Izvorni znanstveni rad
Rukopis primljen 23. 10. 2023.
Prihvaćen za tisak 24. 1. 2024.
https://doi.org/10.22210/govor.2023.40.10

Ana Vidović Zorić
anvidovi@ffzg.unizg.hr
Faculty of Humanities and Social Sciences, University of Zagreb
Croatia

Sara Blažeković
blazekovicsara11@gmail.com
Osijek, Croatia

Pragmatic function of speech disfluencies in high-functioning children with autism spectrum disorder

Summary

Disfluencies in speech are usually viewed as manifestation of problems at the level of inner cognitive processes underlying speech production. However, it has been suggested that some types of disfluencies may also have a pragmatic function. The current study examines the distribution of disfluencies and speech rate in spontaneous speech of 10 high-functioning children with autism spectrum disorder (diagnosed as Asperger syndrome or high-functional autism) as compared to the typically developing controls. The task of participants was to retell the cartoon which they had previously watched. Their speech was audio recorded and transcribed. The duration of the analyzed speech segment was 60 seconds. Our results show that there is no significant difference in the total number of disfluencies between the two groups. Furthermore, participants of both groups produced the same types of disfluencies. Given that disfluencies reflect troubles in speech planning, these findings indicate that disfluencies are indeed a universal phenomenon in speech production. As regards different types of disfluencies, our study shows that participants with autism spectrum disorder (ASD) produced significantly more disfluent silent pauses and significantly fewer filled disfluencies than the controls. As has already been suggested in the literature, this could be due to the deficiency in the pragmatic ability of individuals with ASD. That is, individuals with ASD are less engaged in the maintenance of speech flow, which will result in more silent pauses.

Moreover, speech rate in the ASD group was significantly lower than in the controls, which could also be explained in light of the pragmatic deficit in individuals with ASD.

Keywords: speech disfluencies, Asperger syndrome, high-functioning autism, autism spectrum disorder, pragmatics

1. DISFLUENCIES IN TYPICAL SPEECH

Fluent speech is one of speaker's main goals during the process of speech production. Fluency could be defined in many ways, depending on the approach to this phenomenon. In psycholinguistic literature, fluent speech is usually comprehended as "rapid, smooth, accurate, lucid, and efficient translation of thought or communicative intention into language" (Lennon, 2000: 26). This definition is based on the perspective of language as cognitive process, underlying speech production, whose final product is a spoken utterance. In typical native speakers, this process is highly automatized and effortless (Levelt, 1989; Schmidt, 1992; Segalowitz, 2003; Tavakoli & Wright, 2020). Nevertheless, disfluencies in spontaneous speech are almost inevitable. According to Fox Tree (1995), there are around 6 disfluencies per 100 words. In typical speech, disfluencies are usually viewed as signals of difficulties at some level of the planning process (Clark & Fox Tree, 2002; Corley & Stewart, 2008; Levelt, 1989), i.e., at the level of conceptualization, lemma retrieval, grammatical or phonological encoding, or motor planning. The main categories of disfluencies are silent pauses, filled pauses (such as *uh* and *um*), interruptions, self-repairs, repetitions of phonemes, syllables, words or longer phrases and sound prolongations (Postma, Kolk, & Povel, 1990).

However, the psycholinguistic aspect on disfluencies is not the only one. It has been suggested that disfluencies are not just the manifestation of troubles in the speech planning process, but at the same time they also provide certain metalinguistic information about the speaker, e.g., speaker's mental state (Bortfeld, Leon, Bloom, Brennan, & Schober, 2001). Furthermore, they could improve the listener's comprehension of the message. For example, Brennan and Schober (2001) found that listeners responded faster to target words when they were followed by disfluency than if they were not. Fox Tree (2001) found that *uh* could improve the recognition of the upcoming word, while *um* does not have this kind of effect on word recognition. These results indicate that different types of fillers do not have a unique communicative function. Finally, disfluencies could also have an important role in coordinating conversational interaction (Bortfeld et al., 2001; Clark, 1994; Clark & Fox Tree, 2002; Fox Tree, 2001; Levelt, 1989). For example, filled pauses such as *uh* and *um* could signal to the listener that the speaker is not yet finished with own conversational turn, but just has a delay or difficulty at some stage of speech production. Clark (1994) concluded that *uh* and *um* have a function to warn the listener about the upcoming problem. Similarly, Fox Tree (2001) assumed that *uh* is a signal for shorter upcoming delay, as compared to the *um* which indicates longer upcoming delay in speech production. Furthermore, Levelt (1989: 482) pointed out that editing expressions, such as *er*, *that is*, *sorry*, *I mean* "play a significant role in signaling to the addressee that there is a trouble". Levelt (1989: 481) also concluded that "by interrupting a word, a speaker signals to the addressee that the word is an error. If a word is completed, the speaker intends the listener to interpret it as correctly delivered". Levelt interpreted this kind of signal as pragmatic. In fact, all three later aspects of disfluencies (providing additional information of speaker's mental state, improving listener's comprehension of the message, and coordinating conversational interaction) could be considered within the pragmatic function. In line with this pragmatic view on disfluencies are some findings (Oviatt, 1995; Shriberg, 2001) that the rate of some types of disfluencies is higher in human–human than in human–computer dialogue. That is, a computer has no awareness or knowledge of the communication context and the relationship between communication participants. Thus, it does not have the ability to interpret metalinguistic signals, which in this case are disfluencies. Bearing this fact in mind, the speaker will use this type of signal less often in the dialogue with the computer.

However, it should be noted that the assumption of the communicative function of disfluencies has been questioned by some researchers. For example, Finlayson and Corley (2012) investigated the number of disfluencies between a dialogue and a monologue, as well as the distribution of different types of disfluencies. If disfluencies indeed have a communicative function, one would expect a higher disfluency rate in a dialogue than in a monologue. But the results of Finlayson and Corley showed no difference between these two types of narratives. The authors concluded that there was no straightforward evidence that disfluencies are intentional signals to the listeners. However, both narratives were obtained under experimental circumstances. Therefore, participants were aware that their recordings will be listened by someone at some point, which raises the question whether these two conditions really are different from the speaker's perspective.

