The relationship between boundary markers and audible inhalation in Hungarian read speech

Valéria Krepsz1,2, Anna Huszár1, Katalin Mády1

1Hungarian Research Centre for Linguistics
2Humboldt-Universität zu Berlin

huszar.anna@nytud.hu, krepsz.valeria@nytud.hu, mady.katalin@nytud.hu

Abstract

It is well known that prosodic features – next to meaning-modifying and meaning-clarifying functions – can also act as boundary markers, however, it is unclear that which prosodic cues and in what combination. The results of previous research, have yielded contradictory results as to which are the strongest, most frequently occurring, or most consistent prosodic cues, in part due to different languages and methods. Dissimilarly from most languages analyzed earlier, Hungarian is a left-headed head-edge prominence language. Therefore, the aim of the present study was to investigate the difference in the relationship of prosodic boundary markers and the audible inhalation between the two kinds of clauses in Hungarian read speech and according to the length of the first clauses of the sentences.

Based on 20 young adult male and female Hungarian speakers’ text and sentence reading, altogether 320 (coordinated, CO and subordinated, SUB) clause end positions were examined. Silent pauses, the appearance of high/low tones, creaky voice, and the realization of the audible inhalation (based on the agreement of 2 annotators) were labeled.

The results revealed differences between boundary marking in CO and SUB clauses: SUB clauses were more often realized with no marking, while CO boundaries were labeled with low tone and creaky voice. Besides, notable differences were also observed in the boundary marking cues after long and short clauses.

Index Terms: prosody, boundary markers, read text, sentence structure, audible respiration

1. Introduction

The definitions of boundary markers are diverse. According to one type of interpretation, emphasizing the separation function of elements, fluent speech “is broken into units” by boundaries. The other definitions suggest that the structure of spontaneous speech is indicated by prosodic boundary cues [1]. Examination of the features of speech (such as intonation, timing etc.) is greatly encumbered by the fact that these elements are multi-dimensional and contain several acoustic parameters that overlap in terms of form and function. In fact, none of the prosodic boundary markers per se is responsible for syntactic boundaries, as each has other additional functions, and it would indicate ambiguity even when appearing in the boundary position (based on read speech) [2].

Prosodic reflection of syntactic and semantic structure is not a universal property [3]. Previous research contradicted results regarding the most consistent prosodic feature of boundary marking: the results showed notable differences in different languages (such as Hindi, German [4], Slovenian [5], English [6] or French [7]). Dissimilarity was found in the realization of the given boundary markers, in their frequency and combination.

Beside language differences, the chosen methodology might affect the examined features: The most commonly used procedures for examining boundary markers are grouping tests (e.g. with names or numbers; cf. [4], production based on parentheses [8] or reported speech, and the examination of spontaneous speech.

Hirschberg and colleagues [9] found a relationship between quoted phrases and low intensity, but not for boundary tone. According to Peters and colleagues [8], the change of pitch and final lengthening occur most frequently before boundaries, while pauses are rare. In another study, Zygis and colleagues [10], found the following tendency based on 16 manipulated sentences: duration was the most preferred cue in boundary marking, pauses, while intensity and f0 showed weaker effect.

Several concepts have appeared (e.g. the Scopally Determined Boundary Rank, detailed description can be found in [11]) in the last decades to determine which groups of elements belong together and where and in what way the belonging can be expressed, for example by prosodic cues. Nowadays, the Proximity and Similarity model (built up on the previous conceptions) is considered to be the most comprehensive and acceptable, taking into account both the characteristics of the syntactic structure and the prosodic features [4], [11].

The linguistic units belonging to a given group based on the syntactic constituency show a closer proximity to each other, which is expressed by pitch and duration. The effect of proximity can operate in contrast with anti-proximity: proximity is achieved by significantly weakening the prosodic boundary cues, the elements of anti-proximity are realized with a prosodic distance between each other.

In the case of Similarity: Those linguistic units that are on the same level of embedding, are realized with a similar prosodic pattern, independent of the complexity of the given structure. So the linguistic units that are realized similarly belong together, and those linguistic units belong together that are realized with similar prosodic features.

Examining the form and function of boundary markers from another approach, other experiments found a primary interplay between prosodic cues and the pattern of respiration: inhalation often coincided with syntactic structural boundaries (in reading, this ratio was 88%; [12]; while in spontaneous speech the ratio was 70%; [13]). Additionally, deeper inhalation appeared at
sentence boundaries more frequently than at weaker boundaries (e.g. clauses).

Fuchs and colleagues [20] sought an answer to the question of how the length of the sentence and the syntactic complexity influences the breathing patterns in German speakers’ reading task. Results showed that the length of the sentences had a significant effect: longer pauses appeared after longer sentences, and deeper and longer inhalation came before longer sentences, however, the sentence-initial f0 peak showed no significant effect.

