ACOUSTIC CUES OF PROSODIC BOUNDARIES IN GERMAN AT DIFFERENT SPEECH RATE

Marzena Żygis[1], John M. Tomlinson, Jr.[2], Caterina Petrone[3], Dominik Pfütze[4]

[1,2]Leibniz-Centre General Linguistics, Berlin; [3]CNRS, Aix-en-Provence; [4]Potsdam University

zygis@leibniz-zas.de, tomlinson@leibniz-zas.de, caterina.petrone@lpl-aix.pl, dpfuetze@uni-potsdam.de

ABSTRACT

This paper reports results of a production study on prosodic phrase boundaries at different speech rates. Our research question was to examine which acoustic cues (F0, intensity and duration) change as a function of speech rate both before and after prosodic phrase boundaries. We investigated identical sentences that differed in the placement of the prosodic boundary, i.e. before or after the critical word. Based on the data of 25 German speakers, duration was the main acoustic cue that speakers altered when marking prosodic boundaries at different speech rates. More specifically, words which appear before the boundary had longer durations, similarly to pauses at phrase-final boundaries, even if the speech rate increases. Intensity does also play a role: in faster speech words are produced with a higher intensity. Finally, F0 max was not changed on the critical word but rather it altered in comparison to F max of the preceding word.

Keywords: prosodic phrase boundaries, acoustic cues, German, speech rate

1. INTRODUCTION

For intonation languages like German, prosodic boundaries play an essential role in understanding spoken language in real time communication. They signal syntactic constituents and are critical for disambiguating syntactic structure and understanding meanings that arise from this disambiguation ([11]).

For example, the meaning in the sentences in (1) and (2) is disambiguated by a prosodic phrase boundary (presumably IP boundary) at the end of a parenthetical construction indicated by a comma.

(1) Anja hat, sagte Emily **auch,** eine Torte gebacken.
‘Anja has, said Emily also, baked a cake’.

(2) Anja hat, sagte **Emily**, auch eine Torte gebacken.
‘Anja has, said Emily, also baked a cake’.

As shown in previous studies, German speakers employ different acoustic cues to mark prosodic phrase boundaries. For instance, as reported in [3] and [4] names preceded by strong IP boundaries are characterized by a longer duration (measured as duration of a whole name) and f0 upstep. Furthermore, as shown in [7], phrase boundaries (IP) excerpted from the Kiel Corpus of German spoken language, are signaled by f0 reset or separating contour (74%), final lengthening (66.2%), and pauses (38.3%).

However, the studies on German phrase boundaries are based either on lists, i.e., speakers utter sequences of three and/or four names with different syntactic structure, e.g. *(Nino und Willi) oder Mila* “(Nino and Willi)” or Mila vs. *Nino oder (Willi und Mila)* “Nino or (Willi and Mila)” [4], see also [3], [9], [12] or they take into account uncontrolled sentences from spontaneous speech corpora [2], [6], [7], [8].

In our investigation we utilize the mode of presentation that does not use list-based stimuli but rather sentences as the ones presented in (1) and (2). In doing so, we can better test whether the role of different cues in boundary marking might have been task dependent and at the same time provide more experimental control in terms of syntactic and prosodic structure than spontaneous speech.

One further way we would like to investigate the acoustic cues involved in signaling prosodic boundaries is the influence of speech rate on the choice of acoustic cues. We predict that more robust acoustic cues should be retained at higher speech rates, whereas less robust cues will not.

Although it has been occasionally mentioned for German that speech rate might influence the prosodic boundaries [7], no study according to our knowledge has investigated this question in detail.

Our task, which allows more control over how productions are elicited productions, provides a new window into the following research questions:

1) What are the acoustic correlates of prosodic phrase boundaries in German?
2) To what extent do the acoustic cues change as function of the varying speech rate?

2. EXPERIMENT

2.1. Items

We manipulated our items for prosodic word boundaries appearing after and before a critical item. These included: *auch* “also”, *wohl* “well”, *trotzdem* “despite”, and *immerhin* “at least”. *Auch* and *wohl* are monosyllabic words, *trotzdem* is a bisyllabic word and *immerhin* is a three-syllabic word.