2. DISFLUENCIES IN AUTISM SPECTRUM DISORDER

Autism spectrum disorder (ASD) is described as "a neurodevelopmental disorder, characterized by persistent impairment in reciprocal communication and social interactions as well as restricted repetitive pattern of behaviors, interests or activities" (American Psychiatric Association [APA], 2013; Campisi, Nazish Imran, Nazeer, Skokauskas, & Waqar Azeem, 2018: 92). Clinical manifestations of autism are in the range from severe to high-functioning autism (Mazzone, Ruta, & Reale, 2012). According to the latest classification of American Psychiatric Association (2013), diagnoses like Asperger's syndrome are no longer separated, but are considered to be a specific manifestation of the same disorder – autism spectrum disorder (ASD). Furthermore, in the same diagnostic manual it is stated that "the symptoms of people with ASD will fall on a continuum, with some individuals showing mild symptoms and others having much more severe symptoms" (American Psychiatric Association, 2013). Still, many scientific papers investigated speech and language abilities in high-functioning children with ASD who were previously diagnosed with Asperger's syndrome (AS) or high-functioning autism (HFA). Moreover, in literature, the term high-functioning autism (HFA) often overlaps with the term Asperger's syndrome (AS) (Mazzone et al., 2012) or they are used interchangeably (e.g., Colle, Baron-Cohen, Wheelwright, & van der Lely, 2008; Freitag, Kleser, Schneider, & von Gontard, 2007; Lee, Liang, Hou, Tse, & Chan, 2023). Unlike classic autism, individuals with HFA/AS are described as having language and cognitive abilities within the limits of typical development (Bishop, 2003; Freitag et al., 2007; Gibson, Adams, Lockton, & Green, 2013; Shriberg et al., 2001), although some studies reported language impairments in some cases of HFA (Dai, He, Chen, & Yin, 2022). However, behavioral patterns typical for classic ASD, such as problems in social interaction and communication, repetitive interests, and behaviors, are characteristic features of HFA/AS (Mazzone et al., 2012). These difficulties in social interactions and communication are attributed to the deficit of the theory of mind (ToM) (Baron-Cohen, 1995, 2000). ToM is explained as "the ability to attribute mental states to another person and to infer their underlying intentions, thoughts, emotions and motivation" (Colle et al., 2008: 28). In other words, one of the main characteristics of ASD is the lack of pragmatic function of language. In order to investigate whether this pragmatic deficit would affect narrative skills in individuals with ASD, Colle et al. (2008) compared narrative discourse of adults with HFA or AS with the typically developing individuals. Their results showed that participants with ASD used fewer anaphoric pronouns, as well as temporal and referential expressions in the narration than the controls, i.e., expressions which require pragmatic knowledge. Given that these expressions directly depend on the ToM ability, the authors found these results as evidence that language difficulties in individuals with ASD are mainly a consequence of the impaired pragmatic function, since "pragmatics and ToM cannot be separated" (Colle et al., 2008: 30).

ASD participants did not differ in other narrative abilities, such as following the main plot, quantity of information, number of produced words and episodes, as well as in morpho-syntactic complexity, from the controls. Some other studies (Losh & Capps, 2003; Seung, 2007) also found that individuals with autism do not have difficulties in telling a story. The length of their narratives and grammatical complexity did not differ from their typically developing peers (Tager-Flusberg & Sullivan, 1995). However, Rumpf, Kamp-Becker, Becker, and Kauschke (2012) reported that children with AS produced shorter narratives than typically developing children. Shriberg et al. (2001) reported that the speech rate of individuals with HFA and AS is within the normal range of the widely accepted 4–6 syllables per second. Still, utterances of individuals with HFA were significantly more coded as too slow or slow than those of typically developing individuals.

As some findings indicated the pragmatic role of disfluencies in spontaneous speech (Bortfeld et al., 2001; Clark, 1994; Clark & Fox Tree, 2002; Fox Tree, 2001; Levelt, 1989; Shriberg, 2001), the question arose whether their distribution would be different in speakers with ASD. Lake, Humphreys, and Cardy (2011) investigated the production of disfluencies in speakers with high-functioning autism as compared to the typically developing controls. Results revealed that participants with ASD use more silent pauses and disfluent repetitions than the control group, but fewer revisions and filled pauses such as *um* and *uh*. Moreover, ASD participants used the same rate of silent pauses as the controls used filled pauses, which authors found as an indication that ASD participants used silent pauses in the same positions within the utterance where the controls use *uh* and *um*. Bearing in mind the possible pragmatic function of filled pauses in terms that they signal the speaker is not finished with his conversational turn, which is not the case with silent pauses, the authors interpreted their results as an indication that participants with ASD are less speaker-oriented than the controls, which would be due to the ToM deficit. Irvine, Eigsti, and Fein (2016) explored the function of *uh* and *um* in spontaneous speech in three different groups of participants. The first group consisted of participants with a history of ASD in childhood who lost their ASD diagnosis in the later age. The second group consisted of participants with high-functioning autism (ASD group), and the third group consisted of typically developing individuals. The results showed that participants with ASD produced significantly fewer *ums* than the other two groups, while there was no significant difference in the rate of *uh* between the groups. The authors interpreted their results as the evidence that *uh* serves as a pragmatic cue to the listener and it is more listener-oriented, whereas *um* is more inward-oriented, that is, it does not have communicative role in conversation. Given that individuals with ASD have specific pragmatic deficits, these results were expected, that is, individuals with ASD are less aware of various communicative cues, including *ums*.

3. AIM AND RESEARCH HYPOTHESES

Given the evidence for pragmatic function of disfluencies, the aim of this study was to examine the distribution of the main types of typical disfluencies in high-functioning children with ASD during spontaneous speech in Croatian, and to compare the results with those of typically developing individuals. In order to elicit the spontaneous speech, the participant's task was to retell a story about the Pink Panther, after watching the cartoon. The main research questions were: (1) Do individuals with ASD use the same number of disfluencies as their typically developing peers? (2) Do ASD and typically developing (TD) groups use the same disfluency categories in spontaneous speech and with equal frequency? (3) Do these two groups of participants differ in speed fluency? Finally, the aim was to consider the obtained results in the light of pragmatic and cognitive perspective on disfluencies.

In this study, the following hypotheses are set:

H1: Speech rate in the ASD group will be significantly lower than in the typically developing participants. Assuming that individuals with ASD are less listener-oriented (Irvine et al., 2016), they will be less engaged in maintaining speech flow using spoken disfluencies, which will result in longer silent pauses. As silent pauses are included in the speaking time, it should be expected that this will have repercussions on the duration of the speech interval. Furthermore, this hypothesis is in line with the study of Shriberg et al. (2001) which reported that speech rate in individuals with HFA was more often coded as too slow or slow.

H2: Participants with ASD will produce a higher number of disfluent silent pauses than the controls. This hypothesis is based on the previous research (Lake et al., 2011) on disfluencies in ASD, which indicates that participants with ASD use silent pauses instead of filled ones at the same place in discourse, probably due to the lack of pragmatic skills.

H3: Participants with ASD will produce lower number of filled disfluencies than the controls. As Lake et al. (2011) have shown, participants with AS compensate filled pauses for silent ones, which is interpreted as the consequence of deficiency in pragmatic skills.

