish prominence

As it is common for Slavic languages, Polish exhibits a free word order [8], though a tendency for an SVO pattern has been found [9, 10, 11]. As for sentence accent placement in Polish, some preferences can be pointed out: In broad focus it is the new information which is in the focal domain and it tends to be the rightmost constituent [12]. Narrow focus can be evoked by *wh*-questions, similarly to the contrastive focus which can moreover be appointed to any constituent, depending on the context [9]. Narrow focus is context-dependent, too, and it can be realized either in a syntactically marked or an unmarked position, yet it does not influence the intonation contour, as shown by [9]. However, word order does have an influence on prominence in broad focus constructions.

As for word-level prominence, Polish primary stress has been characterized as being highly predictable, with some well-known exceptions [13, 6, 14, 15, 16, 17, 18]. The major function of primary stress in Polish is considered as delimitative or demarcative, i.e. signaling an upcoming word ending [19, 7, 16, 20, 21]. Theoretical and experimental work on the acoustic realization of Polish stress identified various acoustic correlates, with results differing between the studies. **Intensity** has been considered or found as a prominence correlate by [13, 22, 23, 18, 21]; **duration** by [16, 18, 21], though [16] states it only serves an “accessory role”; **F0 range** by [13, 6, 16, 18]; **F0 maximum** by [18, 21]; and **spectral tilt** by [24]. Following a point of view that regards prominence realization as being achieved by a set of correlates in a cumulative or complementary way, it can be assumed that *all* of the established correlates play a role in the acoustic marking of Polish stress. It is however important to scrutinize the methodology employed by the various authors, as these are not fully comparable.

Most notably, numerous studies attempting to establish correlates of stress in Polish did not account for higher levels of prominence, such as accent, confounding with stress. For instance, [18] examined word stress correlates in target words placed in accented positions in a carrier sentence. Thus, all measured prominence correlates may have been realized as a consequence of accent and or stress marking. A similar confounding of stress and accent can be found in a study by [24], as the material at least indicates that some target words are likely to have been accented, too.

The only study known to us that carefully disentangled word-level stress and phrase-level accent in Polish has been carried out by [6]. The authors controlled for presence or absence of stress and accent by manipulating prosodic focus (narrow, broad, no focus) on the targets. An analysis of multiple acoustic parameters (F0, intensity parameters, duration, vowel quality) on all syllables of one particular words revealed that word-level stress is at best weakly realized in Polish. They further hypothesized that word stress in Polish is marked only when ac-

accompanied by a phrase-level accent. This hypothesis received support by a corpus study on spontaneous Polish [25], where no reliable acoustic cues for minor prosodic prominences, more or less correlating with word stress, could be identified. However, this study lacks the rigor of experimental control. Given these contradictory results and no further control studies, we still do not know whether Polish stress is acoustically marked independently of accent, and if yes, what acoustic correlates are used for its marking.

1.2. Motivation and goals

In a recent article, [26] argue that a strict disentanglement of word-level and phrase-level prominence is crucial when investigating the corresponding acoustic correlates on separate levels. As argued above, this demand is not met by the majority of studies focusing on Polish prosodic prominence. Furthermore, [27] indicated a lack of phonetically reliable studies on Polish prosody in general, despite its having been widely investigated on various phonetic aspects throughout the years. This is confirmed by the contradictory results and on the still open issue of whether Polish stress is expressed independently of accent. Our study therefore has two main goals: First, we plan to address the open question regarding the acoustic realization of word- and phrase-level prominence in Polish in a way that disentangles word stress and sentence accent, while maintaining a necessary level of experimental control. Second, we want to shed light on whether typological differences between prosodic systems lead to similarities in phonetic implementations. More specifically, we are interested in whether the predictability of word stress leads to a lack of its acoustic realization.

2. Methods

2.1. Participants and materials

Twenty-nine native speakers of Polish (16 female; median age = 22) took part in a production study. All participants were monolingual native speakers of Polish and students at the Faculty of Modern Languages and Literatures of Adam Mickiewicz University in Poznań. Two speakers reported to have a reading disorder (i.e. *dyslexia*), yet it did not obviously influence their subsequent behavior during the experiment.

