PROSODIC BOUNDARY STRENGTH AND PHRASE-INITIAL SYLLABLE DURATION

Gerrit Kentner[1,2], Isabelle Franz[2,3], Christine A. Knoop[2], Winfried Menninghaus[2]

[1]Goethe University Frankfurt, [2]Max-Planck Institute for Empirical Aesthetics, [3]HS Gesundheit Bochum
kentner@lingua.uni-frankfurt.de, isabelle.franz@hs-gesundheit.de, christine.knoop@ae.mpg.de, w.m@ae.mpg.de

ABSTRACT

Research on speech timing has produced conflicting results regarding the duration of phrase-initial syllables. The articulation of these syllables has been shown to be subject to *initial strengthening*, suggesting that phrase-initial segments (and by transitivity: phrase-initial syllables) have longer durations than comparable segments in other positions. In contrast, the notion of *anacrusis* holds that phrase-initial syllables are shortened, at least when these syllables are non-prominent. We explore a large sample of German literary prose texts read aloud, with a focus on the duration of phrase-initial syllables. Comparing five predicted levels of boundary strength (level 0: no break; level 1: simple phrase break; level 2: short comma phrase; level 3: long comma phrase; level 4: sentence boundary), we found evidence for initial strengthening and for a higher degree of strengthening after stronger boundaries. In contrast, we don't find systematic evidence for anacrusis.

Keywords: syllable duration; prosodic boundary; anacrusis; initial strengthening.

1. INTRODUCTION

Speech streams are divided into larger constituents (e.g., utterances, sentences) which contain smaller constituents (phrases within sentences, words within phrases, syllables within words). This hierarchical prosodic grouping of constituents underlying the sequence of speech sounds manifests itself in the rendition of prosodic boundaries, with stronger boundaries set between the higher-order constituents, and weaker boundaries set between the smaller embedded constituents. The distribution and varying strength of the prosodic boundaries in spoken texts reflect both coherence and division between speech chunks, and give rise to the temporal rhythm of the text. Prosodic boundaries are marked by a variety of cues [1-3]: The most obvious phonetic expression of a prosodic boundary is a pause between two chunks of speech. Strong boundaries are generally likely to be realized with a pause, and pauses at strong boundaries tend to be longer than pauses at weak boundaries [4]. However, not all prosodic boundaries result in a pause. Apart from pauses, prosodic boundaries are generally marked by a decrease in loudness, by boundary tones (pronounced excursions of the pitch contour preceding a boundary), and by a slow-down in speech rate approaching the boundary (pre-boundary lengthening) [5] (but see [6] for evidence of pre-boundary shortening at stronger prosodic breaks).

The effect of prosodic boundaries on the duration of the following syllable(s) is disputed: The π gesture model holds that phrase-initial segments are subject to the same "prosodically induced local slowing" of articulatory movements as pre-boundary syllables [7]. This slow-down is predicted to be greater the greater the boundary strength is. Other researchers report an articulatory strengthening of phrase-initial segments that manifests itself in an increase in the contact area between active and passive articulators [8]. Both slowing and strengthening (while conceptually quite different) result in a durational increase which – according to most researchers – mainly affects the initial segment, but has been shown to extend to the entire syllable [9].

In contrast to these findings, other researchers suggest a speed-up on pre-ictic phrase-initial syllables in English, i.e., relatively short duration of phrase-initial syllables when these syllables are unaccented. This phenomenon is termed *anacrusis* [10]. We are not aware of studies probing effects of boundary strength on anacrusis. Whereas *initial strengthening* is said to merely affect the initial post-boundary segments, the domain of anacrusis is wider, as it involves the entire pre-ictic stretch, i.e., potentially more than a single syllable. The conflicting results and the different domains motivate a fresh look at how boundary strength and prominence affect the duration of phrase-initial syllables. Here, we explore a large corpus of read aloud German prose texts to shed light on this issue.

2. EXPERIMENT: METHODS

We applied the coding manual by [11] which – on the basis of punctuation and certain syntactic features of written texts – guides the annotation of several degrees of prosodic boundaries. Here, we apply this

tool, together with a corpus of read aloud prose texts, to address the following questions:

- What are the effects of prosodic boundary strength on the duration of post-boundary, phrase-initial syllables?
- What effect does the presence/absence of a pause and pause duration have on post-boundary syllable duration?
- What effect does the degree of prominence of the phrase-initial syllable have on post-boundary syllable duration, and does it interact with boundary strength?

