CLINICAL LINGUISTICS AND PHONETICS

Speech disfluencies in typically developing Finnish children

Jansson-Verkasalo, E.[1], Lehtiö, I.[1], Silvén, M.[2] & Eggers, K.[1, 3]

[1] Department of Psychology and Speech-Language Pathology, University of Turku, Finland

[2] Department of Teacher Education, University of Turku, Finland

[3] Department of Speech-Language Pathology, Thomas More University College, Belgium

Correspondence

Eira Jansson-Verkasalo

email: eira.jansson-verkasalo@utu.fi

Abstract

We investigated the speech disfluencies of 54 typically fluent Finnish-speaking children: 14 children randomly selected from a longitudinal study (age levels 2, 3, and 4), and 40 children from a cross-sectional study (age levels 6, 7, 8, and 9). Speech samples, collected during a semistructured conversation, were analysed for disfluencies per 100 words and 100 syllables. No significant within-age effect was found for the total frequency of disfluencies or disfluency types among the 2- to 4-year-olds. Across the 6- to 9-year-olds, between-group differences were found for the total frequency and type of disfluencies. A clinically relevant notion was that the often-used criterion to distinguish normally fluent children from those who stutter, i.e., <3 stuttering like dysfluencies (SLD) per 100 syllables, was applicable in all age groups whereas the criterion < 3SLD per 100 words seems not to be valid in a language such as Finnish which has long words. Consequently, these preliminary results suggest that different guidelines are needed for defining normal disfluency from stuttering in different languages.

Introduction

Interruptions in the flow of speech, commonly referred to as disfluencies, are the most obvious feature of stuttering. Disfluencies, however, are also present in the speech of individuals who are not stuttering. The past 50 years of research have produced information on the information on the general features of the normal disfluencies mainly in the English language (e.g. (e.g., Ambrose & Yairi, 1999; Pellowski & Conture, 2002; Tumanova et al., 2014; Wexler & Mysak, 1982; Yairi & Ambrose, 2005)) but also in Dutch Boey et al., 2007; Eggers & Elen, 2018), French (Leclercq et al., 2018), German (Natke et al., 2006), and Spanish (Carlo & Watson, 2003; Watson & Anderson, 2001). A normative reference for fluent speech in children speaking Finnish, very different from the English language, does not exist.

The typological features of the ambient language may affect to the guidelines which are applicable when defining normal disfluency from stuttering. In Finnish there are only about 50 one-syllable words when particles are not included (Hakulinen, 2000). Mostly children's first words are two-syllable words (Savinainen-Makkonen, 2000). Kunnari (1998) compared the syllable number in the early words of 10 Finnish-speaking children with same-aged peers learning different types of languages. The comparison revealed that, at the 25-word-point, French, Japanese, Swedish, and especially Finnish children produce more disyllabic words than English-speaking children. In children under the age of 2 years, 3-syllbale or longer words are often truncated to 2-syllable words. By the age of 4 years most children are able to produce correctly even 4-5-syllable words (Kunnari et al., 2012). The Finnish language is a heavily inflected language compared to English. Instead of mainly using prepositions and word collocations, various suffixes are added to word stems (e.g., table/on the table = pöytä/pöydällä) which increases the word length (e.g., in my houses, too = talo+i+ssa+ni+kin). Even the earliest one-word utterances produced by Finnish infants can carry a morphological structure (e.g., I eat = syön) (Laalo, 2011). Thus, in this example, the word "syön" is one word but includes two morphemes.