The elicitation materials were designed in accordance with [1] and consisted of seven sentence quadruples, resulting in 28 target sentences in total, plus 18 fillers. Each sentence within one quadruple contains an identical target syllable, for which the two binary factors \pm word stress and \pm sentence accent are manipulated by varying word and sentence/information structure. An example of a quadruple is given in the following; the target syllable is underlined, stress is marked bold, and accent domain is written in capitals:

- (1) + lexical stress, + sentence accent
Jeśli chcesz **GOTOWAĆ**, nie używaj lewego palnika.
'If you want to cook, don't use the left burner.'
- (2) + lexical stress, – sentence accent
Nie musisz w OGÓLE **gotować** jeśli nie chcesz.
'You don't have to cook at all if you don't want to.'
- (3) – lexical stress, + sentence accent
Zamiast **GOTOWANIEM**, lepiej zajmij się **SPRAZATANIEM**.
'Instead of cooking, better take care of cleaning.'

Jeśli chcieć



gotować



nie używać

lewy palnik

Figure 1: An example instruction as presented to the participants.

- (4) – lexical stress, – sentence accent

Do **gotowania** nie używaj też **PRAWEGO** palnika.
'For cooking, don't use the right burner as well.'

To shift stress while preserving it on the same syllable and in a maximally similar context, either a part of speech was modified from verb to noun (as presented above), or declension was used. The target syllable was always placed word-medially to avoid prosodic boundary effects. Phrase-level accent was controlled for according to Polish information structure tendencies and, where possible, relied on additional structural enhancements such as topicalization, narrow, or contrastive focus [9, 28]. The sentences were controlled by two expert native speakers and two non-expert native speakers to ensure a high likelihood of the intended realizations.

2.2. Elicitations and recordings

To avoid a problematic read speech style [29], the paradigm elicits semi-spontaneous utterances while maintaining experimental control over the conditions. For this, participants are faced with a fictional scenario: they are going on a vacation and are now faced with the task of instructing another person (who is a confederate) to take care of their apartment during their absence. These instructions have to be formulated based on a combination of pictures and uninflected word forms that are presented on a computer screen. These have to be constructed into the target sentences from left to right and top to bottom (see and example in Figure 1). That way, a semi-spontaneous quasi-dialogue evolves, as they directly face the confederate who actively provides feedback.

The order of the presented instructions was fixed to ensure semantic and contextual coherence of the utterances. Also, the participants were asked not to change the predetermined word order within the sentences; if they did so anyway, they were requested to repeat the utterance on the plea that the experimenter had not heard or understand it well. All recordings took place in a sound-proof booth in the psycholinguistics lab of the Faculty of Modern Languages and Literatures of Adam Mickiewicz University in Poznań using a Neumann TLM 103 condenser microphone.

All target syllables were manually annotated and the following potential acoustic correlates of prominence were extracted using a Praat script: syllable duration (ms, logarithmically transformed), maximum vowel intensity (dB), F0 max (st, z-normalized by speaker), and spectral balance (SPLH-SPL). For these dependent variables, linear mixed effects models were fitted with R [30]. The models included *stress*, *accent*, the *interaction of stress and accent*, and *part of speech* as fixed effects

and a minimal by-item and by-subject random effects structure, following [31]. Part of speech was included in the models, as it varied across the target words and was a potential confounding variable (cf. Section 2.1, example sentences (1-2) with target syllable embedded in verbs vs. (3-4) with target syllable embedded in nouns).

3. Results

The detailed results of the fitted models are presented in Table 1. For duration (cf. Figure 2 for its distribution across conditions), two models were fitted – one containing the interaction of stress and accent in the random effects structure and one without, due to a near to 1 correlation of accent with the interaction of stress and accent in both by-item and by-target random slopes. The two fitted models for duration differ significantly from each other, yet both indicate very similar effects, namely a highly significant effect of stress, accent, and their interaction, on duration.

In the models fitted for intensity (cf. Figure 3), spectral balance (cf. Figure 4) and maximum F0 (cf. Figure 5), only accent and part of speech have a significant effect, while neither stress nor an interaction of stress and accent does contribute significantly to their variation.

4. Discussion and conclusions

A first goal of our study was to better understand the acoustic realization of prominence in Polish. Results of the linear mixed-effect models indicate that a reliable prominence marking in Polish occurs on phrase-level accents only. Even if our data indicates a significant contribution of stress on duration, this effect is modulated by a significant interaction of stress and accent: As can be observed in Figure 2, stress alone does not enhance duration, neither does accent on its own; in fact, only if both stress and accent occur together, syllable duration is increased. The results obtained for fundamental frequency, intensity, and spectral balance indicate that it is accent that dominates the acoustic realization of prominence in Polish, while stress has no independent effect.