2.1. Recordings, annotation of data

We evaluated recordings of eight professional speakers (four male, four female). At the time of the recordings, all speakers were enrolled in a university program for rhetoric and professional stage reading.

2.1.1. Recordings

Each speaker was seated in a sound-attenuated recording booth with a Neumann U87 Ai Studio microphone placed ~30 cm from their mouth, and read aloud four prose text samples (excerpts from novels by Fontane, Goethe, Kafka, and Kleist). All speakers were familiar with the texts and had diligently prepared the reading before the recordings took place. The excerpts contained roughly ~1500-1800 words each. When slips of the tongue or disfluencies occurred during the recordings, speakers were asked to re-read the affected sentence or paragraph. The faulty segments were later erased and replaced by the corrected renditions, with the original pauses that preceded the affected sentences preserved. In total, the recordings comprised 52697 spoken words in roughly 6 hrs of speech.

2.1.2. Prosodic boundary and prominence prediction

We predicted prosodic boundaries using a coding system for the annotation of prominence and phrasing of written prose texts [11]. For the transitions between words, this system assumes four degrees of prosodic boundary strength: the strongest prosodic boundary stipulated in this annotation protocol is a sentence boundary (level 4, determined by sentence-final punctuation marks). Commas demarcate weaker breaks: Commas preceded by three or more words are assigned level 3 (long comma phrase), whereas short comma phrases (with only one or two preceding words) are assigned level 2 breaks. Level 1 boundaries are word transitions that do not come with punctuation marks but still afford prosodic boundaries due to the syntactic configuration and/or to split up exceedingly long constituents. All other word transitions were assumed to be not conducive to a prosodic break and therefore considered neutral (level 0).

Apart from boundary annotation, the coding system is used for the annotation of predicted syllable prominence. The coding system assumes, apart from further minor distinctions, the following prominence levels: schwa syllables, unstressed syllables with full vowels, lexical stress, phrasal/sentence accent.

2.2. Measurements

The recordings were automatically segmented for words and syllables using Web-MAUS [12]. Using praat [13], we extracted duration values for each annotated segment in the text, i.e., for syllables as well as pauses. In order to correlate post-boundary syllable duration with boundary strength, all word-initial syllables were coded according to the strength of the preceding boundary.

3. RESULTS

In order to evaluate potential effects of initial strengthening and/or anacrusis, we examined the duration of all word-initial syllables and compared those following a neutral word-boundary with syllables following various degrees of prosodic boundaries. In doing so, we consider potential effects of pausing at the boundary, as well as prominence and segmental content of the syllable.

The plot in Figure 1 shows post-boundary syllable durations broken down by boundary strength (level 0 to level 4) and presence/absence of a pause (top panel: no pause at break; bottom panel: post-pausal syllables). This plot does not suggest a systematic increase in syllable duration along the boundary strength scale. Instead, for the phrase-initial syllables not preceded by a pause (top panel), there is a noticeable decrease in duration for boundary strength level 1 (minor phrase break) when compared to neutral boundaries. Only for level 4 breaks, there is an evident durational increase relative to neutral word boundaries. Hardly any change in syllable duration along the boundary strength scale is detectable for post-pausal syllables.

The plot in Figure 2 shows post-boundary syllable durations broken down by boundary strength and prominence (top panel: unaccented syllables, bottom panel: accented syllables). This plot reveals a stepwise increase in syllable duration along the boundary strength scale for unaccented syllables (top panel Figure 2). For accented syllables, the increase appears to be less marked.

To model the effects of boundary strength, pausing and prominence on syllable duration, we fitted a generalised mixed effects model, assuming a gamma distribution with log link to account for the non-normal distribution of syllable duration. Along with the main effects *boundary strength*, *pausing*, and *prominence*, we considered effects of the *number of segments per syllable*, and the interactions between i. *pause duration and boundary strength*, and ii. *prominence and boundary strength*.

