Izvorni znanstveni rad
Rukopis primljen 11. 10. 2014.
Prihvaćen za tisak 12. 12. 2014.

Tilda Neuberger, Mária Gósy
neuberger.tilda@nytud.mta.hu, gosy.maria@nytud.mta.hu
Research Institute for Linguistics, Hungarian Academy of Sciences, Hungary

A cross-sectional study of disfluency characteristics in children's spontaneous speech

Summary

Since both the erroneous and the well-formed, norm-following utterances are produced by the same rules of production, the analysis of disharmonic phenomena and self-repairs can help the researcher to look into the hidden operation of speech planning processes, also in the stages of language acquisition. The aim of the present study is to analyze age-specific patterns of children's disfluencies and repair strategies. Occurrences of various types of disfluencies, the repair ratio and the duration of editing phases were examined in spontaneous speech samples of seventy monolingual Hungarian-speaking children aged between 6 and 13. Results seem to confirm that the ratio of self-repairs depends on the children's age and the type of disfluency. Findings of our cross-sectional analysis shed light on the changes of speech fluency which may be associated with the developing language usage.

Key words: disfluency, self-repair, spontaneous speech, monolingual Hungarian children

1. INTRODUCTION

Speech development is relatively rapid until the age of six. The fluency of children's utterances is affected by various factors, including syntactic complexity and discourse complexity. Children are assumed to acquire the forms and functions of disfluencies from adults' spontaneous utterances together with the acquisition of the semantic, syntactic, phonological, etc. rules of their first language (MacLachlan & Chapman, 1988). Disfluent utterances tend to be those that are relatively complex within the child's own system (Gaines et al., 1991; Yaruss et al., 1999). Long silent pauses that are characteristic of children's speech at the very beginning (Ambrose & Yairi, 1999) will later on serve processes like selection, retrieval or syntactic and other planning before an utterance is articulated. Children encounter disfluencies more frequently as they get older because adults use more complicated utterances with older children, and adults' speech contains various instances of disfluency. The rate of disfluency varies as a function of several factors like speaker familiarity, utterance length, and speech rate, both in adults (Shriberg, 1994) and in children (Kidd et al., 2011).

The present study addresses age-specific patterns of Hungarian-speaking children's disfluencies and repair strategies from the age of 6 until the age of 13. The aim of the study is to detect the occurrences and types of disfluent episodes as well as their repairing characteristics produced by children with typical language acquisition. Although children's spontaneous utterances differ from those of adults in many respects, particularly in complexity and fluency, their disruptions and disfluent episodes show quite a few similarities to corresponding phenomena in adult speech (e.g., Hudson Kam & Edwards, 2008). Several studies in the area of children's speech fluency have been primarily motivated by the necessity of interventions for children with language disorders (e.g., Yaruss et al., 1999; Guo et al., 2008). Therefore, these studies mainly focus on preschool ages (Tumanova et al., 2014). The goal of a large number of studies that analyze children's speech disfluencies is to better understand stuttering (Yairi & Ambrose, 2005; Boey et al., 2007; Howell & van Borsel, 2011). There are strong interactions between fluency and grammaticality in stuttering children's spontaneous speech (Hall et al., 2007) as well as between lexical factors and stuttering (Newman & Bernstein Ratner, 2007). During the early stages of language development (generally between the ages of 2

and 3), a lot of children undergo a period of typical disfluency (e.g., Ambrose & Yairi, 1999). Previous studies seem to be controversial concerning the effect of age on the occurrence and types of disfluency. There are studies demonstrating that complexity of utterances would increase with age and would appear to be correlated with occurrences of disfluency (e.g., Evans, 1985) while others show no age-dependent differences (e.g., DeJoy & Gregory, 1985; Carlo & Watson, 2003; Watson et al., 2011).

Adults' speech disfluencies tend to occur when speakers encounter difficulties in forming concepts, activating syntactic frames, or retrieving the syntactic and semantic information of lexical items (Levelt, 1989). The question arises whether corresponding phenomena in speech samples of children between the ages of 6 and 13 show the same patterns or, on the contrary, children's disfluencies have their own specific characteristics related to language acquisition. In addition, age could be a relevant factor in children older than 6 years, as opposed to the data reported about younger children in the literature (e.g., Watson et al., 2011). Although our study is based on a cross-sectional analysis, we think that the data obtained add relevant information about the nature of disfluencies in children's speech in a period where speech production is different from both young children's (younger than 6) and young adults' speech.

The processes of monitoring and repair must be highly relevant during the early stages of speech development, already at about 2 years of age (Forrester, 2008). Children must have an ability to process socially and linguistically motivated self-repairs following the monitoring of their own speech. Although children's self-repairs are by and large universal, differences are reported in the literature depending on language, context and speech style (Salonen & Laakso, 2009). Beside disfluencies, this paper reports data on self-repairs found in children's spontaneous speech samples between the ages of 6 and 13.

Since both the erroneous and the well-formed, norm-following utterances are produced by the same rules of production, the analysis of disharmonic phenomena and self-repairs can help the researcher to look into the hidden operation of speech planning processes in the language acquisition period analyzed. Our hypotheses are that (i) the occurrence of disfluencies would increase with age between 6 and 13 years, (ii) the ratio of self-repair would depend on the children's age and the type of disfluency, and (iii) the durations of editing phases would decrease with age.

2. METHODOLOGY

2.1. Participants

Seventy typically developing Hungarian-speaking monolingual children participated in this study. Thirty-three of them were boys and thirty-seven were girls. The analysis was cross-sectional, including five age groups: (i) 6-year-old preschool children, (ii) 7-year-olds (first-graders), (iii) 9-year-olds (third-graders), (iv) 11-year-olds (fifth-graders), and (v) 13-year-olds (seventh-graders), see Table 1. There were 14 children in each age group. None of them had any hearing or speech disorders (based on the tests' data that had been administered to each child before their speech was recorded, cf. Juhász, 1999). The IQ and cognitive functions of all participants were within normal limits. All school-age children came from the same elementary school, while the preschool children came from a kindergarten close to the school in Budapest, the capital of Hungary. They had similar social and cultural background.