We created sixteen different experimental sentences (four with each critical item) as well as twelve filler sentences. An experimental item was embedded into two sentences: one with a prosodic boundary before the critical word, see example in (3b), and one with the prosodic boundary after the critical word, see example in (4b). A context sentence was displayed on the screen prior to the experimental item; see (3a) and (4a), respectively. Note that the boundary is indicated by a comma.

(3a) Frank hat einen Artikel geschrieben.
‘Frank wrote an article.’

(3b) Olli hat, sagt Emily], [**auch** einen Artikel geschrieben.
‘Olli has, says Emily, also written an article.’

(4a) Olli hat, sagt Peter, einen Artikel geschrieben.
‘Olli has, says Peter, written an article.’

(4b) Olli hat, sagt Emily **auch,**] [einen Artikel geschrieben.
‘Olli has, says Emily also, written an article.’

The syntactic construction of the tested sentences was identical, i.e. a host sentence circumventing a parenthetical construction. The critical word was either inside the parenthetical (as in (4)) or outside it (as in (3)) and this was indicated by the placement of a comma. In order to present an item more than once during the experiment, we added four different proper names to the subject position of each experimental item, hence creating four different versions of an item. We also varied the object and the past participle of the host sentence. One of these versions was assigned to one of the levels of our two experiment conditions (comma before+normal, comma after+normal, comma before+fast, comma after+fast).

The syntactic forms of filler sentences varied, though all used other syntactic embeddings at the same positions of the parenthetical, e.g. relative clauses, infinitive clauses, and others.

In summary, each participant had 33 test trials and 36 fillers. The experiment lasted between 15 and 20 minutes.

2.2. Participants

25 native speakers of German (16 female; age range 18 to 60, mean: 31) took part in the production experiment.

2.3. Procedure

The participants received either one of two experimental lists from a Latin-squared design and these were presented using Eprime 2.0. They were asked to read the sentences (experimental items) either at the normal or fast speech rate. Participants were presented with two blocks: either the *normal* or *fast* speech rate block, so that the participant did not have to switch between different speech rates all the time. Overall, participants saw only two versions of a single item (one in each block), though were exposed to all levels of the conditions throughout the experiment. Block order was counterbalanced. All items were presented in randomized order, though block type had a fixed order, e.g. either normal before fast or fast before normal.

The sentences as well as their preceding contexts were displayed on the backdrop of a white screen. Participants started a trial by pressing the space bar and read the context sentence. After reading the context sentence, participants again pressed the space bar to read the test sentence and ended the trial (after reading the test sentence) by also pressing the space bar. The context sentences remained on the screen above the test and filler sentences.

2.4. Annotation

All tested sentences were annotated in PRAAT (version 6.0.43, [1]). First, word boundaries were marked and by means of a script the following parameters were excerpted: (a) duration of the critical word of the tested sentence, (b) duration of the pauses before and after the critical word, (c) duration of each phrase (3 phrases), (d) max and mean intensity of the critical word, (e) max F0 of the critical word and the word preceding it.

2.5. Statistics

All statistical analyses were conducted in the R Studio software (version 1.1.453 [10]). Linear mixed effect models were fitted to the data to examine the influence of the PROSODIC BOUNDARY [before the critical word, after the critical word], SPEECH RATE [normal, fast], and Sex [male, female] as well as the interaction of PROSODIC BOUNDARY and SPEECH RATE on duration, max intensity and F0 max of all the critical words.

In our models we included random structure: random intercepts for participants and items as well as their slopes for prosodic boundary and speech rate. The maximized models were tested against less complex models by means of likelihood ratio tests and the best fit model was taken as the final model.