H4: The total number of disfluencies will not be different between participants with ASD and the controls. As some studies have shown (Colle et al., 2008), general narrative abilities in individuals with ASD do not differ from the typical population. The only difference is in those expressions that have additional pragmatical function. As disfluencies are an inevitable part of spontaneous speech, it should be expected that participants with ASD will be as disfluent as the controls.

H5: Participants with ASD will use the same types of disfluencies as the controls. Given that all types of typical disfluencies are manifestations of inner psycholinguistic processes during speech production, all of them are expected in spontaneous narratives of both groups.

4. METHODS

4.1 Participants

The group of high-functioning children with ASD (experimental group) consisted of nine boys with AS (N = 8) or HFA (N = 1) and one girl with AS[1]. All of them had a clinical diagnosis of autism assessed by a team of clinical experts (therapist, speech therapist and psychologist) and they were diagnosed by a pediatric neurologist. Their mean age was 12.7 years (SD = 1.84); range 8–16 years. They were all native speakers of Croatian. The participants were recruited through Croatian autism associations and by personal contacts of the authors. They were from Zagreb. Their parents did not report any other cognitive difficulties of their children. Furthermore, the children included in the study had no reported hearing or speech difficulties (e.g., stuttering) or any other neuropsychiatric diseases or learning disabilities. No stuttering disfluencies were recorded in the speech of participants, which was one of the criteria for selecting participants. The participants of the control group consisted of typically developing (TD) nine boys and one girl who were matched in terms of age (mean age was 12.7 years; SD = 1.84; range 8–14 years), sex, native language, and all other controlled variables with the target group (children with ASD). They had no neurodevelopmental disorders. They had no reported learning disabilities.

4.2 Procedure

The experiment was approved by the Ethics Committee of the Department of Phonetics of the University of Zagreb. The TD children were recruited from one

[1] In the current paper, the terms AS and HFA were used interchangeably, which is in line with some other studies (e.g., Colle et al., 2008; Freitag et al., 2007; Lee et al., 2023; Mazzone et al., 2012).

elementary school in Zagreb. The recruiting procedure was as follows: the school pedagogue gave parents the introductory letter with the described procedure and goals of the experiment. If parents agreed for their child to participate in the experiment, they contacted the researcher directly. Prior to the experiment, the parents completed the questionnaire to provide the researchers with the information on their children's health status, psychosocial and cognitive development. Written consent was obtained from the parents. All participants were asked if they were willing to participate in the experiment freely and they were informed that they can withdraw from it any time without any consequences. Participants did not receive any kind of compensation for their participation in the experiment. The testing took place either at the participant's home or in the school of the TD group, in a quiet room. Participants were asked to watch a five-minute Pink Panther animated movie, and after that to retell the story with as many details as possible in Croatian language. In order to ensure 1 minute of speech per participant, retelling was boosted by follow-up questions of the experimenter, if needed. These questions were "open-ended", i.e., they encouraged a broad answer, for example: "And what happened next?", "Can you explain it in more details?". Their speech was audio recorded using Zoom H2 audio recorder (with sampling frequency of 44,100 Hz, and 16-bit resolution).

4.3 Data analysis

To ensure equal duration of analyzed speech, only the first minute of speech per speaker was included in the analysis. If the duration of the continued speech was less than one minute, the next part of the discourse was added to the speech material (obtained by follow-up questions), until around one minute of speech per speaker was obtained. Table 1 shows the duration of audio clips for each participant. T-test showed that the duration of analyzed clips between two groups is not statistically significant: $t(18) = 0.0586$, $p = 0.9539$.

Table 1. Duration of analyzed audio clips for each participant, along with average duration (Mean) and standard deviation (SD)

Tablica 1. Trajanje analiziranih zvučnih odsječaka po ispitaniku uz prosječno trajanje (prosjek) i standardnu devijaciju (SD)

TD		ASD	
Participant / Sudionik	Duration (s) / Trajanje (s)	Participant / Sudionik	Duration (s) / Trajanje (s)
TD1	58.61	ASD1	58.95
TD2	64.23	ASD2	61.60
TD3	60.38	ASD3	65.33
TD4	60.40	ASD4	60.99
TD5	61.66	ASD5	61.98
TD6	60.30	ASD6	58.22
TD7	62.66	ASD7	62.01
TD8	60.00	ASD8	62.05
TD9	60.72	ASD9	61.04
TD10	63.14	ASD10	59.43
Mean / Prosjek	61.21	Mean / Prosjek	61.16
SD	1.61	SD	1.91

The recorded speech was transcribed by two trained phoneticians. The transcription included all fillers and disfluencies. Silent pauses longer than 120 ms were annotated in Praat (Boersma & Weenink, 2022). The frequency of disfluencies was expressed as the number of disfluencies per 100 words. The speech rate was calculated as the number of uttered syllables in one second. Statistical analysis included chi-square test (since the data on frequency of disfluencies did not meet the criteria for parametric statistics) and t-test for calculating the speech rate.

In order to test some interpretations that filled pauses are more speaker-oriented than the silent ones (Lake et al., 2011), as well as to test H3, disfluencies were divided into two major categories: (1) silent pauses (SP) and (2) filled disfluencies (FD). There is a lot of disagreement in the literature about the threshold for silent pauses. Goldman-Eisler (1958, 1972) proposed 250 ms as the threshold for a disfluent silent pause, arguing that this is the cut-off point for excluding all silent periods during speech caused by occlusion as the part of some articulatory processes. Zellner (1994) refers to this type of silent periods as intra-segmental pauses. The threshold of 250 ms became broadly accepted in silent pause research (e.g., de Jong, 2016; Kormos & Dénes, 2004; Kovač & Vickov, 2018, 2019). However, Hieke, Kowal, and O'Connell (1983) challenged this threshold, arguing that pauses of 130–250 ms may also have a psychological function, and that their exclusion from the pause analysis is therefore unjustified. Rosen (2005) and Rosen at al. (2010) also note that the exclusion of pauses shorter than 250 ms could have repercussions for the reliability of the results regarding pause distribution. Furthermore, not all silent pauses are disfluent. Many of them are also constitutive elements of speech prosody, therefore it is difficult to distinguish between disfluent and fluent (i.e., with prosodic function) silent pauses exclusively on their duration. In other words, the same duration of the silent pause in one position within the utterance could be annotated as a prosodic component, and in another one as a disfluency. For example, Škarić (1991) argues that pauses between two intonational units are shorter than pauses between two sentences. In order to avoid silent breakdowns in speech continuity whose origin is of the articulatory nature or silent pauses with a prosodic function, but still to capture shorter silent pauses caused by speech planning troubles (i.e., disfluencies), we decided to use a 120 ms threshold. However, silent pauses were excluded from the analysis if both authors assessed that they had a prosodic function in terms of boundary markers between larger speech units. In other words, if pauses were between two utterances or phrases, they were interpreted as boundary markers.

