View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Repository of the Academy's Library

Lengthenings and filled pauses in Hungarian adults' and children's speech

Andrea Deme[1,2] *& Alexandra Markó*[1]

[1] Department of Phonetics, Eötvös Loránd University, Hungary
[2] Research Institute for Linguistics, Hungarian Academy of Sciences, Hungary

Abstract

In the present paper vowel lengthenings and non-lexicalized filled pauses were studied in the spontaneous speech of children and adults (focusing more on the much less studied phenomenon: vowel lengthening). The results revealed different usage and appearance of lengthenings in the two age groups, therefore, differences in speech skills and strategies can be concluded. LEs and FPs differ mostly in their position in the speech session between the age groups, which has implications regarding different planning strategies of adults and children. We also draw conclusions regarding the methodological considerations in the issue of identifying vowel lengthening supporting a previously formulated conception.

Index Terms: lengthening, (non-lexicalized) filled pause, spontaneous speech, speech planning, discourse management.

1. Introduction

Although lengthenings and pauses (silent and filled) are the two most common disfluencies of spontaneous speech, lengthenings have received little attention so far in the literature, while silent and filled pauses are more often placed in the focus of research. Lengthenings (hereafter LE) are generally considered to be similar to non-lexicalized filled pauses (hereafter FP) as both use the least amount of energy and similar duration to afford the speaker time for thought and speech planning [1]. However, there is also a growing body of evidence suggesting that despite their obvious similarities, they also differ significantly in their acoustic properties as well as their positions and functions in discourse.

Previous findings have already implied that children may differ from adults in terms of the acoustic features, and the place of occurrence of LE [2] and the frequency of occurrence of FP [3], which might be the result of differences in speech planning skills and strategies. Based on these observations, the aim of the present study was to compare the patterns of LE and FP within and between two different age groups.

In our investigation different frequency of LEs and FPs between age groups was hypothesized, with children leaning toward lengthening, and adults being more prone to hesitations. Differences in phonetic and positional characteristics of both phenomena between adults and children also were assumed. The differences were thought to originate from the presumed differences of speech production strategies of the speakers of the two age groups.

2. Subjects, material and method

Spontaneous speech samples of 8 children (4 boys, 4 girls, aged 7 to 8 years) and 8 women (aged between 20 and 32 years) were used for analysis. The total duration of the speech material was approximately 1 hour.

For identifying prolonged vowels a perception test was used, introduced in [2]. As [2] points out, two approaches are often blended in the literature: to regard LE as a perceptual phenomenon, or separate lengthened vowels on the grounds of physical duration alone. Therefore, in our present study we adhere to the notion, which regards LE as perceptual phenomena; hence the designation of LEs was carried out with a listening test including 10 linguists. They were asked to listen to the speech samples through headphones, played by a computer, in a quiet room, and meanwhile to follow the transcription of the played texts presented on the screen (the transcription was orthographic, but without punctuation). The subjects' task was to mark those vowels in the written texts, which were considered lengthened based on their subjective auditory judgment. The vowels marked by at least 6 linguists were counted as lengthened. FPs in the texts were identified by means of auditory perception of the authors, together with the visual confirmation of a spectrogram.

The two phenomena (LEs and FPs) were analysed both in children's and adults' speech, and compared according to the following parameters:

- frequency of occurrence,
- duration,
- fundamental frequency (f_0 in semitones),
- position in the clause,
- the position in the "speech session"(separate part of the utterance bordered by silent pauses), and
- the presence of adjacent disfluencies.

Though the following parameters were only applicable to the inspection of LEs, they were included in the research, since they seemed to reveal important aspects of this (less studied) phenomenon. These additional features were:

- the duration of lengthened and non-lengthened vowels,
- the extent of lengthening,
- the rate of lengthened vowels,
- the LE's position within the word, and
- the ratio of word types (content vs. function) involved.

Praat 5.3 was used for acoustic analysis and SPSS 15.0 for statistical analysis (*t*-test, Pearson-correlation).