The interaction of respiration and prosodic planning has been investigated for English and German, but not for Hungarian. Hungarian prosody is typologically different from the prosody of Germanic languages. First, information structure in Hungarian is primarily expressed by word order, i.e. logical functions are linked to certain sentence positions. The position of focus is defined syntactically (it is immediately pre-verbal), while prosodic prominence marking plays only a secondary role and is partly optional [15]. Second, pitch accents and boundary tones are varied to express pragmatic, semantic and scopal meaning. Additionally, Hungarian is a language with left-headed prosody, both on the word and the sentence level.

The aim of the present study was to investigate the relationship between prosodic boundary markers and audible inhalation between two kinds of clauses in Hungarian read speech according to the length of the first clauses of the sentences. The question arises whether there is a difference in the realization of the examined boundary markers according to whether they separate subordinated or coordinated clauses in complex sentences.

Our hypotheses are the following: 1. There is a weaker boundary marking in the subordinated (SUB) clauses than in the coordinated (CO) complex sentences. 2. Audible inhalation occurs more often in the stronger marked clauses, thus in CO clauses, and less frequently in the SUB clauses.

2. Methodology

10 female and 10 male participants’ (age: 20–29 years) recordings were selected from the BEA Hungarian Spoken Language database (for details, see [17]). Speakers didn’t have any known speech or hearing disorders. Participants were asked to read 25 sentences and a short scientific text, including 7 CO and 9 SUB sentences. Complex sentences (coordinated, CO and subordinated, SUB) were examined regarding the structure of their intonational phrases (IP).

We split the sentences into two groups according to the length of the clause before the examined boundary: 1. short: 5–8 syll. (6.6±1.1 syll), 2. long: 12–42 syll (24±9.9 syll). The following 4 categories were analyzed: 1. CO-long (n = 6), 2. CO-short (n = 1), 3. SUB-long (n = 5), SUB-short (n = 4). In terms of the previous results, that showed that inhalation mostly coincided with the end of the sentence. Thus, in this study, only clause boundaries were analyzed, where the occurrence of inhalation is questionable, and it might have an interplay with prosodic boundary values. If there were prosodic boundary markers in other positions in the sentences, the boundary markers in other positions were excluded. The following features were labeled manually in Praat [18] between the two clauses (e.g. Fig. 1): whether the boundaries were realized with 1. pause, 2. high/low boundary tone, 3. creaky voice (whether creaky voice appeared in the last syllable of the clause), 4. audible inhalation.

3. Results

Based on the fact that pauses appeared most often at the boundary, the description of results will begin with that parameter.

74.7% of the clause boundaries were realized with pauses. The mean values in the two main groups are quite similar (CO: 73.5%, SUB: 71.05%). Based on the values, larger differences can be seen regarding the length of the clauses. The mean values for the occurrence of pauses are: CO-long: 87.0%, CO-short: 60.0%, SUB-long: 96.1%, SUB-short: 46.0%. 52.2% of the clauses were connected by inhalation: CO-long: 70.8%, CO-short: 25%, SUB-long: 61%, SUB-short: 20% (Fig.2).
Additionally, on average, 59.1% of the clauses were realized with a high tone. This ratio alternated as the following regarding the types of the clauses: CO-long: 70.7%, CO-short: 42.5%, SUB-long: 75.3%, SUB-short: 44% (Fig. 3).

Furthermore, the occurrences of boundary cues were much lower for low tone and creaky voice: 1.9% of all the cases were realized with low tone: low tone only occurred in the case of CO-long sentences (32.1%). Moreover, 2.5% of clauses boundaries realized with creaky voice: CO-long: 5.7%, CO-short: 0%, SUB-long: 1.3%, SUB-short: 1% (Fig. 4).

On average, the most frequent combination of the prosodic cues were pause and high tone (pause+H). A remarkable difference can be seen between the long and short first clauses: in the case of the long first clauses, pause+H is much more frequent (CO-long: 66.4%, SUB-long: 72.4%), while in the case of short first clauses the occurrences are much lower (CO-short: 40%, SUB-short: 31.3%).

In case of the short first clauses, nearly one third of the boundaries were prosodically unmarked (CO-short: 25%, SUB-short: 41.4%). The lack of prosodic marking was very rarely observed in the case of the long ones (CO-long: 5.2%, SUB-long: 0%).

The boundaries which were marked only by a pause did not differ remarkably from each other across the 4 categories (CO-long: 17.2%, CO-short: 20%, SUB-long: 23.7%, SUB-short: 14.4%).

On average, the occurrence of the boundaries marked with high tone (H) without a pause was low, not higher than 15%. In the case of the short first clauses, H appeared a bit more frequent (CO-short: 15%, SUB-short: 13.1%) than in the case of long first clauses (CO-long: 8.6%, SUB-long: 3.9%).