Our findings contradict previous studies stating that primary word stress in Polish is acoustically expressed by intensity [23, 18], duration [18], fundamental frequency [18], and spectral tilt [24]; however, current results are in line with [6, 25]. Furthermore, our findings corroborate with an assumption that in fixed stress languages, word stress serves as a “landing site” for phrase-level prominence expressed as accents. Interestingly, we found that accent was not expressed by increased duration unless it coincides with lexical stress. A similar idea has previously been put forward by [32].

Our results mirror previous findings for Hungarian, a fixed stress language for which no clear-cut acoustic correlates of word stress could be detected. A potential explanation for this lack of prominence marking could be an underlying production economy principle, as put forward by H&H theory [33]. In a communicative situation, speakers are exposed to great cognitive effort, hence they behave economically and “are expected to vary their output along a continuum of hyper- and hypospeech” [33, p. 403]. As stress marking in Polish is not distinctive or otherwise motivated by semantic considerations, and a lack of its marking does not compromise comprehension, its marking must have an independent purpose such as signaling phrase accent and information structure.

One could argue that the richness of acoustic cues to signal

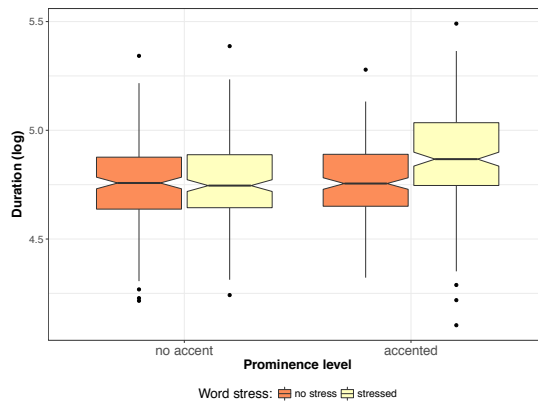


Figure 2: Duration depending on prominence level.

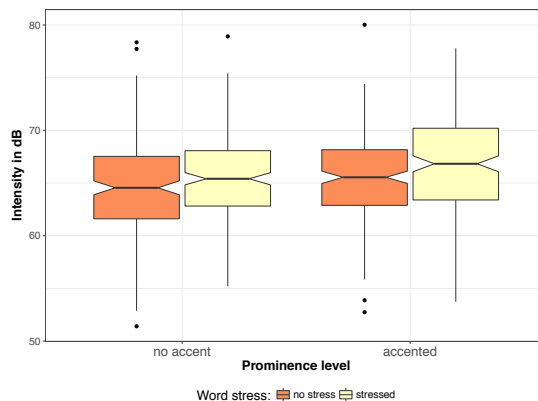


Figure 3: Intensity depending on prominence level.

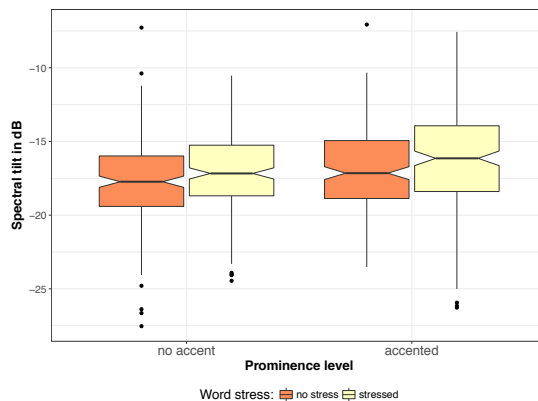


Figure 4: Spectral balance depending on prominence level.

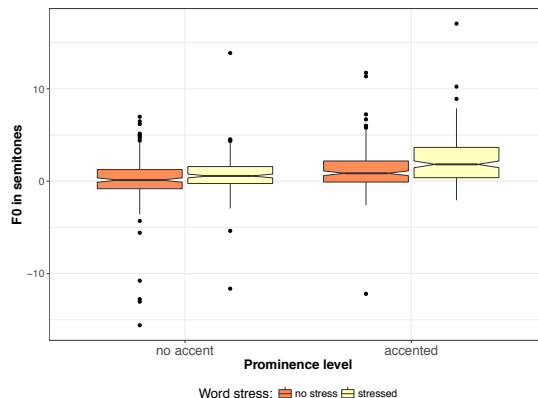


Figure 5: Fundamental frequency depending on prominence level.

Table 1: The results of all linear mixed-effects models. Per factor, the first line represents the estimate value and the second line is the standard error (in parentheses). Models 1 and 2 account for duration: in (1), the interaction of stress and accent was removed from the RE structure, in both by-item and by-subject slopes. In all models, accent has a significant effect, in models 3-5, part of speech plays a significant role, and it is only in models for duration that both the stress*accent interaction and stress have a significant impact.