Table 1. Age and gender distribution of the participants
Tablica 1. Distribucija ispitanika prema dobi i spolu

Experimental groups	Age (year; month)	Number of children	Number of boys	Number of girls
6-year-olds	6;1–6;11	14	8	6
7-year-olds	7;2–7;7	14	7	7
9-year-olds	9;4–9;10	14	6	8
11-year-olds	11;4–11;10	14	5	9
13-year-olds	13;1–13;9	14	7	7

2.2. Method and material

Spontaneous speech samples were recorded directly to computer at the children's school (in a quiet room) using a Sennheiser ME2 microphone, and were digitized at a 44.1 kHz sampling rate and 16-bit resolution. Participants were recorded individually in the mornings. The interviewer was the same female researcher in each case. The task of the children was to talk about their free-time activities, hobbies and everyday life. Speaking time was not limited. The total duration of the

corpus was 371.2 minutes (more than 6 hours). The shortest recording was 1.9 minutes long, while the longest one was 8.7 minutes long, the mean duration of the recorded speech samples was 4.4 minutes. Table 2 summarizes the sample durations, speech tempi and standard deviations across ages.

Table 2. Temporal data of the recorded material (SD = standard deviations)
Tablica 2. Trajanja snimljenoga materijala (SD = standardne devijacije)

Experimental groups	Data of speech materials				
	Mean duration (minute)	SD (minute)	Mean speech tempo (words/min)	SD (words/min)	Number of words produced (mean/SD)
6-year-olds	4.2	1.6	82.1	16.0	345/139
7-year-olds	3.8	1.0	81.3	13.7	312/107
9-year-olds	5.7	1.6	88.9	14.1	506/178
11-year-olds	4.5	1.5	90.1	16.2	422/220
13-year-olds	4.1	1.1	99.6	24.4	411/163

The recordings were annotated using Praat 5.3 software (Boersma & Weenink, 2011). The frequency of disfluency types, the ratio of repaired and unrepaired disfluencies, and the duration of editing phases were coded. The term "editing phase" refers to the interval between the interruption point and the onset of the repair (Figure 1). This phase is characterized by pausing, hesitation and/or editing terms (Levelt, 1983). All incidents of disfluency were categorized by the two authors separately (Table 3). Given that a variety of terms is used for them in the literature, Table 3 contains the definition of each disfluency type followed by an example. In cases of very rare disagreement (less than 3% of the cases), the phenomenon at hand was excluded. Disfluency types were not considered and were excluded from further analysis if their occurrence did not reach 40 instances (for example, incidents of restarts and "tip of the tongue" phenomena). Each repeated stretch of speech was coded as one instance of disfluency, even if the repeated phrase consisted of more than one word.

Table 3. Types and definitions of the analyzed Hungarian disfluencies with English examples

Tablica 3. Tipovi i definicije analiziranih mađarskih disfluentnosti s primjerima na engleskom

Type of disfluency	Definition	English examples
Filled pauses	non-lexical one-syllable vocalizations reflecting various functions including signals of some trouble in the speech planning processes	*um*, *uh*, *er*
Filler words	conventional words or phrases which do not contribute to the semantic content of the discourse, and operate as discourse markers with various functions	*I mean*, *well*, *you know*
Repetitions	repeated portions of speech (e.g., words or phrases) that do not signal emphatic meaning	*I've been to to America*
Grammatical errors	linguistic forms that do not follow the morphological or syntactic rules and norms of the dialect that children of this study use	*I were was ill*
False starts	word fragments, where the speaker realizes, before articulating a complete word, that it does not correspond to the intended (target) word	*I have a do cat*
False words	completely articulated words that were not intended to be produced	*I have a dog a cat*
Phonological errors	sublexical errors (anticipation, perseveration, substitution, deletion)	*blocks of flowers* (box of flowers); *great rist* (great risk); *tennis bacquet* (tennis racquet); *unamity* (unanimity)

The corpus contained 3,870 disfluent incidents in total; 707 of them uttered by 6-year-olds, 471 uttered by 7-year-olds, 1,050 uttered by 9-year-olds, 786 uttered by 11-year-olds, and 856 uttered by 13-year-olds. Examples from the corpus and their English equivalents (disfluencies are written in bold):

Filled pauses: ***mmm*** *a kedvencem talán amit eddig olvastam A gyűrűk ura* "**mmm** my favorite perhaps of what I've read so far is The Lord of the Rings" (13-year-old boy); ***öööö*** *barátokkal szoktunk moziba menni* "**erm** we usually go to the cinema with friends" (13-year-old boy).

Filler words: ***ugye*** *nekik* ***ilyen*** *nagy tévéjük van* "**now** they **kind of** have a large TV set" (6-year-old girl); ***hát*** *egy kicsit tanultam úszni de nem tudok* "**well** I learned to swim a little but I can't really" (6-year-old boy).

Repetitions: *lemegyek a játszótérre vagy* ***vagy*** *macskázom* "I go down to the playground or **or** play with my cat" (11-year-old girl); *van benne kihívás is hogy* ***hogy*** *ne essél el* "there's also the challenge that **that** you don't trip over" (11-year-old girl).

Grammatical errors: *van egy kutyájuk* ***akik*** *öö aki mindig ijesztget minket* "they have a dog and **they** er and it always scares us" (9-year-old boy); ***nem lehet leérni*** *a földre a lufi* "**impossible** to go down to the ground the balloon" (6-year-old boy).

False starts: *az is nagyon* ***hu*** *vicces* "it is also very **hu**… funny" (11-year-old girl); *van ilyen lift meg* ***eme*** *ilyen de én lépcsővel szoktam felmenni* "there's an elevator and **floo** things but I usually go up by the escalator" (6-year-old girl).

False words: *az a* ***méz*** *vagyis az a méhecske* "that **honey**, that is, that bee" (6-year-old girl); ***három*** *testvérem van a Lali a Szilvi és én* "I have **three** bothers/sisters: Louie, Sylvia and me" (9-year-old girl).

Phonological errors: *gurulja gurulnak rajta* "they roll **it** roll on it" (6-year-old boy); *é***s***re észre vegyem* "so that I noti**sh** notice it" (6-year-old boy).

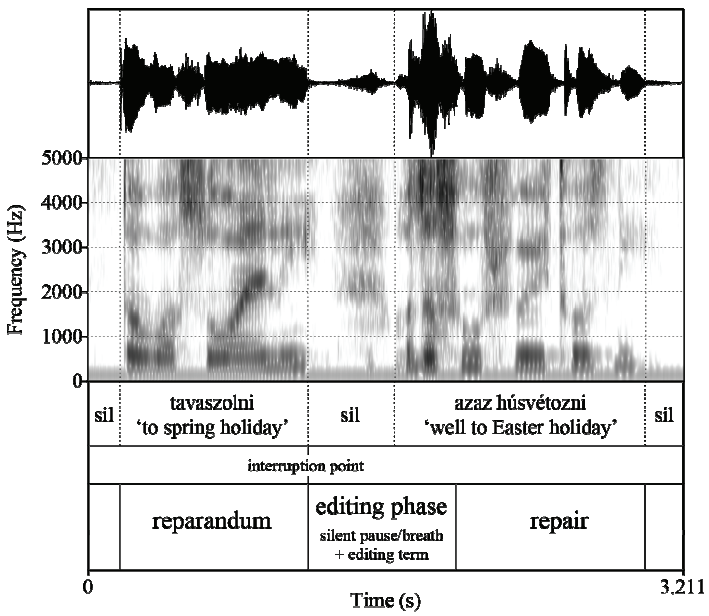

Figure 1. Sample of annotation of a false word in Praat
Slika 1. Ilustracija anotacije pogrešne riječi u programu Praat

Statistical analysis was conducted using SPSS 17.0 software (Univariate and Multivariate ANOVAs, repeated measures ANOVA, Tukey post hoc test, Pearson's correlation analysis, Kruskal-Wallis test, as appropriate). The confidence level was set at the conventional 95%.