The frequency and type of disfluencies are often used as indicators of stuttering and in instruments determining the severity of the disorder in clinical practice and research (e.g., Ambrose & Yairi, 1999; Conture, 2001; Cooper & Cooper, 1985; Curlee, 1993; Eichorn & Fabus, 2012; Riley, 1994). Johnson et al. (1959) introduced a disfluency classification alongside normative data regarding the speech disfluencies of English-speaking children who stutter (CWS) and children who not stutter (CWNS). The classification consists of eight categories: 1) sound and syllable repetition (e.g., m m my, thi thi this), 2) revision (e.g., Mom ate fixed dinner), 3) word repetition (e.g., and and), 4) incomplete phrase (e.g., I want Hey look at that), 5) phrase repetition (e.g., I want I want to go), 6) broken word (e.g., o…pen), 7) interjection or filled pauses (e.g., uhmm), and 8) prolongation (e.g., mmmy, cooookie). Later, the type of tense pause and dysrhythmic phonation (a combination of prolongation and broken word) were added to the classification (Williams et al., 1968). In 1981, Yairi made a distinction between mono- and multisyllabic word repetition. The earlier mentioned classifications resulted in the Illinois Disfluency Classification System (Ambrose & Yairi, 1999; Yairi & Ambrose, 2005) consisting of six types of disfluencies, namely 1) part word repetition, 2) monosyllabic word repetition, 3) dysrhythmic phonation, i.e., prolongation, block, and broken word, 4) interjection, 5) multisyllabic word repetition and phrase repetition and 6) revision or abandoned utterance. Tense pauses were not, however, included in the classification because of potential problems with identification of this category.

So far, only limited data are available about the frequency and types of disfluencies, and the developmental path of speech disfluencies in typically developing children. There is some evidence that the child's age may have an effect on the frequency and type of disfluencies (Eggers & Elen, 2018; Watson & Anderson, 2001; Wexler, 1982; Wexler & Mysak, 1982; Yairi & Clifton, 1972). As shown in the overview of prior cross-sectional research in Table 1, Leclercq et al. (2018) reported disfluency data for only one age level, whereas some cross-sectional studies have reported data for a single age range (Boey et al., 2007; Natke et al., 2006; Tumanova et al., 2014). Only a few studies reported findings for different age levels to define possible developmental differences across age groups (Ambrose & Yairi, 1999; Carlo & Watson, 2003; Haynes & Hood, 1977; Watson & Anderson, 2001; Wexler & Mysak, 1982; for details see Table 1). In these studies, the total frequency of disfluencies ranged between 2.6 to 14.6% per 100 words in 2- to 8-year-old children. No significant differences in the total frequency of disfluencies was found between 2-, 3-, and 4-year-old (Ambrose &Yairi, 1999), 2-, 4-, and 6-year-old (Wexler & Mysak, 1982) or 4-, 6-, and 8-year-old (Haynes & Hood, 1977) English-speaking children. In line with this, Carlo and Watson (2003) did not find any significant differences for the total frequency of disfluencies between 3- and 5-year-old Spanish-speaking children. Only Watson and Anderson (2001) reported a significantly higher total frequency of disfluencies in 3-year-old Spanish-speaking children when compared to 2-year-olds both per 100 syllables and per 100 words. Moreover, they found that the 3-year-olds exhibited significantly more revisions, phrase repetitions, and incomplete phrases than the 2-year-olds per 100 syllables. Similarly, Wexler and Mysak (1982) found differences in the type of disfluencies between the age groups they studied (Table 1). Furthermore, Haynes and Hood (1977) found that the frequency of interjections was higher among 8-year-old children than in 4-year-olds. In contrast, Ambrose and Yairi (1999) and Carlo and Watson (2003) did not find any differences in the type of disfluencies. To construct a reference base for typical early disfluency, more studies are needed, especially longitudinal ones and across older age groups (see e.g., summary by Yairi, 1997). To our knowledge there are no other longitudinal studies investigating the effect of age on disfluencies apart from Yairi's follow-up study (1982) in a group of 24- to 26-month-old and 29- to 33-month-old children.