Filled disfluencies included all types of disfluencies which were vocalized and audible to the listener. They were further divided into six categories: filled pauses (FP), prolongations of the last vowel (Pr), interruptions (I), repetitions (Rp), revisions (Rv) and filler words (FW). Filled pauses included fillers like nonphonemic vowel [ə], or nasalized sounds like [hm] and [um] which are common fillers in Croatian. Another way to avoid breakdowns in speech continuity is by prolongating some speech segments. In Croatian, this is usually the last vowel of the word. Interruptions are sudden stops of utterance production which may or may not be followed by revision (Levelt, 1989). Revisions are kind of self-repairs in which speaker abandons the current utterance and starts a new one. Repetitions are another kind of self-repairs which consist of repeating a previous syntagma, word, morpheme, syllable, or segment (Levelt, 1989). Filler words are lexicalized forms of fillers which consists of frequent use of the same word without any semantic or syntactic need for it (Badurina & Matešić, 2013). The most common filler words in Croatian media are: *ono* (eng. *that*), *pa* (eng. *so*), *ovaj* (eng. *this*), *zapravo* (eng. *actually*), *dakle* (eng. *therefore*) and *znači* (eng. *it means*) (Golub & Vidović Zorić, 2022). Examples of all dysfluency types analyzed in this paper are given in Table 2.

Table 2. Examples of analyzed dysfluency types with English translation. Disfluencies are underlined in each example.

Tablica 2. Primjeri analiziranih vrsta disfluentnosti uz prijevod na engleski jezik. Disfluentnosti su podcrtane u svakome primjeru.

Dysfluency type / Vrsta disfluentnosti	Example in Croatian / Primjer na hrvatskome jeziku	Translation in English / Prijevod na engleski jezik
Silent pauses > 120 ms / Bezvučna stanka > 120 ms	Pink Panther // je vidio…	The Pink Panther // saw…
Filled pauses (ə, *hm*) / Zvučna stanka (ə, *hm*)	…ə nije uspio u tome.	…ə he didn't make it.
Prolongations at the last vowel / Duljenja posljednjega vokala	Sjećam se da je bilooo…	I remember it waaas…
Interruptions / Prekidi	Onda se autom# auto se…	Then by car# car was…
Repetitions / Ponavljanja	Pink Panther je htio htio biti superjunak.	Pink Panther wanted wanted to be a superhero.
Revisions / Revizije	Pokuš#… počela ga je mlatit.	She tri#… started beating him.
Filler words / Poštapalice	Znači, prvo je Pink Panther video…	It means, Pink Panther first saw…

Speed fluency is measured by three variables: the duration of silent pauses, the duration of sounding interval, and speech rate. The duration of silent interval and sounding interval was calculated in Praat. In order to compare the sounding and silent ratio in analyzed speech, this measure included all silent pauses, not just disfluent ones. Speech rate was measured as the number of pronounced syllables per second. Other variables included in the analysis were: number of all disfluencies per 100 words (normalized number of disfluencies), number of disfluent silent pauses and filled disfluencies per 100 words, number of different types of filled disfluencies per 100 words (filled pauses, prolongations of the last vowel, interruptions, repetitions, revisions, and filler words).

5. RESULTS

Both groups of participants managed to perform the task. However, in some cases, retelling was boosted through follow-up questions. The ASD group produced 1,250 words in total, and the TD group produced 1,543 words. The maximum number of produced words per participant in the ASD group was 201, and in the TD group it was 334. The minimum number of produced words in the ASD group was 86, and in the TD group it was 103. The average number of produced syllables in the ASD group was 196.90 (SD = 49.77), and in the TD group it was 247.80 (SD = 48.08). For the ASD group, the highest number of produced syllables per participant was 287, and the lowest number was 142. For the control group, the highest number of produced syllables per participant was 334, and the lowest number was 171.

5.1 Speed fluency

5.1.1 The duration of silent intervals vs. sounding intervals

For the ASD group, the duration of silent intervals vs. duration of the sounding intervals was 24% vs. 76% in the total duration of speech. For the TD group this proportion was 19% (silent intervals) vs. 81% (sounding intervals). These results suggest that TD participants are keener to maintain the communication using continued speech sound, while the ASD group is more prone to break that flow with silent pauses.

5.1.2 Duration of silent intervals

Further analysis of silent intervals included their duration in the speech corpus. The average duration of SP in the ASD group was higher than in the TD group (Figure 1). T-test showed that the difference between the two groups was significant: $t(396) = 3.9051$, $p = 0.00$.

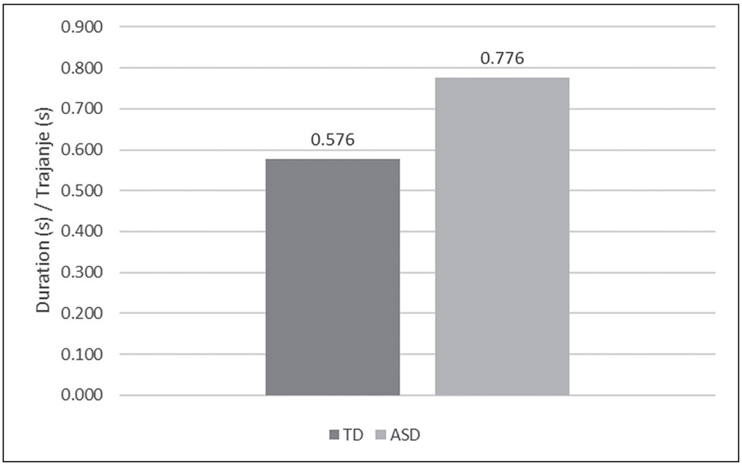

Figure 1. Average duration of silent intervals for TD (SD = 0.354) and ASD (SD = 0.632) group in seconds (s)

Slika 1. Prosječno trajanje bezvučnoga intervala za TD (SD = 0,354) i ASD (SD = 0,632) grupu u sekundama (s)

The maximum duration of silent intervals per speaker in the ASD group was 4.584 s, and the minimum was 0.197 s. In the TD group, the maximum (2.648 s) and the minimum (0.121 s) duration was lower as compared to the ASD group. These results additionally indicate that ASD participants are less prone to maintain the continuation of speech flow than the control group.

5.1.3 Duration of sounding intervals

Maximum duration of sounding intervals per speaker in the ASD group was 9.282 s, and the minimum was 0.012 s. In the TD group, the maximum duration was 8.144 s, and the minimum was 0.08 s. The average duration of a sounding interval between two silent intervals was slightly higher for the TD than for the ASD group (Figure 2), although the difference between the two groups was not statistically significant: $t(401) = 0.278$, $p = 0.78$.