	Dependent variable:				
	Duration (log)		Intensity (max.)	F0 (max.)	Spectral balance
	(1)	(2)	(3)	(4)	(5)
Stress	0.070*** (0.020)	0.071*** (0.020)	-0.824 (1.104)	-0.273 (0.631)	-0.404 (0.607)
Accent	0.063*** (0.020)	0.063*** (0.011)	1.211*** (0.192)	1.349*** (0.307)	0.913*** (0.138)
S*A Interaction	0.099*** (0.022)	0.098** (0.047)	0.254 (1.426)	0.309 (1.082)	0.034 (1.025)
Part of speech	-0.010 (0.023)	-0.011 (0.023)	2.572** (1.196)	1.443** (0.639)	1.586** (0.702)
Intercept	4.786*** (0.028)	4.786*** (0.028)	65.938*** (0.704)	1.236*** (0.214)	-16.784*** (0.555)
Observations	808	808	808	801	808
Log Likelihood	314.454	317.679	-2,036.448	-1,760.583	-1,763.259
Akaike Inf. Crit.	-610.908	-615.358	4,092.896	3,543.165	3,546.518
Bayesian Inf. Crit.	-568.657	-568.413	4,139.842	3,594.710	3,593.464

Note:

*p<0.1; **p<0.05; ***p<0.01

phrase-level prominence contradicts the claims of H&H theory. However, this cue richness may have two independent reasons: (1) the various prominence-lending acoustic correlates are not independent and “cue redundancy” may simply be the result of local hyperarticulation, and (2) the communicative situations are dynamically changing and might give rise to various optimal strategies of realizing prominence. For instance, a noisy environment may leverage F0 or duration as more suitable cues to prominence expression than intensity. Thus, “cue redundancy” enables speakers to fine-tune their strategy in a way satisfying both listener’s needs and production economy.

Our research provides further evidence that the acoustic expression of prominence is modulated by its distinctive or functional load (cf. [1]). It would be interesting to investigate further languages where word stress is free and strongly distinctive, or further languages with fixed word stress to see if prominence marking is in fact governed by its semantic value. Apart from increasing our knowledge on prosodic typology, such investigations may also be interesting for second language education. L2 learners are prone to generalize phonetic realizations from their native language to the L2 [34, 35, 36]. Thus, a false prosodic transfer between typologically distant languages could be reduced by instructing the learners about the existing contrast among prominence realizations.

5. Acknowledgments

We would like to warmly thank our colleagues Jolanta Bachan, Maciej Karpiński, and Katarzyna Klessa at Adam Mickiewicz University Poznań, who let us use their recording studio and helped with finding the participants, and Zofia Malisz, who helped with the experimental design.

6. References

[1] Á. Szalontai, P. Wagner, K. Mády, and A. Windmann, “Teasing apart lexical stress and sentence accent in Hungarian and Ger-

man,” in *Tagungsband der 12. Tagung Phonetik und Phonologie im deutschsprachigen Raum*, 2016.

- [2] D. B. Fry, “Experiments in the perception of stress,” *Language and speech*, vol. 1, no. 2, pp. 126–152, 1958.
- [3] Y. Hasegawa and K. Hata, “Fundamental frequency as an acoustic cue to accent perception,” *Language and Speech*, vol. 35, no. 1-2, pp. 87–98, 1992.
- [4] 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.
- [5] P. Wagner, A. Origlia, C. Avesani, G. Christodoulides, F. Cutugno, M. D’Imperio, D. Escudero Mancebo, B. Gili Fivela, A. Lacheret, B. Ludusan, H. Moniz, and A. Ni Chasaide, “Different parts of the same elephant: A roadmap to disentangle and connect different perspectives on prosodic prominence,” in *Proceedings of the 18th International Congress of Phonetic Sciences*, Glasgow, UK, 2015.
- [6] G. Dogil and B. Williams, “The phonetic manifestation of word stress,” *Word prosodic systems in the languages of Europe*, vol. 4, pp. 273–334, 1999.
- [7] M. Gordon and T. Roettger, “Acoustic correlates of word stress: A cross-linguistic survey,” *Linguistics Vanguard*, vol. 3, no. 1, 2017.
- [8] G. Corbett and B. Comrie, *The Slavonic Languages*. Routledge, 2003.
- [9] A. Eschenberg, “Polish Narrow Focus Constructions,” in *Topic and focus: cross-linguistic perspectives on meaning and intonation*. Dordrecht, The Netherlands: Springer, 2008, pp. 23–40.
- [10] Z. Saloni, *Składnia współczesnego języka polskiego*. Państwowe Wydawnictwo Naukowe, 1998.
- [11] S. Szober, *Gramatyka języka polskiego*. Warszawa: Państwowe Wydawnictwo Naukowe, 1957.
- [12] J. Rubach and G. E. Booij, “A grid theory of stress in Polish,” *Lingua*, vol. 66, no. 4, pp. 281–320, 1985.
- [13] M. Dłuska, *Fonetyka polska*. Warszawa: Państwowe Wydawnictwo Naukowe, 1950.
- [14] 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, p. 439, 2012.