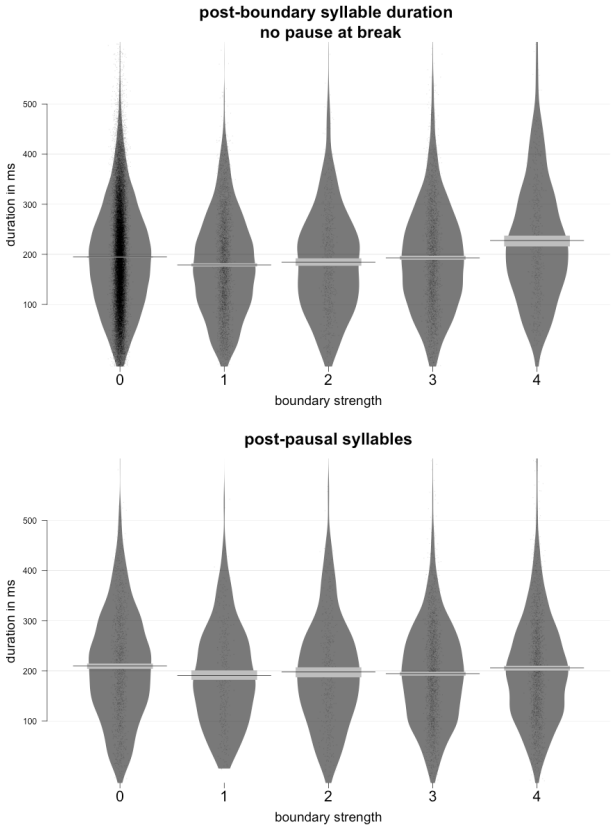

Figure 1: Duration of word-initial syllables broken down by boundary strength (0: neutral word boundary; 1: minor phrase break; 2: small comma phrase; 3: long comma phrase; 4: sentence boundary) and presence/absence of a pause preceding the break (top panel: no pause at break; bottom panel: syllables preceded by a pause).

The model confirms significant effects for boundary strength levels 2-4 on syllable duration (no significant effect of boundary level 1); significant main effects of prominence and number of segments per syllable; a significant negative interaction between pause duration and boundary strength (the effect of boundary strength on post-boundary syllable duration is weakened for post-pausal syllables); a significant negative interaction between boundary strength and prominence (the effect of boundary strength on post-boundary syllable duration is weakened for accented syllables). The coefficients of interest are tabulated in Table 1.

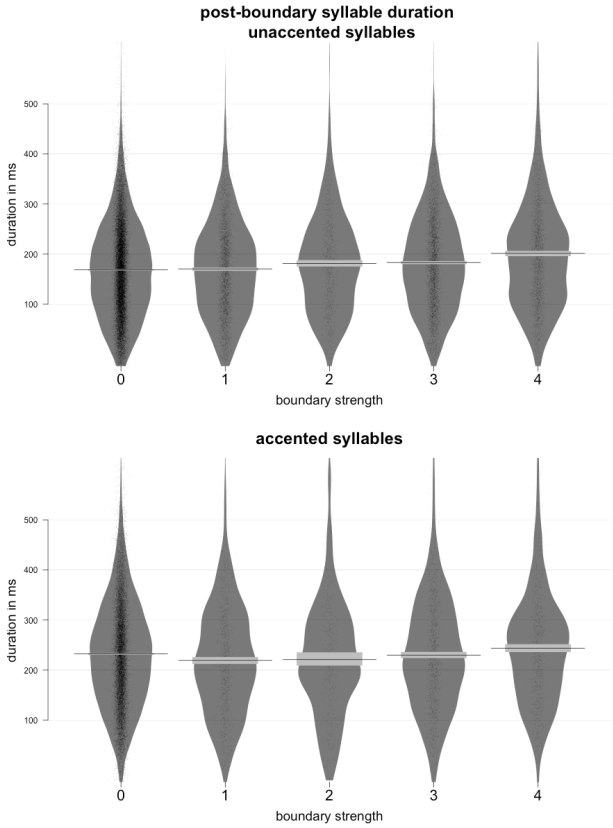

Figure 2: Duration of word-initial syllables broken down by boundary strength and prominence (top panel: unaccented syllables; bottom panel: accented syllables).

Coefficient	Estimate	SE	t-value
Boundary 1	0.007	0.005	1.2
Boundary 2	0.058	0.01	5.7***
Boundary 3	0.062	0.006	10.6***
Boundary 4	0.182	0.01	17.6***
Pause	0.005	0.006	0.9
Accent	0.181	0.004	51.3***
Segments/Syll	0.286	0.002	163.3***
Bndry : Pause	-1.3e-5	2.4e-6	-5.3***
Bndry : Accent	-0.031	0.003	-11.4***

Table 1: Model coefficients of interest (see main text for details).