Low tone appeared in every case with a pause, but only in CO-long sentences, the occurrence was much lower than prosodic markers’ combinations mentioned above (2.6%).

Low tone (in every case with pause) only appeared in CO-long sentences, the occurrence was much lower than prosodic markers’ combinations mentioned above (2.6%).

Smaller differences can be found between co- and subordinated sentences than between longer and shorter ones (Fig. 5). The largest difference was observed in the case of the pause and high tone, where the proportion was more than twice as high for long first clauses as for short ones. In contrast, in the case of the short sentences, the ratio of the not marked boundaries were ten times more in the short sentences, than in the long ones.
Pause with high tone appears most often in both cases (Fig. 6), but with a greater ratio in CO sentences (CO: 62.5%, SUB: 49.1%). Boundaries not marked by any cue occurred nearly three times more in the case of SUB than CO sentences (CO: 8.1%, SUB: 23.4%). Pauses with low tone appeared very rarely and only in CO sentences (2.2%). Similar occurrences can be found in the case of boundaries realized with only a pause (CO: 17.6%, SUB: 18.3%), and with only high tone (CO: 9.6%, SUB: 9.1%).

4. Conclusions

In this paper, the occurrence of prosodic boundary markers in two kinds of clauses was investigated with regard to the length of the first clause in 20 Hungarian young adult speakers’ text and sentence reading task. Silent pauses, changes in fundamental frequency and the occurrence of creaky voice were labeled and analyzed, as well as their co-occurrence with audible inhalation in possible boundary positions.

In general, the results showed that although speakers used the same boundary marking cues, the patterns of the given feature showed large individual differences.

Our first hypothesis concerned differences in the boundary marking between the two different kinds of sentence structures (SUB and CO). We hypothesized that CO clauses are stronger marked by prosodic boundary markers than the SUB clauses.

This hypothesis was corroborated by the fact that the realization of the CO and SUB sentences differed from each other to some extent. More prosodic cues were used at the same time by the speakers to denote CO clauses than SUB clauses. In SUB sentences, the clause boundaries were marked by a pause in fewer cases than in CO sentences. Additionally, CO clauses were realized with low tone and creaky voice more frequently than SUB clauses, which were realized more often without any boundary marking than the COs.

Additionally, much remarkable differences were found between long and short sentences than between CO and SUB clauses: weaker boundary marking was found after short first clauses. It can be assumed that the shorter a clause, the less necessary it is to separate it strongly from the next one, since two smaller language units can even be proceeded together, while longer units are more complex and time-consuming.

The second hypothesis concerned the occurrence of the audible inhalation. We assumed that audible inhalation appears more often in the stronger marked clauses, thus in CO clauses, and less frequently in the SUB clauses. This assumption was not corroborated, for it was not the type of sentence but its length that was decisive for the frequency of inhalation. Not just the ratio of the inhalation, but the different combinations of the prosodic features were more common after a long first clause than a short one. So in the clauses where the boundary of the clauses was more strongly marked, the proportion of inhalation was also higher: In the case of long clauses, inhalation occurred in about 80% of the boundary marked with a pause. This proportion was around 30% for short clauses.

Although the speaker has the opportunity during speech production to breathe during the pause (and vice versa, the speaker can produce a pause during the utterance due to inhalation), it can be seen that pause and audible inhalation coincided in just over half of the examined boundaries (on average 57.5%). These two events coincided most often (81.4%) in long CO clause, with a smaller proportion (63.5%) in long SUB clauses and less than half of all occurrences after short clauses (CO-short: 41.6%, SUB-short: 43.5%).

The trends reported in the research show a strong correlation with the results of previous studies. However, Peters and colleagues [8] did not find pauses being a consistent signal of the boundaries, it is well-known that pauses can have several roles in speech planning and speech production next to their boundary marker function. Pause plays an important role in the articulation, in planning the speaker’s next utterances, and also supports the listeners’ perception by giving enough time for processing.

The question is whether speech planning reflects on respiratory planning or, on the contrary, respiration reflects on speech planning, is currently being addressed in a number of studies. Due to the small number of target sentences in the present research, and the fact that the analysis was conducted based on audible inhalation, it cannot be answered in this study. Apart from this, however, the effect of the recovery hypothesis ("predicts that the probability of an intake increases with the length of the preceding utterance rather than with the length of a subsequent utterance" [19: 1413]) is visible: the ratio of inhalation was notably higher after the long clauses than after the short ones.

The possible influencing effect of the length of the following sentence is also worth examining in a subsequent study (another conception [20]).

This shows that prosodic phrasing is not entirely dependent on syntactic phrasing, but partly on the economics of production and the planning of perception. This is supported by the frequent appearance of pauses as a boundary cue in longer utterances.

Acknowledgements

Acknowledgements: the study was funded by NRDI K 135038. Additional funding was provided by the Alexander von Humboldt Foundation.
References