- [15] L. Dukiewicz, *Intonacja wypowiedzi polskich*, ser. Prace Instytutu Języka Polskiego. Wrocław [u.a.]: Ossolineum, 1978, vol. 30.
- [16] W. Jassem, *Akcent języka polskiego*, ser. Prace językoznawcze. Wrocław: Ossolineum, 1962, vol. 31.
- [17] A. Nagórko, *Zarys gramatyki polskiej*. Państwowe Wydawnictwo Naukowe, 1997.
- [18] L. Newlin-Lukowicz, "Polish stress: Looking for phonetic evidence of a bidirectional system," *Phonology*, vol. 29, no. 02, pp. 271–329, 2012.
- [19] A. Bell, "Accent placement and perception of prominence in rhythmic structures," *Studies in stress and accent*, vol. 4, pp. 1–13, 1977.
- [20] D. Ostaszewska and J. Tambor, *Fonetyka i fonologia współczesnego języka polskiego*, 2nd ed. Warszawa: Wydawnictwo Naukowe PWN SA, 2012.
- [21] B. Wierzchowska, *Opis fonetyczny języka polskiego*. Państwowe Wydawnictwo Naukowe, 1967.
- [22] M. Dłuska, *Prozodia języka polskiego*. Warszawa: Państwowe Wydawnictwo Naukowe, 1976.
- [23] B. Łukaszewicz and B. Rozborski, "Acoustic correlates of word stress in child and adult Polish and the acquisition of phonological rhythm," *Generative Linguistics in Poland*, vol. 6, 2008.
- [24] K. Crosswhite, "Spectral tilt as a cue to word stress in Polish, Macedonian, and Bulgarian," in *Proceedings of the 15th International Congress of Phonetic Sciences*, Barcelona, Spain, 2003, pp. 767–770.
- [25] Z. Malisz and P. Wagner, "Acoustic-phonetic realisation of Polish syllable prominence: a corpus study," in *Rhythm, melody and harmony in speech. Studies in honour of Wiktor Jassem.*, 2012, vol. 14.
- [26] T. Roettger and M. Gordon, "Methodological issues in the study of word stress correlates," *Linguistics Vanguard*, vol. 3, no. 1, 2017.
- [27] Z. Malisz and M. Żygis, "Special Issue: Slavic Perspectives on Prosody," *Phonetica*, vol. 73, no. 3-4, pp. 155–162, 2017.
- [28] A. J. Szwedek, *Word order, sentence stress and reference in English and Polish*. Linguistic Research, 1976, vol. 10.
- [29] P. Wagner, J. Trouvain, and F. Zimmerer, "In defense of stylistic diversity in speech research," *Journal of Phonetics*, vol. 48, pp. 1–12, 2015.
- [30] R Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, 2017. [Online]. Available: <https://www.R-project.org/>
- [31] D. Bates, R. Kliegl, S. Vasishth, and H. Baayen, "Parsimonious mixed models," *arXiv preprint arXiv: 1506.04967*, 2015.
- [32] W. Jassem and D. Gibbon, "Re-defining English accent and stress," *Journal of the International Phonetic Association*, vol. 10, no. 1-2, pp. 2–16, 1980.
- [33] B. Lindblom, "Explaining phonetic variation: A sketch of the H&H theory," in *Speech production and speech modelling*. Springer, 1990, pp. 403–439.
- [34] J. E. Flege, "Second language speech learning: Theory, findings, and problems," *Speech perception and linguistic experience*, pp. 233–277, 1995.
- [35] B. Grzeszczakowska-Pawlikowska, "Probleme beim Rhythmuswerb-Ausgangssprache Polnisch und Zielsprache Deutsch," *Zeitschrift für Interkulturellen Fremdsprachenunterricht*, vol. 12, no. 2, 2007.
- [36] K. Nimz, *Sound perception and production in a foreign language: Does orthography matter*. Universitätsverlag Potsdam, 2016.