

Segment prolongation in Hungarian

Mária Gósy¹ and Robert Eklund²

¹Dept. of Phonetics, Research Institute for Linguistics, Hungarian Academy of Sciences, Budapest, Hungary

²Department of Culture and Communication, Linköping University, Sweden

Abstract

Segment prolongation (PR) has been shown to be one of the most common forms of non-pathological speech disfluencies (Eklund, 2001). The distribution of PRs in the word (initial–medial–final segment) seems to vary between languages of different syllable-structure complexity, making it interesting to study segment prolongation in languages that exhibit different syllable structure characteristics. Previous studies have studied languages with complex syllable structure, such as English and Swedish (Eklund & Shriberg, 1998; Eklund, 2001, 2004) where affixation creates complex consonant clusters, and languages with very simple syllable, such as Japanese (Den, 2003) or Tok Pisin (Eklund, 2001, 2004), as well as Mandarin Chinese (Lee et al., 2004). In this paper we study PRs in Hungarian. Our results indicate that PRs in Hungarian are more similar to English and Swedish than it is to Japanese, Tok Pisin or Mandarin Chinese, which lends support to the notion that underlying morphology plays a role in how PRs is realised.

Introduction

Research on non-pathological disfluency has been carried out for very long, but although formal studies began already in the 1930s it was during the 1950s when extensive and formal studies saw the light. (For an overview of disfluency research the reader is referred to Eklund, 2004:51–171.)

From the very start of this research classification and terminology were at the core, but although 70 years have now passed there is still no general agreement on how to classify the different types of disfluencies in existence. In addition, even the term ‘disfluency’ itself is not generally agreed upon (although it is likely the most commonly used term for the phenomenon discussed here).

One type of disfluency that was recognized early on was *segment prolongation* (PR) – although the terminology varies; see Eklund (2004:163) – i.e. when a speech segment in a word is produced unusually long. Although this is similar to (what is perhaps most commonly called) filled pauses (FPs) in that both are durational and voiced, PRs have been shown to differ from FPs in some respects (e.g. Eklund, 2001).

However, one issue that has been discussed in the literature is *what segment* in the word that tends to get prolonged. A first categorization into different

classes (used by Eklund & Shriberg, 1998) was to analyse PRs in three different positions: word initial (the first segment of a word), word final (the last segment of a word) and word medial (any position that is not initial or final). Eklund and Shriberg (1998) reported almost identical distribution for American English and Swedish, with a 30–20–50% distribution, for initial–medial–final position, respectively. What made these figures interesting, however, was the appearance of studies of other languages. Eklund (2001; 2004:251) reported that the corresponding figures for distribution in Tok Pisin were 15–0–85%. Den (2003) reported 0–5–95% for Japanese and Lee et al. (2004) reported 4–1–95% for Mandarin Chinese.

Swedish is characterized by complex consonant clusters, created by additive affixation of grammatical morphemes, and the maximum allowed complexity of syllables in Swedish is C³VC⁹ (three syllable-initial consonants, and up to nine syllable-final consonants). Given that e.g. Japanese and Tok Pisin are far less permissive in this respect Eklund (2004:251) proposed that PR distribution might be the function of the morphology in the language in which they appear, something Eklund (somewhat misleadingly) called the ‘morphology matters hypothesis’.

Grammar and syntax, too, differ between those languages, so there might be other factors at play, and the ‘acid test’ would then, of course, be to study languages that expand on both the grammar/syntax and the morphology scales.

The goal of this study

In this study we have set out to investigate segment prolongation in Hungarian, a language that is different from all the languages mentioned above. Hungarian is an agglutinative language that belongs to the Finno-Ugric language family with an extremely rich morphology and an extensive system of affixation. The syntactic and semantic functions of noun phrases are primarily expressed via suffixes and postpositions. Case markings are used extensively with Hungarian nouns, but pronouns, adjectives and numerals also take case and number markings. Verbs also have a considerable number of affixes (Kenesei, Vago & Fenyvesi, 2012). Hungarian words are relatively long due to the rich morphology. The number of syllables of words is 3.7 syllables on average in spontaneous speech.

Words can easily consist of 9 or more syllables. The vowel inventory of Hungarian contains 14 vowels and 36 consonants; there are short–long phonemic pairs both in vowels and consonants. Hungarian is a ‘syllable-timed’ language where word stress invariably falls on the initial syllable although in connected speech not all words are stressed (Siptár & Törkenczy, 2000). The goal of the study was to analyse Hungarian PRs to see to what degree that morphology and syllable structure might influence the distribution of prolonged segments in spontaneous speech of the language.