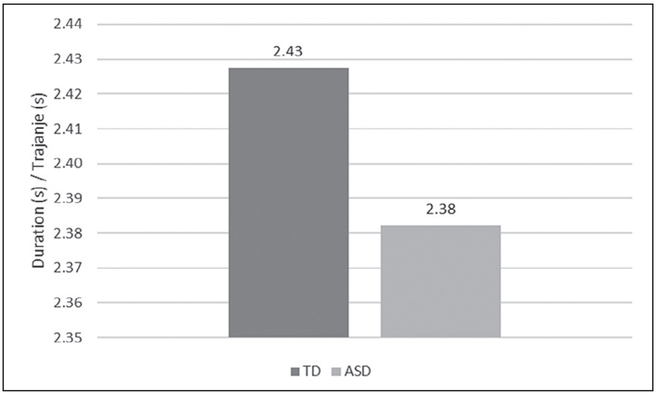

Figure 2. Average duration of sounding intervals between two SPs for the TD (SD = 1.649) and the ASD (SD = 1.624) group in seconds (s)

Slika 2. Prosječno trajanje zvučnih intervala za TD (SD = 1,649) i ASD (SD = 1,624) grupu u sekundama (s)

These results indicate that as far as speed fluency is concerned, the main difference between the two groups arises from the way silent intervals are used and from their distribution. On the other hand, it appears that both groups temporally organize a sounding time in a similar way.

5.1.4 Speech rate

The maximum speech rate in the TD group was 5.70, while the minimum speech rate was 2.60. For the TD group, the maximum speech rate was 4.70, while the minimum was 2.17. The average speech rate for both groups is shown in Figure 3.

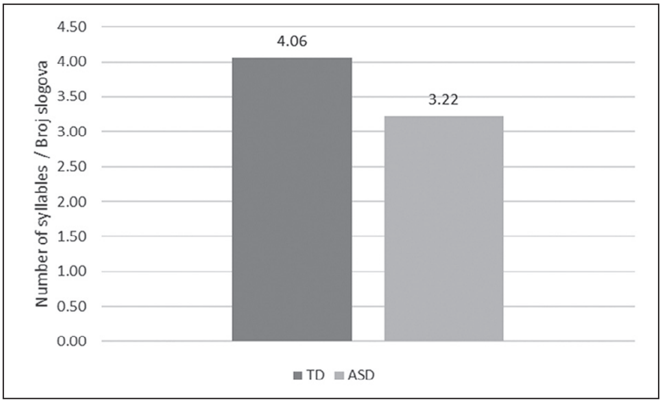

Figure 3. Average speech rate for the TD (SD = 0.82) and the ASD (SD = 0.82) group

Slika 3. Prosječna govorna brzina za TD (SD = 0,82) i ASD (SD = 0,82) grupu

T-test showed that the difference between the two groups is significant ($t(18) = 2.1549$, $p = 0.0450$) at the level of 0.05. It should be noted that the average duration of silent intervals was longer in the ASD than in the TD group. As silent intervals are a constitutive part of the total speaking time, these results are expected. They are also consistent with H1.

5.2 The total number of disfluencies in the ASD and TD group

In order to avoid the influence of discrepancy in the number of produced words between the groups, the number of disfluencies was normalized using the number of disfluencies per 100 words as a comparable measure between the groups. Participants with ASD produced 32.50 disfluencies per 100 words in total, whereas TD participants produced 32.14 disfluencies per 100 words. The difference in the number of produced disfluencies between groups is not statistically significant ($\chi^2 = 0.019$; df = 1; $p = 0.88$). This result indicates that H4 is correct. Furthermore, all analyzed types of disfluencies (see Table 2) were produced by participants of both groups, which is in line with H5.

5.2.1 Disfluent silent pauses and filled disfluencies

There is high interspeaker variability in the normalized number (per 100 words) of disfluent silent pauses and filled disfluencies for both groups. The range in the TD group is between 2.20 and 14.88 for SPs, and between 12.82 and 51.35 for SDs. The range in the ASD group is between 6.78 and 21.51 for SPs, and between 8.23 and 39.53 for SDs. Figure 4 presents a normalized number of disfluent silent pauses and filled disfluencies in both groups.

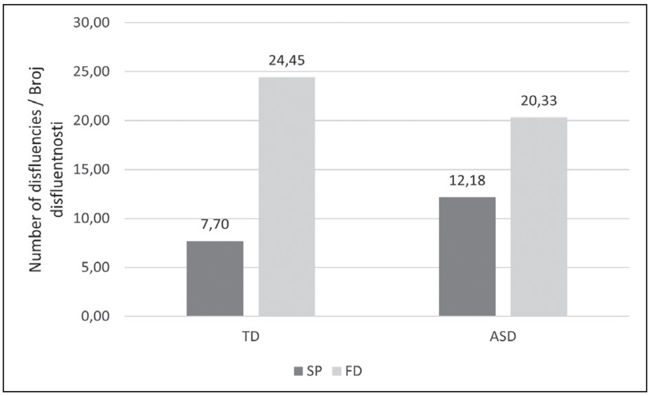

Figure 4. Number of silent pauses (SP) and filled disfluencies (FD) per 100 words in the TD and the ASD group

Slika 4. Broj bezvučnih stanki i zvučnih disfluentnosti na 100 riječi u TD i ASD grupi

The difference in the number of SPs between the TD and the ASD group was found to be statistically significant ($\chi^2 = 13.135$; df = 1; p = 0.00), as well as in the number of SDs ($\chi^2 = 4.703$; df = 1; p = 0.03), at the level of 0.05. These results are consistent with H2 and H3.

5.2.2 Analysis of filled disfluencies

Figure 5 presents the normalized number of different types of filled disfluencies in both groups. It can be seen that the TD group produced a higher number of disfluencies in each analyzed category of spoken disfluencies.