3. RESULTS

3.1. Occurrences of disfluencies

Data of disfluency rate are given by two indicators: (i) occurrences per minute (see Table 4) and (ii) occurrences per 100 words (see Table 5). In the latter case all words, filler words, repetitions, and even word fragments were included in the word count.

Disfluencies occurred least frequently in the 7-year-old children's speech samples, and were most frequent in the 13-year-olds' speech samples. Univariate ANOVA revealed significant main effect of "age" for the occurrences: $F(4, 69) = 2.724$; $p = 0.037$; $\eta^2 = 0.144$. Tukey post hoc test showed significant differences between seven-year-olds and thirteen-year-olds ($p = 0.023$). As expected, we found strong positive correlation between the number of instances and speaking time (Pearson's correlation: $r = 0.710$; $p < 0.001$). The longer the children speak, the more disfluencies occur in their speech samples.

Table 4. Occurrences of disfluencies per minute across ages
Tablica 4. Broj disfluentnosti u jednoj minuti govora u svakoj dobnoj skupini

Age groups (years)	Mean occurrence (item/minute)	SD	Minimum (item/minute)	Maximum (item/minute)
6	14.3	4.81	7.6	25.8
7	10.8	4.87	3.1	18.4
9	14.9	4.08	8.9	22.0
11	14.0	5.49	3.3	21.9
13	16.4	4.23	9.4	23.5

We also analyzed the occurrences of disfluencies per 100 words (Table 5). In this case there were no significant differences among age groups. However, this measurement showed similar results to the former one. Seven-year-olds produced

the fewest disfluencies, and the standard deviation of the occurrences of the disfluencies was the highest in the group of 11-year-olds. The differences between the data obtained by different calculations (time dependency and word count dependency) are rooted in the fact that older children produced more words during the same amount of time than did the younger children. The largest difference between the two calculations can be seen in the cases of 13-year-olds and 6-year-olds. Disfluencies produced by 13-year-olds show that they are the most disfluent group according to the occurrence of disfluencies depending on speaking time while 6-year-olds were proved to be the most disfluent group according to the number of disfluencies per 100 words. Furthermore, there was significant correlation between the number of disfluencies and the number of words produced in spontaneous speech (Pearson's correlation: $r = 0.674$; $p < 0.001$). The more words a speech sample contains, the more likely it is to be interrupted by some kind of disfluency.

Table 5. Occurrences of disfluencies per 100 words across ages
Tablica 5. Broj disfluentnosti na 100 riječi u svakoj dobnoj skupini

Age groups (years)	Mean (item/100 words)	SD	Minimum (item/100 words)	Maximum (item/100 words)
6	17.5	5.20	10.5	31.4
7	13.7	6.88	4.8	24.7
9	17.3	6.04	10.8	30.3
11	16.1	7.35	4.4	27.7
13	17.1	4.90	10.0	25.2

3.2. Type-related occurrences of disfluencies

The most frequent types of disfluencies were filler words and filled pauses, which were produced (at least one instance) by 98.6% and 97.1% of the speakers, respectively. The least frequent ones among the analyzed types belonged to the type of false words, which was produced by about half of the participants (54.3%). Table 6 summarizes the type-related occurrences of the children's disfluencies in the total speech material and the proportion of children whose speech samples contained them.

Table 6. Distribution and frequency of the analyzed disfluency types
Tablica 6. Distribucija i učestalost analiziranih tipova disfluentnosti

	Type of disfluency	Occurrences (instances)	Frequency (% of speakers*)
unrepairable	filler word	1694	98.6
	filled pause	1258	97.1
	repetition	439	78.6
repairable	false start	190	80.0
	phonological error	119	68.6
	grammatical error	120	67.1
	false word	50	54.3

*The percentage of speakers shows the proportion of children whose speech samples contained instances of the given type of disfluency.

3.3. Occurrence of disfluency types per minute

Statistical analysis (multivariate ANOVA) yielded significant main effect of "age" in the three types of unrepairable disfluencies: filled pauses ($F(4, 69) = 5.052$; $p = 0.001$; $\eta^2 = 0.237$), filler words ($F(4, 69) = 3.512$; $p = 0.012$; $\eta^2 = 0.127$), and repetitions ($F(4, 69) = 10.175$; $p < 0.001$; $\eta^2 = 0.347$). Tukey post hoc test showed significant differences between the ages of 7 on the one hand and 9, 11, and 13 on the other in the case of the filled pauses; similarly, between the ages of 13 on the one hand and 6 and 7 on the other in the case of the filler words, as well as between the ages of 6 and all the other age groups in the case of repetitions ($p < 0.05$). The occurrence of filled pauses was similar with 6- and 7-year-olds, while a sharp increase was observed in the use of filled pauses after the age of 7. There were two 6-year-old children who did not produce any filled pauses in their speech samples. We found an increasing tendency of the occurrences of filler words after the age of 7. Repetitions were more frequent in preschool children's speech samples than in school-age children's speech samples. Filler words and filled pauses were the most frequent disfluency phenomena in cases where the children seemed to need more time to overcome their speech production difficulties (Figure 2). The occurrence of repetitions decreased after the age of 6. The data show that repetitions are characteristic of preschool children, while filled pauses and filler words are largely produced by school-age children after the age of 7, with an increasing frequency in the case of filler words.