3. Results

3.1. Comparison of LE and FP between the two age groups

In the adults' corpus 70 LEs and 127 FPs were annotated. In terms of **frequency** of occurrences this means that a FP was produced on average every 17 seconds, and roughly one in every 115 vowels was marked as lengthened. In the children's material we found 67 LEs and 56 FPs, which means an FP every 32 seconds, and a LE occurred once in every 56 vowels. The rate of FPs seems to be similar in the two groups, but the rate of LEs in child speech is almost the double of the count calculated for adults.

When compared to the count of all vowels uttered, adults seem to prefer FPs to LEs, which in adults' speech seem to be relatively rare (0.44%) with a moderate standard deviation (0.33%). In child speech, however, LE seemed to appear more often and diverged (0.78±0.65%). The frequency as well as the deviation of FP was similar in the two groups (adults: 0.78±0.55%; children: 0.66±0.55%) (Figure 1).

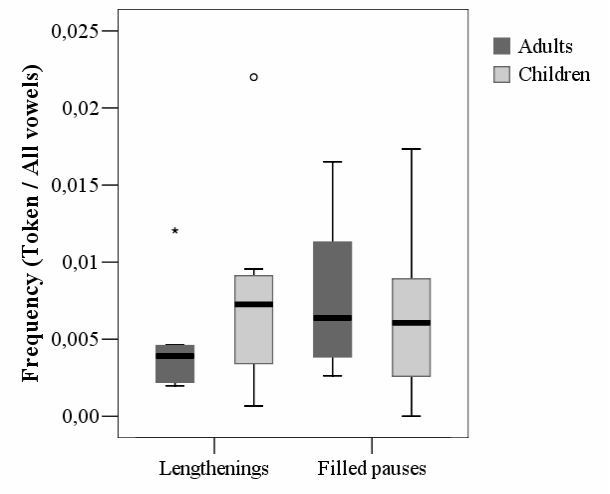

Figure 1: *Frequency of LE and FP in both age groups.*

There was no correlation observed between the ratio of FPs and LEs in the speakers' material. As we have seen, the inter-speaker variability is rather high, however the age groups do not differ in this point of view. Figure 2 shows that there are speakers among both children and adults who preferred LEs (e.g., A3, Ch1, Ch4), but the opposite preference can also be noticed (e.g., A2, A5, Ch8), while balanced ratios are shown as well (e.g., A1, A8, Ch6).

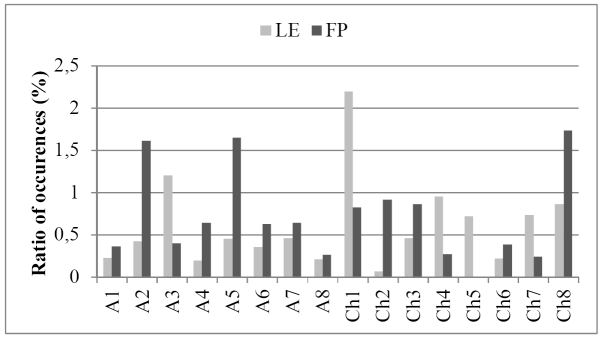

Figure 2: *Ratio of occurrences of LEs and FPs in the speakers' material in % of the total amount of speech sounds (A= adult, Ch= child).*

The pattern of the **duration**al differences between the two phenomena differed according to age groups (with greater variability for FPs in both): while in child speech LE might be slightly longer than FP on average (LE: 344±87 ms; FP: 319±189 ms), in adults it is the opposite (LE 286±105 ms; FP: 374±143 ms) (Figure 3).

The **f_0 values** of FP show a very systematic correspondence: they appear to be very similar within the two groups, but, as it could be presumed, children realized both of the phenomena on a higher fundamental frequency (LE for adults: 10±3 st, for children: 18±2 st; FP for adults: 9±3 st, for children: 19±4 st).