An additional distinction is made between Stuttering-Like Disfluencies (SLD; i.e., part word repetitions, monosyllabic word repetitions, and dysrhythmic phonations) and Other Disfluencies (OD; i.e., interjections or filled pauses, multisyllabic word repetitions, phrase repetitions, as well as revisions and abandoned utterances; Ambrose & Yairi, 1999; Yairi & Ambrose, 2005). OD are suggested to be typical of fluent speech but also occur frequently in the speech of individuals who stutter (Tumanova et al., 2014; see also Adams, 1977). Instances of SLD are found to be more frequent in the speech samples of children diagnosed as stuttering compared to children considered fluent speakers (Ambrose & Yairi, 1999; Pellowski & Conture, 2001: Tumanova et al., 2014). Conture (2001) found that the ratio of SLD to OD was on average 66% for CWS and 28% for CWNS. The mean frequency of SLD in the speech of CWS is reported to vary from 9.6 to 16.9% of total disfluency, while in CWNS the frequencies vary from 1.2 to 3.2% (Yairi & Ambrose, 2005, p. 117). Conture (2001) as well as Ambrose and Yairi (2005) suggested a criterion of 3% SLD per total words to distinguish stuttering. Accordingly, the cut-off of 3% per 100 words was used in investigating English-speaking (Pellowski & Conture, 2002) and Dutch-speaking children (Boey et al., 2007). Boey et al. (2007) found that this criterion resulted in high sensitivity (0.95) and high specificity (0.98) in distinguishing CWS and CWNS. Natke et al. (2006) used a cut-off of 3% SLD per 100 syllables, and found that this criterion to be a powerful measure for the diagnosis of stuttering among German-speaking children. More than 93% of children could be classified correctly into their fluency group. This guideline of 3% seems to be applicable for various languages, including English, Dutch, Spanish and German (Ambrose & Yairi, 1999; Boey et al., 2007; Carlo & Watson, 2003; Natke et al., 2006; Pellowski & Conture, 2002; Tumanova et al., 2014). Recent findings, however, have shown that this guideline may not be used reliably in typically developing children speaking languages other than English (Leclercq et al., 2018). Moreover, a confounding factor is that in some studies a syllable-based metric (i.e. three stuttered disfluencies per 100 syllables) (e.g., Ambrose & Yairi, 1999; Carlo & Watson, 2003), while in others a word-based metric, was used (i.e. three stuttered disfluencies per 100 words) (e.g., Boey et al., 2007; Leclercq et al., 2017).

Normative data for the different types of disfluencies and their evolution throughout the child's development remain scarce until today (Eggers & Elen, 2018; Tumanova et al., 2014). Therefore, the aim of the current preliminary study was to explore age-related changes in disfluencies in typically developing Finnish-speaking children to provide a basis for the clinical need for a differential diagnosis of stuttering from normal disfluency. The second aim is to clarify whether the findings and norms of English or other languages can be applied to Finnish, a heavily inflected language with long words, and to reveal if the 3% SLD guideline used to define stuttering would be an appropriate means for the Finnish language. We used a longitudinal sample of 2- to 4-year-olds and a cross-sectional sample of 6- to 9-year-olds to investigate, whether age-related changes emerge when we assess 1) the total frequency of disfluencies, 2) the type of disfluencies, and 3) the amount of SLD and OD based on 100 words, on one, or 100 syllables, on the other hand. We hypothesised 1) no age-related changes or differences in the total frequency of disfluencies but expected 2) age-related changes or differences in the types of disfluencies. Moreover, we expected 3) no age-related changes or differences in the frequency of SLD, whereas we assumed that the frequency of OD would increase by age. Finally, we assumed that 4) the syllable-based metric is more sensitive than the word-based metric in counting SLD and OD because of the inflectional nature of the Finnish language with long words.

Methods

Participants

In this study, we investigated the disfluent speech of 54 monolingual, Finnish-speaking children. Fourteen typically developing children were randomly selected from a larger longitudinal study on children's development (age levels 2, 3, and 4 years), and 40 typically developing and fluent children randomly selected from a broad cross-sectional study (age levels 6, 7, 8, and 9 years) (see Table 2).

The longitudinal sample comprised first-born children of two-parent Finnish-speaking families. All families that gave birth to their child in January or February living in a middle-to-large southern Finnish city were chosen from the files of the Population Registration Centre and contacted by both mail and phone. Of 105 families that fulfilled all the criteria for the study, 66 families consented to participate in the follow-up (for more details, see Silvén et al., 2003). For the present study, 14 children (7 boys) followed up on a yearly basis from 2 to 4 years of age were randomly chosen from the longitudinal data. The youngest age represents the age level when normally developing children learn the basics of Finnish phonology and morphology (Kunnari, 2000; Laalo, 2011). Moreover, this is also the age range when developmental stuttering is most often diagnosed (Yairi & Ambrose, 2013). According to the parents, none of the children had any untreated visual, auditory, or neurological deficits during early childhood. The speech production of the selected children was assessed at all age levels from video recordings by the first author, a qualified speech-language pathologist, and was found to be typical for the age as indexed by the phonological development and word and sentence production. No indication of stuttering was encountered in any child. Moreover, the development of these children, including language development, was followed up in the health care system from new-borns to school age and was found to be normal.