In sum, whereas inspection of the raw data suggests shortening of phrase-initial syllables in some contexts (level 1 boundaries, see, e.g., upper panel of Figure 1), the statistical model does not corroborate this finding. There is especially no evidence for a shortening of pre-icitic phrase-initial syllables, as predicted in [10] for English. Instead, there is considerable evidence for initial strengthening.

4. DISCUSSION AND CONCLUSION

The results of the analysis reveal longer durations on phrase-initial syllables, providing evidence for initial strengthening. In addition, syllable duration correlates with boundary strength: sentence-initial syllables are longest on average; specifically, they are longer than comma-phrase-initial syllables, which in turn are longer than non-boundary related word-initial syllables. Furthermore, initial syllable duration is subject to an interaction between boundary strength and preceding pause duration, with the generally longer post-pausal syllables showing only little additional effects of boundary strength. Similarly, prominent syllables, which are inherently longer than the non-prominent ones, show little if any additional lengthening along the boundary strength scale. The clearest strengthening/lengthening effect, with increased lengthening along the boundary strength scale, is visible on unaccented syllables (Figure 2, top panel). This latter effect is in conflict with the concept of anacrusis which holds that unaccented or pre-ictic syllables are subject to shortening rather than lengthening.

Finally, we note a number of limitations: First, the current analysis does not provide information about effects on the level of individual segments; the corpus data lacks the control over the segmental content of the syllables. Also, we only investigate acoustic data – the articulatory dynamics leading to the increased duration of phrase-initial syllables is beyond the scope of this study. Finally, the lack of evidence in favour of anacrusis in German should not be taken to invalidate effects of anacrusis found in certain contexts in other languages.

5. REFERENCES

[1] Féry, C. (2017). *Intonation and prosodic structure*. Cambridge University Press.

[2] Ladd, D. R. (2008). *Intonational phonology*. Cambridge University Press.

[3] Cruttenden, A. (1997). *Intonation*. Cambridge University Press.

[4] Gee, J. P., Grosjean, F. (1983). Performance structures: A psycholinguistic and linguistic appraisal. *Cognitive Psychology*, 15(4), 411-458.

[5] Krivokapić, J. (2014). Gestural coordination at prosodic boundaries and its role for prosodic structure and speech planning processes. *Philosophical Transactions of the Royal Society B*, *369*(1658), 20130397.

[6] Kentner, G., Franz, I., Knoop, C.A., Menninghaus, W. (2023). The final lengthening of pre-boundary syllables turns into final shortening as boundary strength levels increase. *Journal of Phonetics*, 97, 101225.

[7] Byrd, D., Saltzman, E. (2003). The elastic phrase: Modeling the dynamics of boundary-adjacent lengthening. *Journal of Phonetics*, 31(2), 149–180.

[8] Cho, T. (2016). Prosodic boundary strengthening in the phonetics–prosody interface. *Language and Linguistics Compass*, 10(3), 120–141.

[9] de Souza, R. F. N. (2019). *The Interaction of Domain-initial Effects with Lexical Stress: Acoustic Data from English, Spanish, and Portuguese*, Ph.D. dissertation, The University of New Mexico, USA.

[10] Jassem, W., Hill, D., Witten, I. (1984). Isochrony in English speech: its statistical validity and linguistic relevance. In: Gibbon, D., Richter, H. (eds.), *Intonation, Accent and Rhythm*. Berlin: De Gruyter, 203-225.

[11] Franz, I., Knoop, C. A., Kentner, G., Rothbart, S., Kegel, V., Vasilieva, J., Methner, S., Scharinger, M., Menninghaus, W. (2022). Prosodic phrasing and syllable prominence in spoken prose: A validated coding manual. *OSF Preprints*. https://osf.io/h4sd5.

[12] Kisler, T., Reichel, U., Schiel, F. (2017). Multilingual processing of speech via web services. *Computer Speech & Language*, *45*, 326-347.

[13] Boersma, P., & Weenink, D. (2020). *Praat: Doing phonetics by computer* v 6.1. 21.