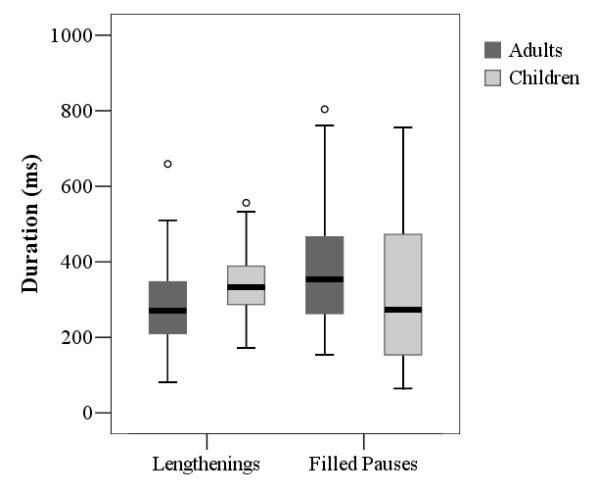

Figure 3: *Duration of LE and FP.*

With respect to the **position in the clause** LE and FP vary marginally in adults, and both phenomena seem to occur mainly on the boundaries (at the conjunction or at the beginning or ending) of the clauses (Figure 4). Contrarily, children tended to use LE and FP more separately: as the distribution of LE is leaning towards the boundary position, FPs seem to be equally distributed.

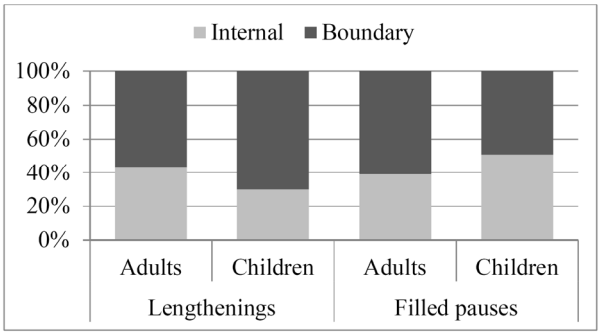

Figure 4: *Position of LE and FP in the clause.*

Position inside the speech sessions can be interpreted as follows. For an FP the position is determined with the help of the adjacent silent pauses (*initial* position is preceded by a pause, *final* position is followed, *isolated* is preceded and followed by a pause, while *medial* means the lack of adjacent pauses). Therefore the results suggest that in child speech half of the occurrences stand between pauses, while adults preferred the initial or final positions in a same ratio (Figure 5). For LE the position refers to the word's place in which the LE occurred (e.g. LE in the initial position is in the first word of the session, and isolated position means a one-word session). For LEs we could not find remarkable differences in the position patterns of the two age groups, both used LE mainly in the final position. This can be a marker of discourse management.

The context of both phenomena was analyzed to see whether any (other) **disfluency phenomena** (like false start, repetition, etc.) occur next to the given LE or FP. According to the results, LEs in adult speech are more likely to occur in the company of other disfluency phenomena than in child speech (yet almost 80% appear without adjacent disfluencies), whereas in child speech they seem to be a more independent means of discourse. In contrast, the pattern is the opposite for FPs, as FPs in children have the highest ratio of neighbouring disfluencies out of all cases (Figure 6).

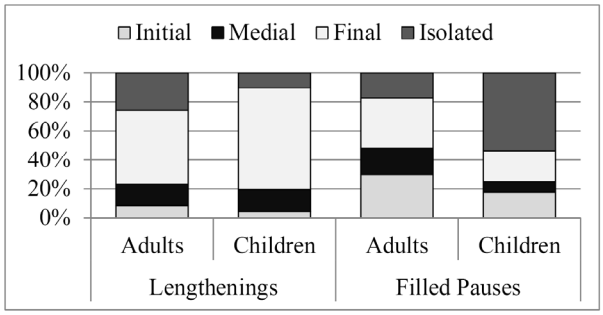

Figure 5: *Position of LE and FP in the speech sessions.*

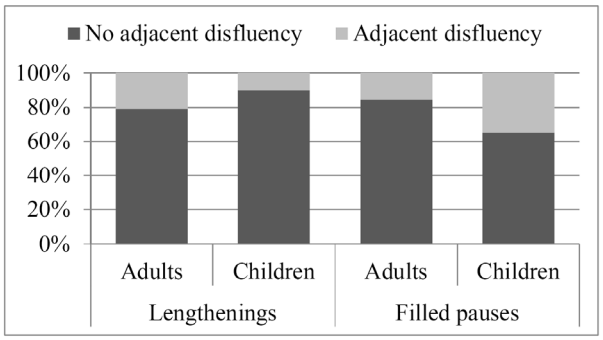

Figure 6. *Distribution of LEs and FPs in terms of adjacent disfluencies.*

The figure does not show those cases when a FP and a LE occurred together. In the children's material 10.4% of LEs were accompanied by a FP, this ratio is 24.3% in adults' speech. At the same time 12.5% of FPs were preceded or followed by a LE in children's speech, and similarly 12.2% in the adults' material.