Some evidence exists that the features of stuttering change remarkably from the early years to the age of 9 (see Bloodstein & Ratner, 2008). However, there are no studies on the Finnish language of the developmental changes in normal speech fluency in this age range. Therefore, a cross-sectional sample of older children, aged 6 (n = 10), 7 (n = 10), 8 (n = 10), and 9 (n = 10), was randomly selected from a broad study investigating central auditory processing and executive functions in fluently speaking children and CWS in northern Finland, Oulu region, and in south-western Finland, Turku region, during 2011–2013 with the permission of the Ethical Committee of Oulu University Hospital. After permissions for the broad study from the Health and Social services in both cities, fluently speaking children for the broad study were recruited through pre-primary and primary schools. An informative letter about the research project was sent to Kindergarten teachers of 6-year-old children and primary school teachers of 7-, 8-, and 9-year-old children asking for their permission to recruit typically developing, monolingual, healthy children for this study from their classrooms. After receiving the teachers' permission, informed consents together with a questionnaire about the child's development were distributed amongst the parents. Based on the completed questionnaires together with the signed written consent from the parents and permission from the child, the examiner selected those children who met the inclusion criteria and contacted both the teachers and the parents to schedule all the examinations for selected children. In addition to the parental report, the examiner checked that the development and hearing of these children, including language development, had been followed up in the health care system from new-borns to school age and was found to be normal. Only children who were monolingual Finnish-speaking children and had no known or questionnaire-reported neurological, psychological, developmental, learning, or hearing problems were included for further measurements. To verify the normal speech and language development of the participants, the first author assessed spontaneous speech samples between the child and the examiner of two video-recorded sessions. Speech fluency was assessed using the Stuttering Severity Instrument for Children and Adults (SSI-3; Riley, 1994). The total score of each participant was less than 6, the lower limit for mild stuttering in SSI-3. To exclude cognitive difficulties, two subtests of the Finnish Wechsler Intelligence Scale for Children–Third Edition (WISC-III; Wechsler, 2005), the Vocabulary verbal subtest and the Block Design performance subtest were administered to the participating children. These subtests were chosen because they correlate highly with the WISC-III overall score (Groth-Marnat, 2009). The hearing level of the children was screened using screening tone audiometry at 500, 1000, 2000, and 4000 Hz (SA 50, Entomed, Sweden), and only children with a normal hearing level were included in the sample. Moreover, all children had to attend the typical Finnish pre-primary and early primary education with no known language or learning problems.

Collection of Speech Samples

The conversational speech sample from the longitudinal data of the younger group of children was collected at ages 2, 3, and 4 during a 10-minute book reading situation and a 10-minute semistructured play situation between the mother and the child. A picture book and small plastic Duplo® toys representing objects from different semantic domains such as animals, people, and furniture, were used to elicit child speech. The mother and the child sat in their own chairs side by side at a table. The mother was instructed to interact with the child as they would normally do. The interactions were video-recorded in a laboratory setting.

The cross-sectional data collection of the four groups of older children (age levels 6, 7, 8, and 9) attending pre-primary and primary education were conducted either in the laboratory or in a silent room in the child's day-care centre or primary school. The children participated in two video-recorded sessions with the examiner: a semistructured conversation based on telling a story of a picture book and on uniform open-ended questions on hobbies, holidays, and other everyday activities to elicit child speech.

Procedure

The speech samples of each participant were literally transcribed by repeated listening, and unintelligible utterances, isolated affirmatives and negatives were deleted (cf. Ambrose & Yairi, 1999). From the transcriptions, 50 utterances per child, 25 consecutive utterances from two different situations, were selected for further analyses. Darley and Moll (1960) stated that a sample of 50 utterances provides adequate reliability. Moreover, Hutchins et al., (2005) referred to a minimum of 50 utterances for a representative sample. Conture (2001) on the other hand used a minimum of 300 words to evaluate the speech disfluency of participants. In the current study, we tried to adhere to both criteria. However, especially in the youngest age groups, 50 utterances did not include 300 words (Table 3). This is mainly due to the nature of Finnish morphology. For example, “my neighbour also has a cat” (“naa-pu-ril-la-ni-kin on kis-sa”) includes three words with nine syllables. The word and syllable counts were done in line with Yairi and Seery (2011).