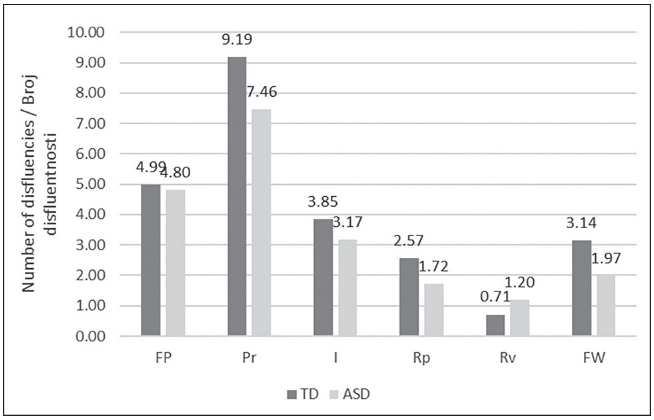

Figure 5. Number of different types of filled disfluencies per 100 words in the TD and the ASD group
FP – filled pauses; Pr – prolongations of the last vowel; I – interruptions; Rp – repetitions; Rv – revisions; FW – filler words

Slika 5. Broj pojedinih vrsta zvučnih disfluentosti na 100 riječi u TD i ASD grupi
FP – zvučne stanke; Pr – duljenja posljednjega vokala; I – prekidi; Rp – ponavljanja; Rv – revizije; FW – poštapalice

However, analysis did not reveal a significant group difference for any of the categories: FP ($\chi^2 = 0.045$; df = 1; p = 0.83), Pr ($\chi^2 = 2.260$; df = 1; p = 0.13); I ($\chi^2 = 0.824$; df = 1; p = 0.36); Rp ($\chi^2 = 1.510$; df = 1; p = 0.219); Rv ($\chi^2 = 2.116$; df = 1; p = 0.14) and FW ($\chi^2 = 3.297$; df = 1; p = 0.06). These results are also in line with H5, according to which participants with ASD produce the same types of disfluencies as the controls, although not at the same rate.

6. DISCUSSION

This paper examined the number and types of disfluencies in individuals with ASD during the spontaneous speech in Croatian and measured the speech rate as an aspect of speed fluency. The results were compared with those for typical individuals. Our results showed that individuals with ASD used a similar total number of disfluencies as their typically developing peers. Moreover, they use exactly the same categories as the controls. Thus, H4 and H5 are confirmed. It should be remembered that disfluencies are primarily viewed as manifestations of troubles at the planning level (Clark & Fox Tree, 2002; Corley & Stewart, 2008; Levelt, 1989). Accordingly, these results indicate that cognitive processes underlying speech production and narrative organization in individuals with ASD do not differ from the same processes in individuals with typical speech. Still, the distribution of those categories of disfluencies is somewhat different between the groups. It is worth noting that some studies (Colle et al., 2008; Losh & Capps, 2003; Seung, 2007) found that participants with ASD usually do not have a problem with telling a story: they are able to follow the main plot, do not differ in quantity of information or in the grammatical complexity of the sentences from the controls. The main difference is the narration arising from these elements that require pragmatic understanding, such as anaphoric pronouns, or temporal and referential expressions (Colle et al., 2008). The primary aim of the current study was to analyze the obtained results in the light of potential pragmatic function of disfluencies. Previous studies had already shown that disfluencies could have a pragmatic function as well (Bortfeld et al., 2001; Clark, 1994; Clark & Fox Tree, 2002; Fox Tree, 2001; Levelt, 1989). Due to the deficit in pragmatic skills, some researchers predicted different disfluency patterns in individuals with ASD (Irvine et al., 2016; Lake et al., 2011). As we presupposed by H2 and H3, our results showed that participants with ASD produced significantly more disfluent silent pauses and significantly fewer spoken disfluencies as compared to the typically developing group, which is in line with the results obtained by Lake et al. (2011). As Lake et al. (2011) suggested, these results could be interpreted in terms of the pragmatic view on disfluencies. More specifically, if a speaker has difficulties in speech planning, but at the same time does not want to lose the floor in conversation, they usually maintain the speech using filled pauses and other kinds of spoken disfluencies. In this case, spoken disfluencies are seen as pragmatic cues in conversation. Individuals with ASD appear to less likely maintain speech flow in this way, which could be attributed to a lack of pragmatic skills in this population. Lake at al. (2011) concluded that participants with ASD are less listener-oriented, that is, they care less about the listener's point of view in the conversational act and potential misunderstandings, which is argued by the ToM deficit in ASD. In line with this explanation is also our additional finding that the speech rate in the ASD group was significantly slower than in the controls, which confirmed H1. This is probably due to the significantly longer duration of silent pauses in the total speech time in the ASD group than in the controls. This finding additionally suggests that participants with ASD are less aware of the communicative role of spoken disfluencies, or, in other words, they are less able to recognize a possible unintended message which could be signaled by a silent pause. It appears that individuals with AS/HF are less concerned about the possible negative effect of longer silent intervals during the conversational turn on the listener's perception of received speech or understanding of a message. As regards the speech rate, Shriberg et al. (2001) reported similar results: utterances of HFA participants were more often perceived as too slow or slow than those of the controls. It is indicative that no significant difference was found in the average duration of the sounding part between the two groups. That is, the difference between the two groups arises only from the silent intervals. Still, it should be noted that there was a large variability in this variable in both groups which could have masked potential group differences. Our results showed that controls produced a higher number of filled disfluencies in five categories, that is, more filled pauses, prolongations, interruptions, repetitions, and filler words, although the difference between the groups is not statistically significant. As regards the number of filled pauses, this finding is not completely in line with the study from Irvine et al. (2016), which reported that ASD participants produced fewer *ums* and *uhs*, although for the latter the difference was not significant. The authors concluded that *um* serves as a pragmatic marker, while *uh* does not have this function in the conversation. Given that in the current study all filled pauses were considered as a single category, further research is needed to investigate whether different types of filled pauses in Croatian, for example *hm*, *ə*, *m*, differ from each other in terms of their role in communication. Contrary to the finding of Lake et al. (2011), participants with ASD produced more revisions than the controls. However, the difference was not significant in any of the analyzed variables. Nevertheless, it would be worth to investigate these opposite results regarding revisions more thoroughly in future studies. Additionally, it should be noted that results again show high variability between speakers in both groups. As some other studies also had similar results (Bortfeld et al., 2001; Clark & Fox Tree, 2002; Golub & Vidović Zorić, 2022; Shriberg, 2001), these results are expected. However, they should be considered in the interpretation of the results, since they can potentially have impact on group comparisons.

This study has certain limitations. Firstly, the sample size is not too large, thus it would be useful to expand research to more participants, especially for the experimental group. Secondly, the elicited speech sample is quasi-spontaneous, since participants were not completely free to choose the topic of speaking, and lexical items were suggested by the cartoon's story as well. Furthermore, the study considers the disfluencies in only one type of discourse – narratives – and does not take into account other types, such as dialogue or argumentation. As Irvine et al. (2016) remarked, telling a story is a type of monologue, thus, this kind of discourse could diminish a potential group difference, since pragmatic ability less comes to the fore in a monologue than in a dialogue. It should be also noted that Finlayson & Corley (2012) expressed doubts about the pragmatic role of disfluencies, as the results of their research showed no difference in the distribution of disfluencies between a monologue and a dialogue. That is, if certain types of disfluencies are indeed more listener-oriented, one should expect their higher number in a dialogue, concluded the authors. Therefore, verifying their results in the Croatian speech corpus would provide a better insight into potential functions of disfluencies.