3.2. Additional differences between adults and children in terms of the characteristics of LE

The **duration** of all vowels was measured in the corpus. The values of duration (as one might expect on the basis of the data, see Figure 7) showed positive correlation with the number of markings in both age groups (Pearson's $r = 0.514$, $p < 0.001$ for children; Pearson's $r = 0.645$, $p < 0.001$ for adults). In most cases, the physical length of the prolonged vowels was significantly longer than the values of the corresponding non-lengthened realizations in both age groups (*t*-test, $p < 0.05$), but at the same time, a relevant degree of overlap can be observed between the lengthened and non-lengthened clusters (Figure 7). This means that some of the non-marked vowels' duration exceeded some LEs' duration. (It should be noted that Hungarian sound system distinguishes phonologically short and long vowels [4], although the physical duration of short and long vowels naturally overlap [5].)

The **extent of lengthening** was determined as follows: the average duration of vowels, which in the perception test were marked as lengthened were compared to the average of those, which were not marked by at least 6 informants (Figure 8). The results showed differences in terms of vowel quantity, but the age groups differed only with respect to the long vowels. In case of (phonologically) short vowels a greater extent of lengthening was observed than for long ones; and the degree of prolongation was very similar in the two age groups (children: 374%, adults: 365%). For long vowels an average prolongation of 262% was found in the children's material, while 307% in adults' speech.

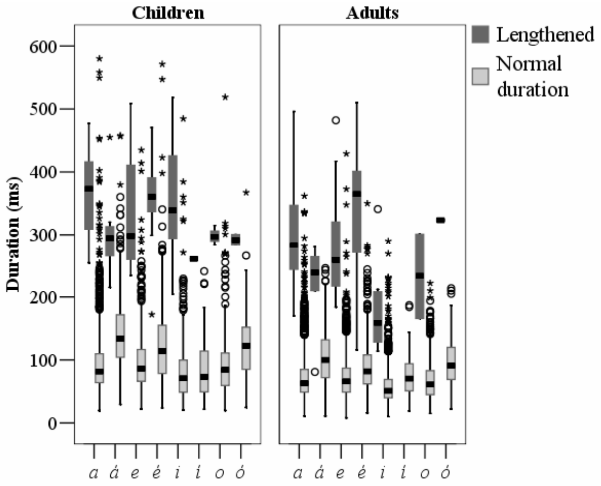

Figure 7: *Average duration of vowels (a* = [ɑ], *á* = [aː], *e* = [ɛ], *é* = [eː], *i* = [i], *í* = [iː], *o* = [o], *ó* = [oː]*).*

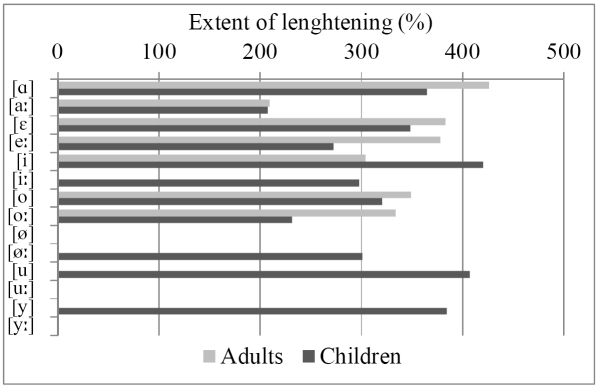

Figure 8: *The extent of lengthening.*

Ratio of lengthened vowels was defined as well. With respect to the vowel quality the differences between the two age groups were spectacular: while in children's speech all vowels of the Hungarian sound system were lengthened in a certain amount, in adults' speech this phenomenon was not documented in case of [ø øː u uː y yː] (Figure 9). It should be noted, that the ratio of the various vowels was similar in the speech samples (and in accordance with the frequency of speech sounds in Hungarian spontaneous speech reported by the literature [5]). Although [ø øː u uː y yː] are relatively rare in Hungarian speech, children lengthened a remarkable proportion of their occurrences (while adults did not at all).