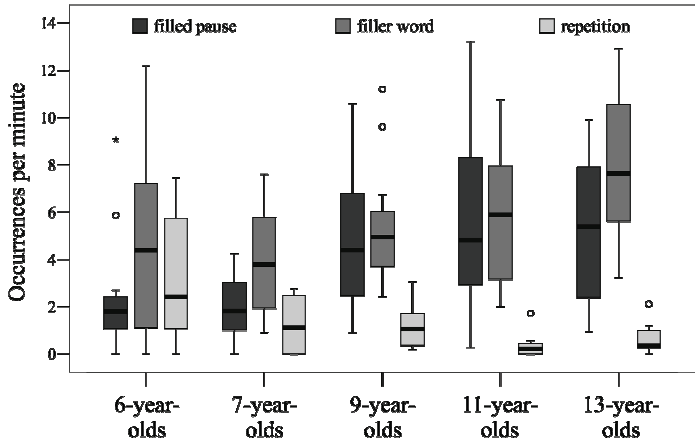

Figure 2. Occurrences of unrepairable disfluencies across ages
Slika 2. Broj nepopravljivih disfluentnosti u svakoj dobnoj skupini

We analyzed the occurrences of the four types of repairable disfluencies: grammatical errors, false words, false starts, and phonological errors. Overall, the occurrences of repairable disfluencies showed a slightly decreasing tendency with age; however, statistical analysis did not show significant main effect of "age" in this respect (Figure 3).

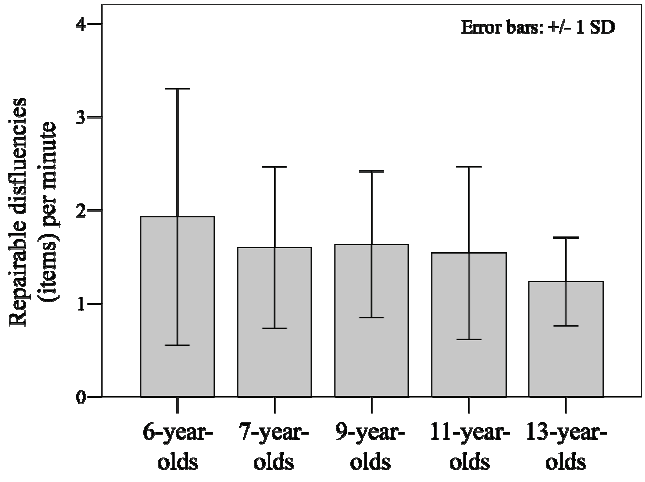

Figure 3. Mean values of repairable disfluencies per minute across ages
Slika 3. Prosjek popravljivih disfluentnosti u minuti govora u svakoj dobnoj skupini

In addition, "age" had no significant main effect on any type of repairable disfluency, either. However, some tendencies could be observed. The data in Figure 4 indicate that false starts tend to occur most frequently in children's speech samples, except for 6-year-olds who produced more grammatical errors than false starts. As expected, grammatical planning and execution seemed to be more difficult for the younger speakers (6- and 7-year-olds) than for the older ones.

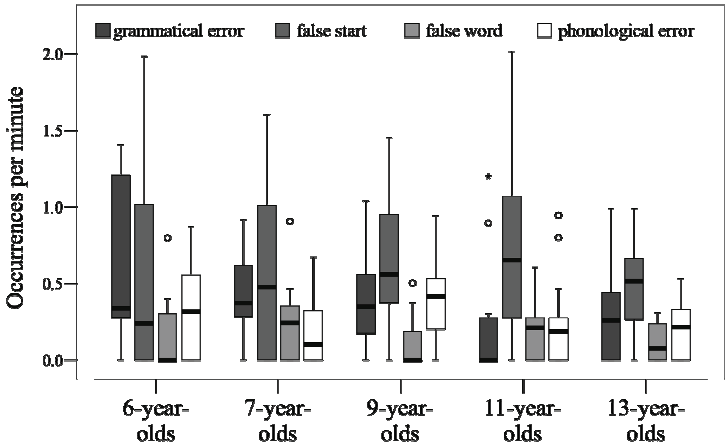

Figure 4. Occurrences of repairable disfluencies across ages
Slika 4. Broj popravljivih disfluentnosti u svakoj dobnoj skupini

Pearson's correlation analysis revealed significant relations among the occurrences (per minute) of various types of disfluency. Some of the disfluency types showed slight but significant correlation (Table 7). Medium-strong correlation was found between repetitions and grammatical errors. The slight negative correlation between the occurrences of filler words and false words indicates that the more filler words occurred, the less false words were produced by the children.

Table 7. Correlation data between the occurrences of various disfluency types (Correlation is significant at the 0.05 level = * and at the 0.01 level = **)
Tablica 7. Koeficijenti korelacija između različitih tipova disfluentnosti (Značajnost korelacije na razini 0,05 označena je sa *, a na razini 0,01 sa **.)

Disfluency types		Correlation coefficient (*r*)
filled pause	false start	0.270*
filler word	false word	-0.333**
repetition	grammatical error	0.448**
	phonological error	0.244*

3.4. Occurrence of disfluency types per 100 words

The calculations of the occurrence of disfluencies per 100 words show another aspect of frequency (Figure 5). The present section describes these data following the same criteria that were considered in the former analysis. Comparing the occurrences of disfluencies per minute and 100 words, there are more similarities than differences (Figures 2 and 5). The differences appear in the range of occurrences suggesting that the individual differences are slightly larger when calculating with words than when calculating with time. Statistical analysis (multivariate ANOVA) yielded significant main effect of "age" in two types of unrepairable disfluencies: filled pauses ($F(4, 69) = 3.093$; $p = 0.022$; $\eta^2 = 0.160$), and repetitions ($F(4, 69) = 11.664$; $p < 0.001$; $\eta^2 = 0.418$), but not in the case of filler words. Tukey post hoc test showed significant differences between the ages of 6 and the other four age groups in the case of the repetitions ($p < 0.005$).

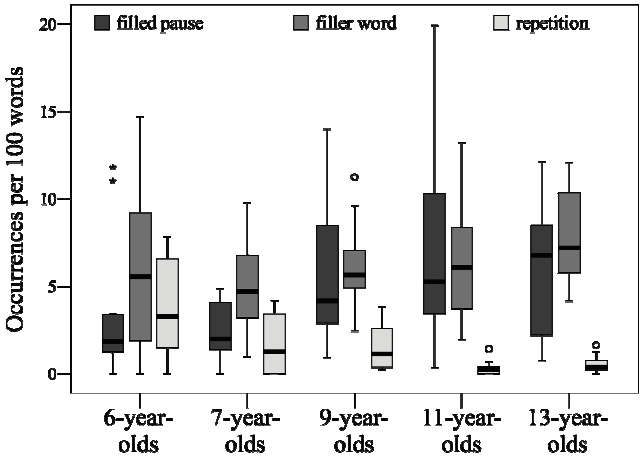

Figure 5. Occurrences of unrepairable disfluencies per 100 words across ages
Slika 5. Broj nepopravljivih disfluentnosti na 100 riječi u svakoj dobnoj skupini

Statistical analysis did not show significant main effect of "age" in terms of the total occurrences of repairable disfluencies per 100 words. The only difference between the data of the two kinds of calculations is the direction of change. Repairable disfluencies decrease continuously between the ages of 6 and 13 with a larger drop between the ages of 11 and 13 when analyzing them in terms of 100 words produced (Figure 6), while they show a drop between the ages of 6 and 7

counting them in terms of speaking time, where there is no change between the ages of 7 and 11, and there are fewer disfluencies at the age of 13 than at the age of 11. These slight differences might support the view that the occurrence of repairable disfluencies really continuously decrease depending on age; however, the occurrence of pauses (figuring in the calculations depending on speaking time) slightly influence the age-dependent differences in the occurrence of these types of disfluencies.