Method

Thirty-six speakers (aged between 22 and 32 years, mean age: 27 years; half of the speakers were females) participated in this study who were randomly selected from the BEA Hungarian Spontaneous Speech Database (Gósy, 2012).

All subjects were native monolingual speakers of Hungarian living in Budapest, and had a similar socio-economic status. Half of both females and males had mid-level education while the other halves had university degrees. There were no indications of language or speech disorders for any of the participants.

Recordings were made in a sound-attenuated room (the same for all), under identical technical conditions using an AT4040 microphone connected directly to a computer using GoldWave to record samples at 44.1 kHz, 16 bits, monaurally. In all recordings the interviewer was the same young female phonetician.

Various types of spontaneous speech materials were used in the analysis including narratives, storytelling and a three-member conversation with each participant. One of the narratives was about the participant’s life, family, job and hobbies, while the participants talked about a topic of current interest in the other narrative and in conversations.

The duration of the analyzed spontaneous narratives was about 24 hours (ca. 40 minutes/speaker).

Target segments

All prolongations were considered occurring in the 24-hour speech material both concerning vowels and consonants. Prolongations were identified by one of the authors and was checked by another phonetician, also a native Hungarian. 0.3% of disagreement was found in the identification of prolongations between the two phoneticians; these cases were excluded from further analysis.

Prolongations were categorized according to their occurrence in the word.

Annotation was done manually using Praat software (Boersma & Weenink, 2015) according to

criteria determined in advance. Vowel boundaries were marked between the onset and offset of the second formants of the vowels. Consonants were identified depending on their acoustic structures considering their voicing part (if any), burst, release, second formant information and the neighbourhood context, as appropriate. Duration measurements were carried out automatically using a specific script. A total of 948 prolongations were found which is 0.66 PRs per minute.

Examples (prolonged segment is marked bold; the English equivalent of the target word containing the prolonged segment, is given right after the Hungarian word): *olyan szülők* ‘parents’ *ismerek meg* ‘I get acquainted with parents that’, *huszonöt nagycsoportos óvodás* ‘twenty five preschool children’, *egy tanító* ‘teacher’ *a faluban* ‘a teacher in the village’, *tudod mert* ‘because’ *nagyon elfáradtam* ‘you know because I got very tired’, *dolgoztam és* ‘and’ *jól éreztem magam* ‘I worked and felt well’, *busszal utaztam* ‘traveled’ *tegnap* ‘I traveled with bus yesterday’, *ez az elektronikus könyvtár* ‘library’ ‘this is the electronic library’, *hogyan* ‘how’ *lehet elérni* ‘how can it be reached’.

Six factors were considered for analysis: 1: Position of the target segment in the word (initial, medial, final); 2: Type of segment (vowel vs. consonant); 3: Word type (content word vs. function word); 4: Number of syllables of the word containing the prolonged segment (from 1 to 7); 5: Duration of the target segment; and 6: Gender.

For statistical analysis, a Kruskal–Wallis test was performed. The confidence level was set at the conventional 95%.

Results

Position

Beginning with distributional patterns (see above), our results are shown in Table 1.

The general distribution observed (when the one-syllable word “a” is excluded from the analysis) is approximately 18–19–63%, i.e. a distribution which is quite similar to that of American English and Swedish, especially compared to the figures reported from Tok Pisin, Japanese and Mandarin Chinese.

Table 1. PR distribution in words. The total number of PRs = 779. Note that the one-syllable word, a definite article, “a”, which arguably falls in all three categories (initial, medial, final) is reported separately.

Position	Number of occurrences	Percentage of total number
Initial	138	17.7%
Medial	148	19.0%
Final	493	63.3%
“a”	169	21.7%

Segments

What type of segments were subject to prolongation is shown in Table 2.

Table 2. Segments subject to prolongation, given in orthography and IPA and relative frequency given as percentages.

Vowels (N=628) (orthography)	IPA	Occurrence (%)
a	ɔ	37.1
e	ɛ	21.9
é	e:	13.0
i	i	10.3
á	a:	8.1
o	o	2.5
ó	o:	2.3
ö	ø:	2.3
í	i:	0.9
õ	ø	0.3
ü	y	0.1
ú	u:	0.1
Consonants (N=320) (orthography)	IPA	Occurrence (%)
s	ʃ	42.8
m	m	19.1
n	n	18.1
z	z	8.1
sz	s	3.7
h	h	1.8
gy	ʒ	1.2
k	k	1.2
f	f	0.9
l	l	0.9
ty	c	0.3
v	v	0.3
tt	t:	0.3
ny	ɲ	0.3
p	p	0.3
cs	tʃ	0.3

As is seen, prolongation affects all possible kinds of segments, similar to what has been reported for English and Swedish.