In analysis of the LE's **position within the word** four categories were determined: *initial* means LE in the first syllable, *final* means LE in the last syllable, *medial* means a LE word internally, and *isolated* stands for one-syllable words, in which the vowel was lengthened (Figure 10). In adult speech final and isolated tokens occurred in the same ratio, while in child speech the frequency of isolated LEs exceeded 70%.

Finally the **type of the word** in which the LE occurred (content or function word) was compared between the age groups. The similarity of the proportion of isolated LEs (in Figure 10) and function words (in which LEs appear) (in Figure 11) is not surprising, as most of both of them can be traced back to the lengthened (one-syllable) articles and connectives, while in content words LEs tend to be positioned in the last syllable.

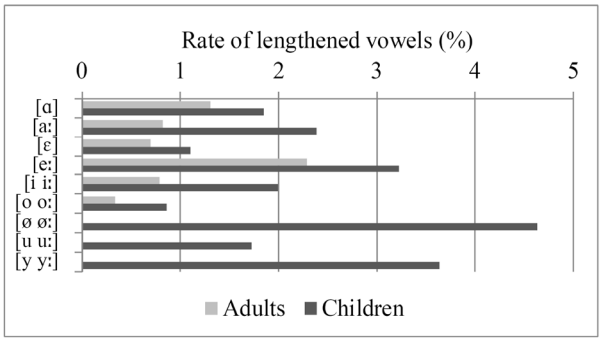

Figure 9: *Rate of lengthened vowels in terms of vowel quality (lengthened/all).*

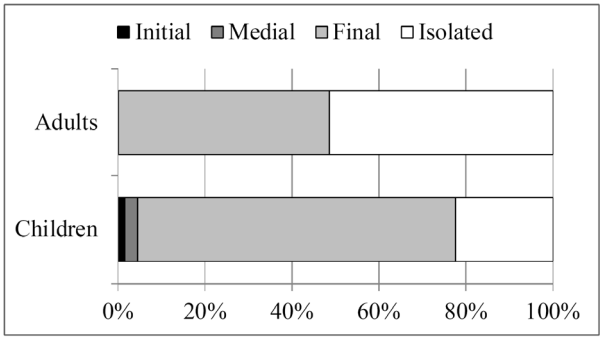

Figure 10: *Position of LEs within the word.*

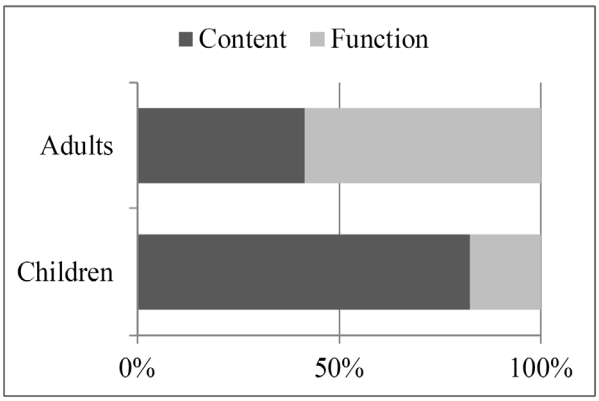

Figure 11: *Distribution of LEs in terms of word type.*

4. Discussion and conclusions

The above analysis revealed that occurrences of LE show different patterns in child and adult speech: despite the great inter-speaker variability children on average used LEs twice as often as adults do and the phenomenon in child speech affected all of the vowel qualities, but mostly the content words (while in adult speech the rare vowels were excluded and mostly the function words were involved, as described for Swedish adult speech as well [1]).

The greater variability of the duration of FPs in child speech might have some implications regarding latter acquisition (thus not yet stabilized strategy) using FP in the discourse.

Both phenomena occurred mainly on the boundaries of the clauses in adult speech, while in the children's material LE was more often in the boundary position, but FPs were equally distributed. On one hand, balance in distribution might denote a more incidental use of a speech unit. On the other hand, as the clause-internal position means a point of grammatical incompletion, pattern differences found in "filling up" this position with a non-lexical item, again, imply differences between the speech planning processes or discourse managing strategies of the speakers. This is especially true, if we accept the generally assumed idea about the function of FP: pausing, or time-gaining (for a basic summary of typical interpretation found in the literature see [6]).