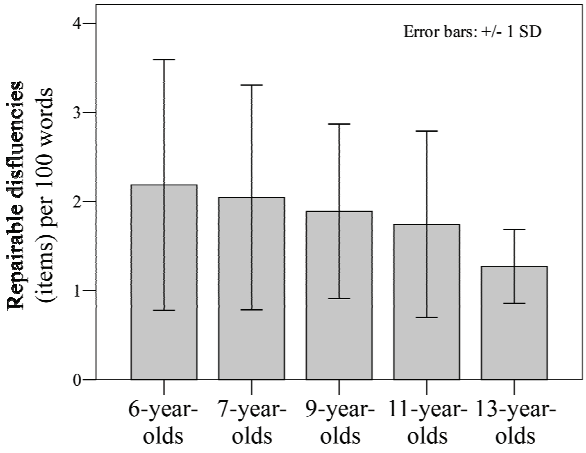

Figure 6. Mean values of repairable disfluencies per 100 words across ages
Slika 6. Prosjek popravljivih disfluentnosti na 100 riječi u svakoj dobnoj skupini

Statistical analysis (considering occurrences per 100 words) revealed significant main effect of "age" on the occurrence of grammatical errors (multivariate ANOVA: $F(4, 69) = 2.798$; $p = 0.033$; $\eta^2 = 0.147$). Tukey post hoc test showed significant differences between the ages of 6 and 7 ($p = 0.043$). No such significant effects were found in the cases of the other three types of repairable disfluencies. Decrease of the grammatical speech errors can be seen between the ages of 6 and 11 while the 13-year-olds produced more grammatical errors than the 11-year-olds did (Figure 7).

Comparing the data of occurrences in the two kinds of calculations, no large differences can be seen. The tendencies are very similar, even in the case of grammatical errors where statistically significant differences were found when the data were analyzed in terms of word count.

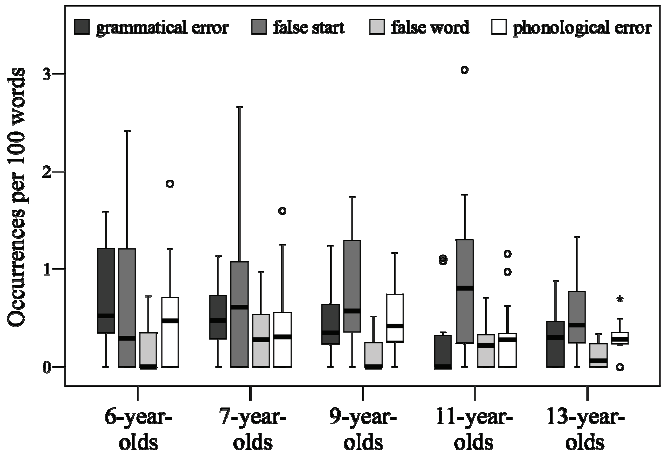

Figure 7. Occurrences of repairable disfluencies per 100 words across ages
Slika 7. Broj popravljivih disfluentnosti na 100 riječi u svakoj dobnoj skupini

3.5. Ratio of self-repairs

The mean ratios of self-repairs were 45.86%, 65.88%, 56.48%, 78.43%, and 65.71% of all speech errors produced in each age group (respectively). The repairs of false starts (total mean: 94%) were the most frequent ones in most age groups, while repairing grammatical errors (total mean: 33%) occurred least frequently in our material. On average, children repaired 85% of all false words, while only 40% of all phonological errors were repaired. Grammatical errors were repaired at a very low rate by preschool children (8.57%), while 13-year-olds repaired half of them. Children repaired almost all of their false starts across all age groups (the few unrepaired instances were followed by a totally different grammatical construction, showing no connection with the false start), 13-year-olds' repairing ratio is a bit lower than that of all the others. There was an increasing tendency to repair false words after the age of 6. Phonological errors were repaired more frequently by the 6- and 7-year-olds than by school-age children older than 9 years (Table 8).

Repeated measures ANOVA showed that the repair rates significantly depended on disfluency type: $F(3, 4) = 31.395$; $p < 0.001$; $\eta^2 = 0.887$). Bonferroni pairwise comparisons showed that repair rates of grammatical errors significantly differed from repair rates of false starts and false words ($p = 0.009$ in both cases), and repair rates of phonological errors significantly differed from those of false starts ($p = 0.003$).

Table 8. Details of repair rates of repairable disfluency types across ages (100% = all incidents of disfluency)

Tablica 8. Učestalost samoispravljanja popravljivih tipova disfluentnosti u svakoj dobnoj skupini (100% = sve disfluentnosti)

Age groups (years)	Repaired rates of repairable disfluency types (%)				
	grammatical errors	false starts	false words	phonological errors	total
6	8.57	93.55	77.78	53.62	45.86
7	40.91	93.33	69.23	46.03	65.88
9	29.03	97.87	87.50	26.67	56.48
11	37.50	98.21	92.31	42.86	78.43
13	50	88.46	100	30.84	65.71
mean	33	94	85	40	62.47

3.6. Editing phases of self-repairs

The durations of the editing phases were analyzed in order to get information about the time span that children's repairing mechanism needed. The mean durations of the editing phases were 472 ms in 6-year-olds, 693 ms in 7-year-olds, 445 ms in 9-year-olds, 545 ms in 11-year-olds, and 430 ms in 13-year-olds. Children may have good strategies for monitoring and repairing their speech since immediate repairs (0 ms editing phases) were found in all age groups. The editing phase was 0 ms in 6% of all repairs in 6-year-olds, 14% in 7-year-olds, 11% in 9-year-olds, 15% in 11-year-olds, and 11% in 13-year-olds. Immediate repair (0 ms editing phase) was observed in 18% of all phonological errors, in 17% of all grammatical errors, in 11% of all false starts, and in 7% of all false words. These figures mean that the repair was carried out without the interruption of the speech flow and they were spotted by the speaker irrespective of age and disfluency type.