Word type

In Figure 1 we report how prolongation occurred as a function of whether the words affected occurred on content words or function words.

As is seen in Figure 1, prolongation on content words is, on average, shorter than it is on function words. This sits well with proposed theories that hesitation occurs whenever important choices are made in speech production, sometimes referred to as the “many-options hypothesis” (see e.g. Eklund & Wirén, 2010:24).

Number of syllables in words

We also set out to find out whether the number of syllables in the affected words played a role in segment prolongation. Our results are shown in Table 3. As can be seen there is a strong linear fall-off as a function of number of syllables in the affected words: the fewer the number of syllables, the more likely the word is to exhibit prolongation.

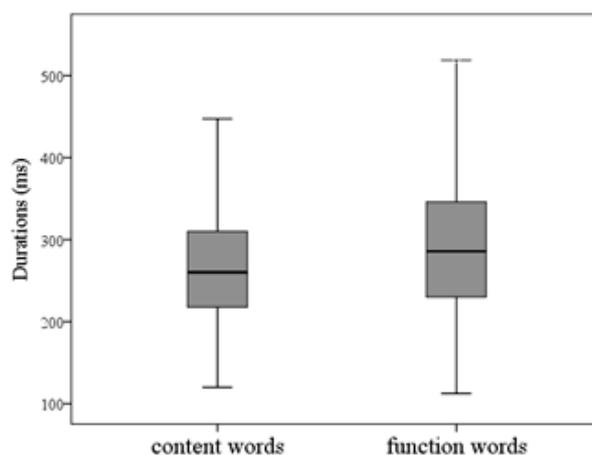


Figure 1. Prolongation as a function word type. Total number of content words = 371. Total number of function words = 577. The difference is significant, chi-square (two-tailed) at $p < 0.001$.

Table 3. Prolongation as a function of number of syllables in the affected word, given both as actual number of occurrences and as relative frequency, as well as the relative frequency of words in spontaneous speech. The total number of words analysed = 948.

Number of syllables of words	Occurrences of the words	Relative frequency (%)	Relative frequency of words in spontaneous speech
1	589	62.1	44.7
2	181	19.1	28.8
3	101	10.6	15.2
4	56	6.0	7.6
5	16	1.7	2.7
6	2	0.2	0.7
7	3	0.3	0.2

Duration of the prolonged segments

In Figure 2 below we show the results of our durational analysis, broken down for vowels and consonants. As is shown in Figure 2, prolongation is generally longer on vowels than on consonants.

Gender

Finally, we observed that there is a small, but significant, tendency for men to produce longer prolongations than females, chi-square (two-tailed); $p = 0.012$.

Discussion and conclusions

Starting with the *Distribution*, there is a remarkable similarity between our results from Hungarian and previous reports on American English and Swedish, especially when compared with the reported figures from Tok Pisin, Japanese and Mandarin Chinese. So, at a first glance it would seem as the proposed ‘morphology matters hypothesis’ is given some support in the present study.

However, recent results from German seem to point in another direction, and suggest that at least a strong version of the morphology matters hypothesis is not supported.

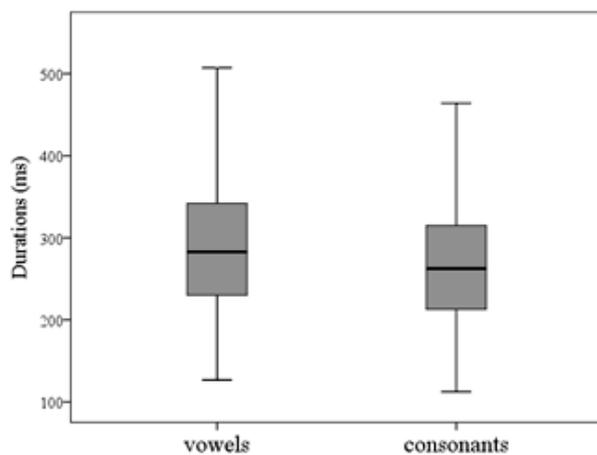


Figure 2. Durations of PRs broken down for vowels and consonants. The difference is significant, chi-square (two-tailed) at $p < 0.001$.

Evidence against such a strong interpretation comes from German, where the distribution 7–15–78% was found (Betz, Eklund & Wagner, 2017). Since German and Swedish have very similar morphology – more similar than that of Hungarian and Swedish – and both exhibit phenomena like frequent and creative compounding it would seem that morphology alone cannot explain the observed differences in distribution.

As for *Segments*, the most striking observation is that tantamount to American English and Swedish, all types of segments are subject to prolongation.