According to the argument of [6], a clitic attached onto the previous element (and any additional lengthening of it) signals that the clitic was planned simultaneously with the element it is stuck to. Adapting their reasonable argument, we suggest that great differences found in position of LE and FP in the speech session indicate planning differences.

LE and FP diverge between the two age groups based on their position in speech sessions: session ending LEs found in both groups, and ending FPs found in adults imply a similar planning strategy while FPs in child speech, which form isolated utterances imply another. Furthermore, it can also be assumed that session final LE and FP can be a marker of discourse management.

Conclusions can also be drawn concerning the methodology of designating LEs applied in the paper. The duration of vowels showed positive correlation with the number of marks collected in the listening test and the lengthened vowels showed statistically longer durations from the normal realization. However, it was also revealed that in some cases lengthened vowels were shorter in duration. These findings support the suggestion introduced in [2], which states the necessity of differentiating the two approaches often blended in the literature: to regard LE as a perceptual phenomenon, or separate lengthened vowels on the grounds of physical duration alone.

Our analysis sheds light on the different usage of LE and FP within and between the two age groups, suggesting that LE and FP maybe a part of a globally different speech production strategy changing with age and can even have functions over and beyond providing time for speech planning (e.g. in discourse management).

5. References

[1] R. Eklund, "Prolongations: A dark horse in the disfluency stable", *Proc. on Disfluency in Spontaneous Speech (DiSS'01)*, Edinburgh, Scotland, UK, August 29–31, pp. 5–8, 2001.

[2] A. Deme, "Magánhangzónyújtások gyermekek spontán beszédében" [Lengthenings in the spontaneous speech of children], in Váradi, T. [ed], *VI. Alkalmazott Nyelvészeti Doktoranduszkonferencia*, Budapest, Hungary, pp. 24–39. Online: http://www.nytud.hu/alknyelvdok12/proceedings12/deme2012.pdf, 2013.

[3] K. Menyhárt, "A spontán beszéd megakadásai az életkor függvényében", in Hunyadi, L. (ed.) *Kísérleti fonetika – laboratóriumi fonológia a gyakorlatban* [Experimental phonetics – laboratory phonology in practice], Debrecen: Debreceni Egyetem Kossuth Egyetemi Kiadója, pp. 125–138, 2003.

[4] P. Siptár, Törkenczy, M., *The Phonology of Hungarian*. Oxford: Oxford University Press, 2000.

[5] M. Gósy, *Fonetika, a beszéd tudománya* [Phonetics, the science of speech]. Budapest: Osiris. 2004.

[6] H.H. Clark and J.E. Fox Tree, "Using uh and um in spontaneous speaking", *Cognition* 84(1), pp. 73–111, 2002.

Proceedings of

DiSS 2013

The 6th Workshop on Disfluency in Spontaneous Speech

KTH Royal Institute of Technology
Stockholm, Sweden
21–23 August 2013

TMH-QPSR
Volume 54(1)

Edited by
Robert Eklund

Conference website: http://www.diss2013.org
Proceedings also available at: http://roberteklund.info/conferences/diss2013

Cover design by Robert Eklund
Front cover photo by Jens Edlund and Joakim Gustafson
Back cover photos by Robert Eklund

Proceedings of DiSS 2013, The 6th Workshop of Disfluency in Spontaneous Speech
held at the Royal Institute of Technology (KTH), Stockholm, Sweden, 21–23 August 2013
TMH-QPSR volume 54(1)
Editor: Robert Eklund
Department of Speech, Music and Hearing
Royal Institute of Technology (KTH)
Lindstedtsvägen 24
SE-100 44 Stockholm, Sweden

ISBN 978-91-981276-0-7
eISBN 978-91-981276-1-4
ISSN 1104-5787
ISRN KTH/CSC/TMH--13/01-SE
TRITA TMH 2013:1

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

Printed by Universitetsservice US-AB, Stockholm, Sweden, 2013