In our corpus, the durations of (non-null) editing phases were the shortest in phonological errors (mean: 276 ms) and the longest in the case of false words (mean: 571 ms) (Figure 8). Self-repairs of false starts had a mean duration of 489 ms, while those of grammatical errors had 455 ms, on average. This finding indicates that repairing a false word seems to be the most difficult process for children. Statistical

analysis (Kruskal-Wallis test) revealed that "disfluency type" had a significant effect on the duration of the editing phases ($\chi^2 = 8.074$; $p = 0.045$).

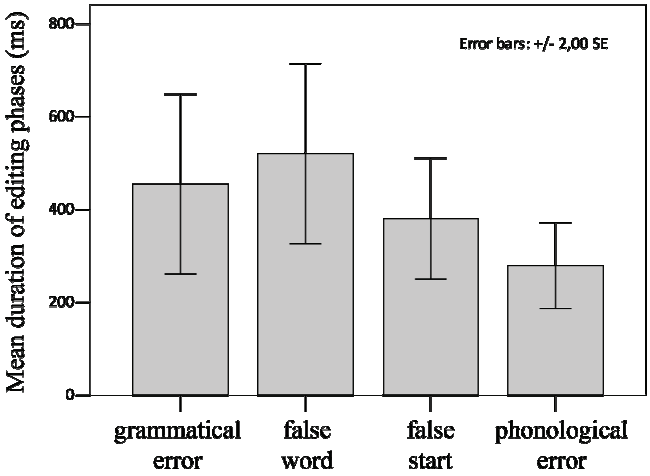

Figure 8. Mean duration of editing phases of each disfluency type
Slika 8. Prosječno trajanje faza kontrole u svakom tipu disfluentnosti

The durations of the editing phases were also analyzed according to disfluency types. Statistical analysis showed neither significant differences related to "age", nor clear tendency of change across ages (Figure 9). Durations of the editing phases of phonological errors and false words seem to be similar across ages, showing shorter durations in the case of 7-year-olds and increasing durations in the case of 13-year-olds. Roughly the opposite tendency can be observed in the durations of the editing phases in false starts. 6-year-olds and 13-year-olds have shorter durations in repairing false starts than in repairing false words, while 9- and 11-year-olds do not show much difference in durations of the editing phases of these two types of disfluency. 7-year-olds show a totally different pattern: they need longer editing phases for repairing false starts than for repairing false words.

There is a decreasing tendency of the durations of the editing phases in grammatical errors across changes. The older the children the shorter time they need to repair their grammatical errors. 7- and 13-year-olds, however, seem to contradict this tendency. In addition, no correlation was found between the duration of editing phases and the frequency of repairing the disfluencies.

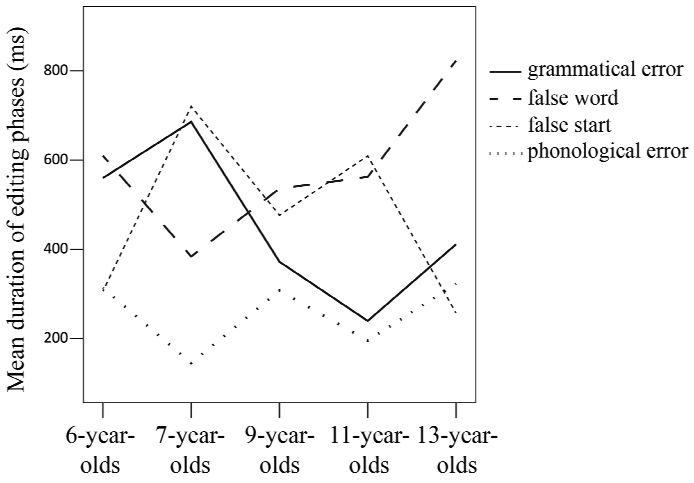

Figure 9. Mean duration of editing phases of each disfluency type across ages
Slika 9. Prosječno trajanje faza kontrole u svakom tipu disfluentnosti u svakoj dobnoj skupini

4. CONCLUSIONS

Our study focused on occurrences and self-repair characteristics of Hungarian children's narratives between the ages of 6 and 13. Data on types and occurrences of disfluencies seem to confirm that these phenomena are related to speech planning processes in children's spontaneous speech similarly to that of adults. As expected, the changes of speech fluency across ages may be associated with developing language proficiency (Haynes & Hood, 1978; Zuckerman Pearl & Bernthal, 1980; Bernstein Ratner & Sih, 1987; McLaughin & Cullinan, 1989; Yaruss et al., 1999; Watson et al., 2011). In addition, individual differences in the occurrences of various types of disfluency experienced in adults (e.g. Boutsen & Hood, 1997; Bortfeld et al., 2001) are characteristic also of children.

We hypothesized that the occurrence of disfluencies would increase with age between 6 and 13 years as a consequence of grammatical complexity and increasing vocabulary of children in these ages. However, we could not unambiguously confirm this hypothesis. The increase of disfluencies was not a linear increase with age, irrespective of the calculation method (focusing on time or on the number of words). This finding might refer to a more or less constant quantity of disfluent episodes in 6-, 9-, and 11-year olds. The only statistically confirmed difference, taking into

consideration all disfluencies, was found between 7-year-olds and 13-year-olds: the former produced less, while the latter produced more of them than children in all the other age groups. These differences might be explained by starting school. We suppose that 7-year-olds behaved during recordings as if in a task situation similar to those they experience at school. They made efforts to suit the task. Older school-age children will have participated in the recording sessions as an everyday situation and not as a task-solving problem. In addition, further statistical analyses revealed significant differences among age groups in terms of the number of occurrences of filled pauses, filler words and repetitions. These findings support the developing language use of children and their longer and more complex utterances (Neuberger, 2013). The strong positive correlation found between the number of instances of disfluency and speaking time suggests that the speech planning process of children is still not well developed when longer continuous speech is required from them.

Filled pauses and filler words are more frequent in the speech samples of older than in those of the younger children, meaning that the older children use these disfluencies more efficiently in order to get some help in their speech planning processes. A characteristic change could be observed in the use of repetitions between younger and older children. The occurrence of the repetitions gets rarer as the child grows older; and they are replaced by filled pauses and filler words. English-speaking three-year-olds produced significantly more repetitions than five-year-olds did (DeJoy & Gregory, 1985), indicating a characteristic change in the use of repetitions during childhood. This finding leads us to think that changes in the occurrences of some disfluency types might show the children's conscious strategies to overcome their speech planning difficulties. Repetitions of words seem to become troublesome for children after the age of 7 on the one hand, and they are supposed to hear far more filled pauses than repetitions also from the adult speakers.