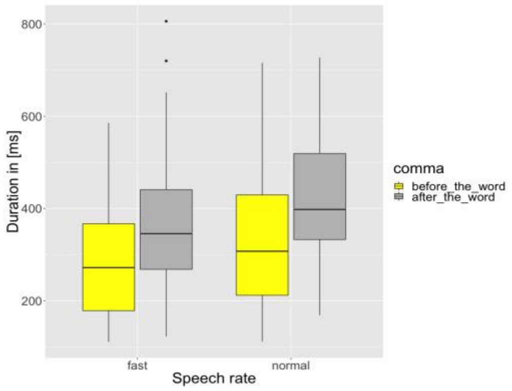
The results are based on 716 observations.

3. RESULTS

3.1. Duration

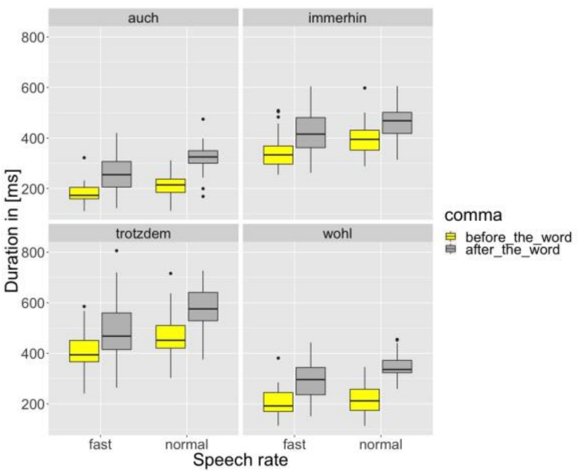
Our results show that duration is the decisive parameter in the production of prosodic boundaries in German. The words are significantly longer if they precede a boundary, i.e., at the end of the parenthetical construction (the comma is placed after the word, t= 9.80, p<.001). They are also shorter in fast speech (t=-8.82, p<.001). Despite the fact that fast speech reduces the duration of word, the words are longer when followed by boundaries and shorter when they follow the boundaries. The interaction speech rate (normal) and boundary (before the word) is significant (t=-3.61, p<.001). Figure 1 illustrates the results. Note that the term *comma* denotes prosodic boundary.

Figure 1: Duration of the critical words before and after the boundary at different speech rates



The same conclusion holds true for each individual word, as shown in Figure 2.

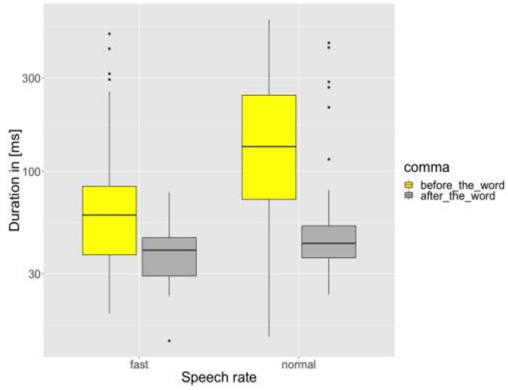
Figure 2: Duration of the individual words before and after the boundary at different speech rates



As far as pauses are concerned, they show a rather consistent pattern in dependence of the appearance of the prosodic boundary. Note that our results are based on 361 instances of a pause after the critical word and 355 before the critical word.

Figure 3 presents the results of pause duration appearing before the critical word. If there is a prosodic boundary before the critical word, its duration is significantly longer in normal speech than in fast speech (t= 5.03, p<.001). Pauses are also shorter when comma appears after the word in both fast (t= -6.58, p<.001) and normal speech (t=-2.81, p<.01). The difference in pause duration is also larger in normal than in fast speech (the interaction speech rate*comma is significant: t= -3.48, p<.001).

Figure 3: Duration of pauses before the critical word depending on prosodic boundary and speech rate



If we examine pauses after the critical word we obtain an inverse picture: pauses are significantly longer when there is a prosodic boundary after a word. In contrast to pauses preceding the critical items, speech rate does not influence the results: the difference in pause duration between preceding and following boundary was not significant. Figure 4 illustrates the results.

The difference in duration between words before and after the boundary depends on speech rate as it was higher in normal than in fast speech (the interaction between speech rate and comma is significant t=3.66, p<.001).