The current findings could have implications for a better understanding of inner processes in speech production in both typical and atypical individuals, specifically in ASD. Furthermore, they could contribute to the understanding of the cognitive organization underlying narration in ASD, as well as indicate similarities and differences in the usage of disfluencies as pragmatic markers between the ASD population and the controls in communication. Considering that less research has been conducted on the potential pragmatic function of disfluencies in languages other than English, this study provides new evidence on this issue, that is, it sought to answer whether the claims about some types of disfluencies that proved true for English were also true for Croatian. Finally, this new evidence on disfluencies as pragmatic markers in Croatian could have a beneficial effect on speech and social therapy of individuals with ASD.

To conclude, the current study showed that, as regards the total number of disfluencies, individuals with ASD did not produce significantly more disfluencies than typically developing individuals. Furthermore, both groups produced the same type of disfluencies which suggests that disfluencies are indeed a manifestation of universal underlying processes of speech production. However, some types of disfluencies are differently distributed between the groups, that is, ASD participants produced significantly more disfluent silent pauses and significantly fewer spoken disfluencies. As previous studies indicated a pragmatic function of some types of disfluencies, more specifically, silent pauses, current results could be attributed to the deficiency of pragmatic ability in ASD individuals. The results are in line with some similar studies on this issue. Finally, speech rate was significantly slower in the ASD group, which could be explained as potential evidence that individuals with ASD are less engaged in maintaining speech flow, resulting in longer silent pauses. Bearing in mind the potential limitations of this research, some aspects of the pragmatic role of disfluencies, i.e., their distribution in different types of discourses, should be investigated in more detail in future studies.

REFERENCES

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders: DSM 5*. Washington, DC: Bookpoint US.

Badurina, L., & Matešić, M. (2013). Poštapalice pod povećalom: o jednoj posebnoj vrsti diskursnih oznaka. In A. Turbić-Hadžagić (Ed.), *Zbornik radova. Tuzla. Grad na zrnu soli* (pp. 9–24). Tuzla: JU Zavod za zaštitu i korištenje kulturno-historijskog i prirodnog nasljeđa Tuzlanskog kantona.

Baron-Cohen, S. (1995). *Mindblindness: An essay on autism and theory of mind*. MIT Press.

Baron-Cohen, S. (2000). Theory of mind and autism: A review. In L. M. Glidden (Ed.), *International Review of Research in Mental Retardation*, Vol. 23 (pp. 169–184). Academic Press. https://doi.org/10.1016/S0074-7750(00)80010-5

Bishop, D. V. (2003). Autism and specific language impairment: Categorical distinction or continuum? *Novartis Foundation symposium*, *251*, 213–297.

Boersma, P., & Weenink, D. (2022). Praat: Doing phonetics by computer [Computer program], Version 6.2.06. Retrieved August, 27, 2023, from https://www.praat.org

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.

Brennan, S. E., & Schober, M. F. (2001). How listeners compensate for disfluencies in spontaneous speech. *Journal of Memory and Language*, *44*(2), 274–296. https://doi.org/10.1006/jmla.2000.2753

Campisi, L., Nazish Imran, N., Nazeer, A., Skokauskas, N., & Waqar Azeem, M. (2018). Autism spectrum disorder. *British Medical Bulletin*, *127*(1), 91–100.

Clark, H. H. (1994). Managing problems in speaking. *Speech Communication*, *15*(3–4), 243–250.

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

Colle, L., Baron-Cohen, S., Wheelwright, S., & van der Lely, H. K. (2008). Narrative discourse in adults with high-functioning autism or Asperger syndrome. *Journal of Autism and Developmental Disorders*, *38*(1), 28–40.

Corley, M., & Stewart, O. W. (2008). Hesitation disfluencies in spontaneous speech: The meaning of um. *Language and Linguistics Compass*, *2*(4), 589–602.

Dai, H., He, X., Chen, L., & Yin, C. (2022). Language impairments in children with developmental language disorder and children with high-functioning autism plus language impairment: Evidence from Chinese negative sentences. *Frontiers in Psychology*, *13*, 926897.

de Jong, N. H. (2016). Predicting pauses in L1 and L2 speech: The effects of utterance boundaries and word frequency. *International Review of Applied Linguistics in Language Teaching*, *54*(2), 113–132.

Finlayson, I. R., & Corley, M. (2012). Disfluency in dialogue: An intentional signal from the speaker? *Psychonomic Bulletin & Review*, *19*(5), 921–928.

Fox Tree, J. E. (1995). The effects of false starts and repetitions on the processing of subsequent words in spontaneous speech. *Journal of Memory and Language*, *34*(6), 709–738.

Fox Tree, J. E. (2001). Listeners' uses of *um* and *uh* in speech comprehension. *Memory & Cognition*, *29*(2), 320–326.

Freitag, C. M., Kleser, C., Schneider, M., & von Gontard, A. (2007). Quantitative assessment of neuromotor function in adolescents with high functioning autism and Asperger syndrome. *Journal of Autism and Developmental Disorders*, *37*(5), 948–959.

Gibson, J., Adams, C., Lockton, E., & Green, J. (2013). Social communication disorder outside autism? A diagnostic classification approach to delineating pragmatic language impairment, high functioning autism and specific language impairment. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *54*(11), 1186–1197.

Goldman-Eisler, F. (1958). The predictability of words in context and the length of pauses in speech. *Language and Speech*, *1*(3), 226–231.

Goldman-Eisler, F. (1972). Pauses, clauses, sentences. *Language and Speech*, *15*(2), 103–113.

Golub, G., & Vidović Zorić, A. (2022). Poštapalice u medijskome govoru. *Fluminensia*, *34*(2), 367–396.

Hieke, A. E., Kowal, S., & O'Connell, D. C. (1983). The trouble with "articulatory" pauses. *Language and Speech*, *26*(3), 203–214.

Irvine, C. A., Eigsti, I. M., & Fein, D. A. (2016). *Uh*, *um*, and autism: Filler disfluencies as pragmatic markers in adolescents with optimal outcomes from autism spectrum disorder. *Journal of Autism and Developmental Disorders*, *46*(3), 1061–1070.

Kormos, J., & Dénes, M. (2004). Exploring measures and perceptions of fluency in the speech of second language learners. *System*, *32*(2), 145–164.

Kovač, M. M., & Vickov, G. (2018). The impact of immediate task repetition on breakdown fluency. *Govor*, *35*(2), 139–160.

Kovač, M. M., & Vickov, G. (2019). The influence of pre-task planning on speed and breakdown fluency. *Vestnik za tuje jezike*, *11*(1), 87–102.

Lake, J. K., Humphreys, K. R., & Cardy, S. (2011). Listener vs. speaker-oriented aspects of speech: Studying the disfluencies of individuals with autism spectrum disorders. *Psychonomic Bulletin & Review*, *18*(1), 135–140.