Our results show that children learn to cope with disfluency phenomena in two ways: (i) they seem to acquire the necessary strategies to deal with disfluencies, using some types of them rather than others and repairing their speech errors, and (ii) they (consciously or unconsciously) change the frequency of using various types of disfluency across ages.

Hungarian-speaking adults' narratives contain 9.47 occurrences of filled pauses per 100 words (Beke et al., 2014). Our children between the ages of 6 and 13 produced 4.79 filled pauses per 100 words. The number of occurrence in all ages is less than that observed in adults. Accepting the main function of using filled pauses

in spontaneous speech irrespective of age, that is to aid the speakers thinking and formulating their thoughts into grammatical and phonological forms, the difference between children and adults seems to reflect the less complex and shorter utterances of children compared to those of adults (Neuberger, 2013). Therefore, children need the help of filled pauses less frequently than adults do.

The occurrence of repairable disfluencies shows a slightly decreasing tendency across ages; this can particularly be seen between the preschool children and the school-age children. Two- and three-year-old children were reported to have seven times more speech errors than adults do (Stemberger, 1989). Our 6-year-old children produced 6.6 times more speech errors than adults did, while the error rate of 13-year-old children was 5 times higher than that of adults (see Gósy & Gyarmathy, 2014). The decrease of these speech errors reflects the developing language acquisition of children even between the ages of 6 and 13.

As expected, the occurrence of grammatical errors changes significantly between the preschool children and the 11-year-old group. The primary reason for this seems to be that children learn to read and write at school and they start using language more consciously than they did earlier. The increased occurrence of grammatical errors in the 13-year-olds' narratives can be explained by their more complex and longer utterances than those of the younger children.

Our second hypothesis was that the ratio of self-repairs would depend on the children's age and the type of disfluency. Both statements could be confirmed. There was a tendency showing that the repairing ratio of grammatical errors, false starts and false words increased as the children grew older. No exception was found in this respect in grammatical errors, while slightly opposite difference was found in false starts in the case of 13-year-olds and in false words in the case of 7-year-olds (in both cases the younger children repaired a little more errors than the older ones). No clear tendency of repairing ratio could be observed in the case of phonological errors across ages. This finding can be explained primarily by the fact that various types of errors occurred in this group (see Table 2). Children and adults seem to behave differently in successfully repairing the phonological errors and false starts. Children repair almost all their false starts (94%) but only 40% of their phonological errors. Adults, however, repair their phonological errors to a larger extent (63.5%, see Gósy & Gyarmathy, 2014) than their false starts (56.35%, see Horváth & Gyarmathy, 2012). The explanation for this difference can be found in children's more conscious operations of lexical retrieval than their phonological planning. The latter seems to

be more automatic and less controlled at these ages. Adults seem to be more attentive to their own phonological errors in order to ensure the listeners' speech perception. As expected, the self-monitoring mechanism of older children seems to be more efficient. The ratios of the children's self-repairs may be affected by individual characteristics, and the level of their language acquisition (e.g. their own fluency in spontaneous utterances), see Laakso, 2010.

Our third hypothesis concerned the duration of the editing phases of repairable disfluencies. Data show that these values depend primarily on the type of the disfluency and not on the children's age. We assumed that the durations of the editing phases would decrease with age; however, this assumption could not be confirmed. Analyzing the repairing patterns of each age group, we can see that repairing false words and false starts needs longer time in all age groups than the necessary durations to repair grammatical and phonological errors. The increase of the durations of the editing phases in repairing false words between the ages of 7 and 13 is a consequence of the developing mental lexicon. The more numerous the words the children can select from, the more difficulties might arise in repairing selection errors. The facts that repairing grammatical and phonological errors is less frequent and needs more time than repairing the two others indicate that children have no efficient strategy to cope with these errors. Since this is different from the adults' patterns, we can assume that it is characteristic of the process of first language acquisition.

Despite the individual differences of children's speech behavior, our findings help to evaluate the typical characteristics of disfluent episodes in children's spontaneous utterances between the ages of 6 and 13. These data can also be used as typical and comparable patterns when analyzing atypical language acquisition in children (Thordardottir & Weismer, 2002).

REFERENCES

Ambrose, N. G., Yairi, E. (1999). Normative disfluency data for early childhood stuttering. *Journal of Speech, Language, and Hearing Research* **42**, 895–909.

Beke, A., Gósy, M., Horváth, V., Gyarmathy, D., Neuberger, T. (2014). Disfluencies in spontaneous narratives and conversations in Hungarian. In S. Fuchs, M. Grice, A. Hermes, L. Lancia, D. Mücke (eds.), *Proceedings of the 10th International Seminar on Speech Production (ISSP)*, 29–32.

Bernstein Ratner, N., Sih, C. C. (1987). Effects of gradual increases in sentence length and complexity on children's disfluency. *Journal of Speech and Hearing Disorders* **52**, 3, 278–287.

Boersma, P., Weenink, D. (2011). Praat: Doing phonetics by computer, http://www.praat.org/praat.org [accessed 10th Sept 2011].

Boey, R. A., Wuyts, F. L., Van de Heyning, P. H., De Bodt, M. S., Heylen, L. (2007). Characteristics of stuttering-like disfluencies in Dutch-speaking children. *Journal of Fluency Disorders* **32**, 4, 310–329.

Bortfeld, H., Leon, S. D., Bloom, J. E., Schober, M. F., Brennan, S. E. (2001). Disfluency rates in conversation: Effects of age, relationship, topic, role, and gender. *Language and Speech* **44**, 2, 123–147.

Boutsen, F. R., Hood, S. B. (1997). Determinants of speech rate and fluency in fast and slow speaking normally fluent children. In W. Hulstijn, H. F. M. Peters, P. H. H. M. Van Lieshout (eds.), *Speech Production: Motor Control, Brain Research and Fluency Disorders*, 557–564. Amsterdam: Elsevier.

Carlo, E. J., Watson, J. B. (2003). Disfluencies of 3- and 5-year old Spanish-speaking children. *Journal of Fluency Disorders* **28**, 1, 37–53.

DeJoy, D. A., Gregory, H. H. (1985). The relationship between age and frequency of disfluency in preschool children. *Journal of Fluency Disorders* **10**, 2, 107–122.

Evans, M. A. (1985). Self-initiated speech repairs: A reflection of communicative monitoring in young children. *Developmental Psychology* **21**, 2, 365–371.