3.2. Intensity

The results for intensity show that maximum intensity was significantly lower in normal speech then in fast speech (t=-3.01, p<.01). Furthermore, prosodic boundary also influences the intensity: it is lower when the boundary appears before the word (t=-2.33, p<.05). This is shown in Figure 5. The interaction of speech rate and the prosodic boundary remains not significant. If we take into account mean intensity, it is only significantly lower in normal speech (t=-5.42, p<.001).

Figure 4: Duration of pauses after the critical word in dependence of prosodic boundary and speech rate

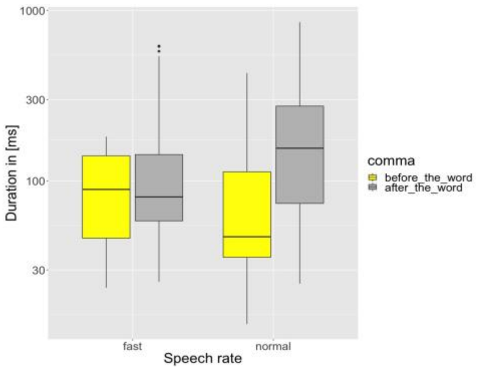
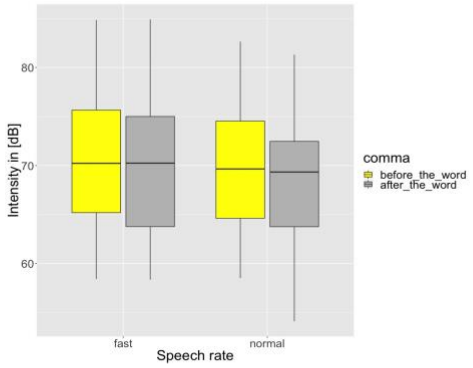


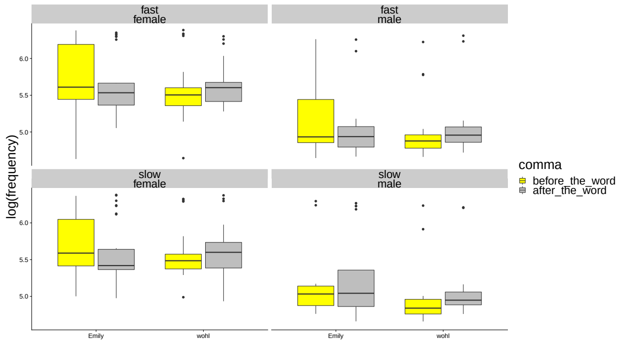
Figure 5: Max intensity of the words placed before and after comma at different speech rates



3.3. Fundamental frequency

Preliminary results on f0 max are shown in Figure 6. Note that in order to impose more control on F0 results in the process of its automatic extraction we extracted one monosyllabic word, i.e. *wohl* "well" and the preceding word *Emily* (181 observations).

Figure 6: Max F0 of the words placed before and after the prosodic boundary at different speech rates



Apart from the obvious male/female distinction (t=-5.30, p<.001), we found no significant difference on f0 max on *wohl* across the two prosodic conditions. Furthermore, when the comma is before, the f0 max is higher on *Emily* than on *wohl*. When the comma is after the critical word, the difference in f0 max between *Emily* and *wohl* is more reduced. However, the interaction is at the level of statistical tendency (t=-1.91, p=.056). Speech rate does not affect F0 max.

4. DISCUSSION AND CONCLUSIONS

Our study shows that the duration is the most preferred cue in marking the prosodic boundaries. If speakers increase their speech rate, the duration of words decreases but the contrast between presence vs. absence prosodic boundary is still maintained: words followed by boundaries are longer than those which follow the boundaries. This effect is, as we hypothesize, due to final lengthening which seems to be immune to varying speech rate. Pauses also signal prosodic boundaries but they are not as strong cues as duration (lengthening) since they do not play a role when the speech rate is increased.

Furthermore, intensity decreases towards the boundary. It is also lower in normal than in fast speech but it not crucial for marking boundaries in fast speech.