As for *Word Type*, the tendency is to prolong function words more than content words, something that sits well with the “many-options hypothesis” of the roll hesitation plays in speech production.

As for *Duration*, vowels are, on the whole and perhaps not surprisingly, more prolonged than consonants in our data.

As for *Sex*, there is a small tendency for male speakers to produce longer PRs than female speakers, supporting the proposed hypothesis that men are less prone to yielding the floor in dialog (see Eklund & Wirén, 2010:23). Females produced more PRs (500 items) than males (448 items) which can also explain their shorter lengthened segments.

We think that our paper not only sheds light on previous research on speech prolongation but also reveals many new details about this sometimes neglected disfluency. It is our hope that future studies will provide even more insights into segment prolongation in non-pathological speech. Finally, it must be pointed out that the reported figures might be indicative of what *kind* of data was used. For example, the American English and Swedish data used by Eklund and Shriberg (1998) were all telephone data, and disfluency in dialog over a telephone line, where interlocutors cannot make use of visual cues, might be different from disfluency in face-to-face dialog.

Acknowledgements

The research was supported by OTKA Project, #108762. Thanks to Beáta Megyesi for comments on Hungarian morphology.

References

- Betz, S., R. Eklund & P. Wagner. 2017. Prolongation in German. In R. Eklund (ed.): *Proceedings of DiSS 2017*, 18–19 August, Royal Institute of Technology, Stockholm, Sweden [this volume], 5–8.
- Boersma, P. & D. Weenink. 2015. *Praat: doing phonetics by computer*. <http://www.praat.org> (Accessed 2014).
- Den, Y. 2003. Some strategies in prolonging speech segments in spontaneous Japanese. In R. Eklund (ed.), *Proceedings of DiSS'03, Disfluency in Spontaneous Speech*, 5–8 September 2003, Göteborg, Sweden. *Gothenburg Papers in Theoretical Linguistics 90*, ISSN 0349–1021, 87–90.
- Eklund, R. 2001. Prolongations: A dark horse in the disfluency stable. In *Proceedings of DISS 2001, Disfluency in Spontaneous Speech*. 29–30 August 2001, Edinburgh, Scotland, 5–8.
- Eklund, R. 2004. *Disfluency in Swedish human–human and human–machine travel booking dialogues*. PhD thesis, Linköping University, Sweden. ISBN 91-7373-966-9, ISSN 0345-7524
- Eklund, R. & E. Shriberg. 1998. Crosslinguistic Disfluency Modelling: A Comparative Analysis of Swedish and American English Human–Human and Human–Machine Dialogues. *Proceedings of ICSLP 98*, 30 November – 5 December 1998, Sydney, Australia, 6:2631–2634.
- Eklund, R. & M. Wirén. 2010. Effects of open and directed prompts on filled pauses and utterance production. In: *Proceedings of Fonetik 2010*, 2–4 June 2010, Lund, Sweden, 23–28.
- Gósy, M. 2012. BEA – A multifunctional Hungarian spoken language database. *The Phonetician* 105/106, 50–61.
- Kenesei, I., R. Vago. & A. Fenyvesi. 2012. *Hungarian*. New York: Routledge.
- Lee, T.-L., Y.-F. He, Y.-J. Huang, S.-C. Tseng & R. Eklund. 2004. Prolongation in spontaneous Mandarin. In *Proceedings of Interspeech 2004*, 4–8 October 2004, Jeju Island, Korea, vol. III, 2181–2184.
- Siptár, P. & M. Törkenczy. 2000. *The phonology of Hungarian*. Oxford: Oxford University Press.



Proceedings of



DiSS 2017

The 8th Workshop on Disfluency in Spontaneous Speech

**KTH Royal Institute of Technology
Stockholm, Sweden
18–19 August 2017**

**TMH-QPSR
Volume 58(1)**



**Edited by
Robert Eklund & Ralph Rose**

Conference website: <http://www.diss2017.org>

Proceedings also available at: <http://roberteklund.info/conferences/diss2017>

Cover design by Robert Eklund

Graphics and photographs by Robert Eklund (except ISCA and KTH logotypes)

Proceedings of DiSS 2017, Disfluency in Spontaneous Speech

Workshop held at the Royal Institute of Technology (KTH), Stockholm, Sweden, 18–19 August 2017

TMH-QPSR volume 58(1)

Editors: Robert Eklund & Ralph Rose

Department of Speech, Music and Hearing

Royal Institute of Technology (KTH)

Lindstedtsvägen 24

SE-100 44 Stockholm, Sweden

ISSN 1104-5787

ISRN KTH/CSC/TMH-17/01-SE

© The Authors and the Department of Speech, Music and Hearing, KTH, Sweden