Forrester, M. (2008). The emergence of self repair: A case study of one child during the early pre-school years. *Research on Language and Social Interaction* **41**, 99–128.

Gaines, N. D., Runyan, C. M., Meyers, S. C. (1991). A comparison of young stutterers' fluent versus stuttered utterances on measures of length and complexity. *Journal of Speech and Hearing Research* **34**, 1, 37–42.

Gósy, M., Gyarmathy, D. (2014). Szublexikális nyelvbotlások jellemzői a spontán beszédben [Characteristics of sublexical errors in spontaneous speech]. *Magyar Nyelvőr* **138**, 3, 297–307.

Guo, L., Tomblin, J. B., Samelson, V. (2008). Speech disruptions in the narratives of English-speaking children with specific language impairment. *Journal of Speech, Language, and Hearing Research* **51**, 3, 722–738.

Hall, N., Wagovich, S., Bernstein Ratner, N. (2007). Language and childhood stuttering. In E. G. Conture, R. F. Curlee (eds.), *Stuttering and Related Disorders of Fluency*, 153–167. NY: Thieme.

Haynes, W. O., Hood, S. B. (1978). Disfluency changes in children as a function of the systematic modification of linguistic complexity. *Journal of Communication Disorders* **11**, 1, 79–93.

Horváth, V., Gyarmathy, D. (2012). Téves kezdések magyarázó elvei a spontán beszédben [Explanations of false starts in spontaneous speech]. In M. Gósy (ur.), *Beszéd, adatbázis, kutatások [Speech, database, research]*, 138–153. Budapest: Akadémiai Kiadó.

Howell, P., van Borsel, J. (eds.) (2011). *Multilingual Aspects of Fluency Disorders*. Bristol: Multilingual Matters.

Hudson Kam, C. L., Edwards, N. A. (2008). The use of *uh* and *um* by 3- and 4-year-old native English-speaking children: Not quite right but not completely wrong. *First Language* **28**, 313–327.

Juhász, A. (ed.) (1999). *Logopédiai vizsgálatok kézikönyve [Manual for Speech and Language Tests]*. Piliscsaba: Új Múzsa Kiadó.

Kidd, C., White, K. S., Aslin, R. N. (2011). Toddlers use speech disfluencies to predict speakers' referential intentions. *Developmental Science* **14**, 4, 925–934.

Laakso, M. (2010). Children's emerging and developing self-repair practices. In H. H. Gardner, M. Forrester (eds.), *Analysing Interactions in Childhood: Insights from Conversation Analysis*, 74–99. Chichester: Wiley-Blackwell.

Levelt, W. J. M. (1983). Monitoring and self-repair in speech. *Cognition* **14**, 1, 41–104.

Levelt, W. J. M. (1989). *Speaking. From Intention to Articulation*. Cambridge, MA: MIT Press.

MacLachlan, B. G., Chapman, R. S. (1988). Communication breakdowns in normal and language learning-disabled children's conversation and narration. *Journal of Speech and Hearing Disorders* **53**, 2–7.

McLaughin, S. F., Cullinan, W. L. (1989). Disfluencies, utterance length, and linguistic complexity in nonstuttering children. *Journal of Fluency Disorders* **14**, 1, 17–36.

Newman, R., Bernstein Ratner, N. (2007). The role of selected lexical factors on confrontation naming accuracy, speed and fluency in adults who do and do not stutter. *Journal of Speech, Language and Hearing Research* **5**, 196–213.

Neuberger, T. (2013) *Életkor-specifikus sajátosságok a spontán beszéd fejlődésében [Age-specific Properties in Spontaneous Speech Development]*. PhD dissertation, Budapest: ELTE BTK.

Salonen, T., Laakso, M. (2009). Self-repair of speech by four-year-old Finnish children. *Journal of Child Language* **36**, 855–882.

Shriberg, E. (1994). *Preliminaries to a Theory of Speech Disfluencies*. PhD dissertation, University of California at Berkeley.

Stemberger, J. P. (1989). Speech errors in early child language production. *Journal of Memory and Language* **28**, 2, 164–188.

Thordardottir, E. T., Weismer, S. (2002). Content mazes and filled pauses in narrative language samples of children with specific language impairment. *Brain and Cognition* **48**, 564–576.

Tumanova, V., Conture, E. G., Lambert, E. W., Walden, T. A. (2014). Speech disfluencies of preschool-age children who do and do not stutter. *Journal of Communication Disorders* **49**, 25–41.

Watson, J. B., Byrd, C. T., Carlo, E. J. (2011). Disfluent speech characteristics of monolingual Spanish-speaking children. In P. Howell, J. van Borsel (eds.), *Multilingual Aspects of Fluency Disorders*, 169–191. Bristol: Multilingual Matters.

Yairi, E., Ambrose, N. G. (2005). *Early Childhood Stuttering: For Clinicians by Clinicians*. Austin, TX: Pro-Ed.

Yaruss, J. S., Newman, R. M., Flora, T. (1999). Language and disfluency in nonstuttering children's conversational speech. *Journal of Fluency Disorders* **24**, 3, 185–207.

Zuckerman Pearl, S., Bernthal, J. E. (1980). The effect of grammatical complexity upon disfluency behavior of nonstuttering preschool children. *Journal of Fluency Disorders* **5**, 55–68.

Tilda Neuberger, Mária Gósy
neuberger.tilda@nytud.mta.hu, gosy.maria@nytud.mta.hu
Istraživački institut za lingvistiku, Mađarska akademija znanosti
Mađarska

Krossekcijsko istraživanje obilježja disfluentnosti u spontanom govoru djece

Sažetak

Budući da su i pogrešni i ispravni govorni odsječci oblikovani prema istim pravilima govorne proizvodnje, analiza pogrešaka i samoispravljanja može pomoći u istraživanjima skrivenih postupaka u planiranju govora, kao i u istraživanju razvoja govora. Cilj je ovog istraživanja analizirati utjecaj dobi na tipove pogrešaka i ispravljanja u govoru djece. Na uzorku spontanog govora sedamdesetero monolingvalne mađarske djece u dobi od 6 do 13 godina analizirani su tipovi disfluentnosti, učestalost samoispravljanja i trajanje faza kontrole. Rezultati potvrđuju da učestalost samoispravljanja ovisi o dobi djeteta i o tipu disfluentnosti. Rezultati krossekcijske analize ukazuju na promjene u fluentnosti govora koje mogu biti povezane s razvojem govora.

Ključne riječi: disfluentnost, samoispravljanje, spontani govor, monolingvalna mađarska djeca