Lee, T. M. C., Liang, L., Hou, W. K., Tse, A. H. Y., & Chan, C. C. H. (2023). The Chinese "Reading the Mind in the Eyes" Test: A study with normal adults, and adults with Asperger syndrome/high-functioning autism. *Asian Journal of Psychiatry*, *89*, 103785. Advance online publication.

Lennon, P. (2000). The lexical element on spoken second language fluency. In H. Riggenbach (Ed.), *Perspectives on fluency* (pp. 25–42). Ann Arbor: University of Michigan Press.

Levelt, W. J. (1989). *Speaking: From intention to articulation*. Cambridge, Massachusetts. London: A Bradford Book. The MIT Press.

Losh, M., & Capps, L. (2003). Narrative ability in high-functioning children with autism or Asperger's syndrome. *Journal of Autism and Developmental Disorders*, *33*(3), 239–251.

Mazzone, L., Ruta, L., & Reale, L. (2012). Psychiatric comorbidities in Asperger syndrome and high functioning autism: Diagnostic challenges. *Annals of General Psychiatry*, *11*(1), 16.

Oviatt, S. (1995). Predicting spoken disfluencies during human-computer interaction. *Computer Speech and Language*, *9*(1), 19–35.

Postma, A., Kolk, H., & Povel, D.-J. (1990). Speech planning and execution in stutterers. *Journal of Fluency Disorders*, *15*(1), 49–59.

Rosen, K. (2005). Analysis of speech segment duration with the lognormal distribution: A basis for unification and comparison. *Journal of Phonetics*, *33*, 411–426.

Rosen, K., Murdoch, B., Folker, J., Vogel, A., Cahill, L., Delatycki, M., & Corben, L. (2010). Automatic method of pause measurement for normal and dysarthric speech. *Clinical Linguistics & Phonetics*, *24*(2), 141–154.

Rumpf, A. L., Kamp-Becker, I., Becker, K., & Kauschke, C. (2012). Narrative competence and internal state language of children with Asperger syndrome and ADHD. *Research in Developmental Disabilities*, *33*(5), 1395–1407.

Schmidt, R. (1992). Psychological mechanisms underlying second language fluency. *Studies in Second Language Acquisition*, *14*(4), 357–385.

Segalowitz, N. (2003). Automaticity and second languages. In C. Doughty, & M. H. Long (Eds.), *The handbook of second language acquisition* (pp. 382–408). Malden, MA: Blackwell.

Seung, H. K. (2007). Linguistic characteristics of individuals with high functioning autism and Asperger syndrome. *Clinical Linguistics & Phonetics*, *21*(4), 247–259.

Shriberg, E. (2001). To ‘errrr’ is human: Ecology and acoustics of speech disfluencies. *Journal of the International Phonetic Association*, *31*(1), 153–169.

Shriberg, L. D., Paul, R., McSweeny, J. L., Klin, A. M., Cohen, D. J., & Volkmar, F. R. (2001). Speech and prosody characteristics of adolescents and adults with high-functioning autism and Asperger syndrome. *Journal of Speech, Language, and Hearing Research*, *44*(5), 1097–1115.

Škarić, I. (1991). Fonetika hrvatskoga književnog jezika. In S. Babić, D. Brozović, M. Moguš, S. Pavešić, I. Škarić, & S. Težak (Eds.), *Povijesni pregled, glasovi i oblici hrvatskoga književnog jezika* (pp. 61–377). Zagreb: HAZU – Globus.

Tager-Flusberg, H., & Sullivan, K. (1995). Attributing mental states to story characters: A comparison of narratives produced by autistic and mentally retarded individuals. *Applied Psycholinguistics*, *16*(3), 241–256.

Tavakoli, P., & Wright, C. (2020). *Second language speech fluency: From research to practice*. Cambridge: Cambridge University Press.

Zellner, B. (1994). Pauses and the temporal structure of speech. In E. Keller (Ed.), *Fundamentals of speech synthesis and speech recognition* (pp. 41–62). Chichester: John Wiley.

Ana Vidović Zorić
anvidovi@ffzg.unizg.hr
Filozofski fakultet Sveučilišta u Zagrebu
Hrvatska

Sara Blažeković
blazekovicsara11@gmail.com
Osijek, Hrvatska

Pragmatička funkcija govornih disfluentnosti u visokofunkcionalne djece s poremećajem iz spektra autizma

Sažetak

Disfluentnosti u spontanome govoru uobičajena su pojava. Najčešće se tumače kao manifestacija poteškoća na razini kognitivne obrade u procesu govorne proizvodnje. Ipak, rezultati pojedinih istraživanja ukazuju na to da neke vrste disfluentnosti mogu imati i pragmatičku funkciju u komunikaciji. U ovome istraživanju ispituje se raspodjela pojedinih vrsta disfluentnosti te govorna brzina u spontanome govoru deset visokofunkcionalnih ispitanika s dijagnozom poremećaja iz spektra autizma (Aspergerovim sindromom ili visokofunkcionalnim autizmom). Rezultati se uspoređuju s kontrolnom skupinom sastavljenom od ispitanika tipičnoga razvoja. Zadatak ispitanika bio je prepričati crtani film koji su prethodno pogledali. Govor je snimljen u zvučnome obliku, a zatim i transkribiran. Trajanje analiziranoga segmenta govora iznosilo je 60 sekundi. Rezultati su pokazali da nema značajne razlike u ukupnome broju disfluentnosti između dviju skupina. Nadalje, sudionici obiju skupina proizveli su iste vrste disfluentnosti. S obzirom na to da disfluentnosti odražavaju poteškoće u govornome planiranju, ovi rezultati pokazuju da su disfluentnosti zaista univerzalna pojava, kako u tipičnome tako i u atipičnome govoru. Istraživanje također pokazuje da su sudionici s poremećajem iz spektra autizma (ASD-om) proizveli znatno više disfluentnih bezvučnih stanki i značajno manje zvučnih disfluentnosti od kontrolne skupine. Ovakav rezultat mogao bi biti posljedica nedostatka pragmatičnih sposobnosti pojedinaca s ASD-om, kao što je već sugerirano u znanstvenoj literaturi. Drugim riječima, moguće je da se osobe s ASD-om manje zalažu u održavanju tijeka govora, što rezultira većim brojem bezvučnih stanki u odnosu na tipične govornike. Istraživanje je pokazalo i da je brzina govora u skupini ispitanika s ASD-om bila značajno manja nego u kontrolnoj skupini, što se također može objasniti utjecajem manjkavih pragmatičnih sposobnosti u osoba s ASD-om.

Ključne riječi: govorne disfluentnosti, Aspergerov sindrom, visokofunkcionalni autizam, poremećaj iz spektra autizma, pragmatika