No strong effects of prosodic boundary were found on f0. The lack of effects might be due to our methodological choice, in that automatic measure of f0 max is not entirely appropriate to capture the intonational changes across prosodic conditions.

Preliminary phonological analyses suggest that, in the case of a comma after the critical word, a high f0 plateau was created spanning from a H* on the accented syllable of *Emily* till a H% on the critical word. On the other side, when the comma was before the critical word, the critical word contained an accent and was under focus. Hence, though f0 max on the critical word did not vary across conditions, the phonological specification of the critical words and the previous material changed. Hence, a more detailed phonological analysis should be done to better understand the contribution of f0 as a cue to phrasing. Note also that some inter-speaker variation in the use of f0 was observed, which might be in line to some extent with the conclusion drawn in [9:88] where "the f0 was the most variable cue across speakers […] depending on the speakers' choice."

Finally, we need to extend our analyses to the remining three critical words and better understand the relationship between the semantics of the adverbs and the interaction with prosody.

Overall, our results are also consistent with perception studies on prosodic boundaries in German which show different role of acoustic cues in boundary perception with a preferred reliance on duration cues over pitch changes in both adult [9] and children perception [5].

Acknowledgements

This research has been supported by the Bundes-ministerium für Bildung und Forschung (BMBF, Ger-many) Grant Nr. 01UG1411 and Leibniz Association. We also thank Hubert Truckenbrodt for help with stimuli construction and Jörg Dreyer for technical support. All errors are our own.

5. REFERENCES

[1] Boersma, P., Weenink, D. 2018. Praat: doing phonetics by computer [Computer program]. Version 6.0.43, retrieved 8 September 2018 from http://www.praat.org/

[2] Döring, S. 2010. *Parenthesen im Deutschen. Untersuchung syntaktischer und prosodischer Aspekte*. PhD thesis, University of Leipzig.

[3] Féry, C., Kentner, G. 2010. The prosody of embedded coordinations in German and Hindi. *Proceedings of 5th International Conference on Speech Prosody*. Chicago.

[4] Kentner, G., Féry, C. 2013. A new approach to prosodic grouping. *The Linguistic Review* 30(2), 277–311.

[5] Männel, C., Friederici, A.D. 2016. Neural correlates of prosodic boundary perception in German preschoolers: If pause is present, pitch can go. *Brain Research* 1632, 27-33.

[6] Peters, B. 2005. Weiterführende Untersuchungen zu prosodischen Grenzen in deutscher Spontansprache. In: Kohler, K.J., Kleber, F., Peters, B. (eds.). *Prosodic Structures in German Spontaneous Speech* 35a. *AIPUK* 203–345.

[7] Peters, B., Kohler, K. J., Wesener, T. 2005. Phonetische Merkmale prosodischer Phrasierung in deutscher spontansprache. In: Kohler, K.J., Kleber, F., Peters, B. (eds.). *Prosodic Structures in German Spontaneous Speech* 35a. *AIPUK*, 143–184.

[8] Peters, J. 2006. Syntactic and prosodic parenthesis. *Proceedings of the 3rd International Conference on Speech Prosody*, 2-5 May 2006, Dresden.

[9] Petrone, C., Truckenbrodt, H., Wellmann, C., Holzgrefe-Lang, J., Wartenburger, I., Höhle, B. 2017. Prosodic boundary cues in German: Evidence from the production and perception of bracketed lists. *Journal of Phonetics* 61, 71–92.

[10] RStudio Team (2018). RStudio: Integrated Development for R. RStudio, Inc., Boston, MA. Version 1.1.453.

[11] Watson, D., & Gibson, E. 2004. The relationship between intonational phrasing and synactic structure in language production. *Language and Cognitive Processes* 19(6), 713–755.

[12] Wellmann, C., Holzgrefe, J., Truckenbrodt, H., Wartenburger, I., Höhle, B. 2012. How each prosody boundary cue matters: Evidence from German infants. *Frontiers in Psychology* 3, 1-13.