The 50 utterances were phonetically transcribed based on acoustic analyses, saved in Microsoft Office Excel and edited using XMediaRecode from VOB files to WAV files. Subsequently, they were analysed using PRAAT 5.4. (Boersma & Weenink, 2012) to get results based on measurable, objective physical features of speech. Each utterance was defined as fluent or disfluent by repeated listening and based on the formant structure in the PRAAT. In addition, each disfluency was labelled using a classification system of nine disfluency types, based on the system by Ambrose and Yairi (1999) and Stes (2004) (see Table 4 for an overview). The transcription of the speech samples and classification of disfluencies were carried out by the first and last author together with four speech-language pathology students after extensive theoretical teaching, practicing with the help of video samples categorized together with detailed written examples of transcriptions as well as online consultation. The four students worked in pairs to analyse and categorise the speech samples to reach consensus on the data analyses and classification. In addition, all the unsure disfluencies were checked by the first writer together with the students.

Data Analyses

The total frequency and type of disfluencies as well as SLD/OD were calculated based on 100 words (see Pellowski & Conture, 2002; Tumanova et al., 2014) and 100 syllables (Ambrose & Yairi, 1999). Because the data did not meet the criteria of parametric tests (see e.g., Tumanova et al., 2014), the statistical analyses were based on nonparametric tests (SPSS version 25). The longitudinal data were analysed using the Friedman test, the cross-sectional data were analysed using the Kruskal-Wallis test and the post-hoc pairwise comparisons using Wilcoxon's tests and Mann-Whitney tests, respectively. Due to multiple comparisons, the significance level was adjusted using the Bonferroni correction.

Results

Total Frequency and Type of Disfluencies

According to the Friedman test, no significant age effect was found for the total frequency of disfluencies per 100 words or per 100 syllables among children followed up at ages 2, 3, and 4, χ^2 (2) = 0, 429, p = .807and χ^2 (2) = 1, 286, p = .526, respectively (for descriptive statistics, see Tables 5 and 6). In contrast, the Kruskal-Wallis test among the older children representing age groups 6, 7, 8, and 9 revealed significant between-group differences for the total frequency of disfluencies both when assessed per 100 words, χ^2 (3) = 10,553 p = .014, and per 100 syllables, χ^2 (3) = 12,083, p = .007. Pairwise comparisons showed that the group of 6-year-old children had more disfluencies per 100 words and per 100 syllables than the 7-year-olds, U = 15, p = .008 and U = 16, p = .010, respectively. In addition, 6-year-old children had more disfluencies per 100 syllables than the 9-year-olds, U = 10, p = .002.

Regarding the type of disfluencies, the Friedman tests showed no significant changes across 2- to 4-year-olds per 100 words or 100 syllables (for descriptive statistics, see Tables 5 and 6). Among the 6- to 9-year-old children, the Kruskal-Wallis test showed significant between-group differences both per 100 words and 100 syllables for revisions, χ^2 (3) = 8,448, *p* = .038 and χ^2 (3) = 7,837, *p* = .050, respectively, and for filled pauses per 100 syllables, χ^2 (3) = 8,400, *p* =.038. Pairwise comparisons showed that 6-year-old-children had more revisions than 7-year-old children, *U* = 18, *p* = .016 and *U* = 20, *p* = .023, respectively. For revisions, the pairwise comparisons did not reveal further difference between the age groups. A significant between-group differences both per 100 words and 100 syllables was found also for prolongations, χ^2 (3) = 9,755, *p* = .021 and χ^2 (3) = 10,685, *p* = .014, respectively. The frequency of prolongations was found to be higher in 7-year-old than in 8-year-old children, *U* = 15, *p* =.004 and *U* = 15, *p* =.004, respectively. Moreover, a significant between-group difference was found per 100 words for part word repetitions, χ^2 (3) = 9,136, *p* =.028, which was due to a higher frequency of part word repetitions in 9-year-old children when compared to 7-year-old children, *U = 12.5, p* = .004. Only one block in one child was encountered.

SLD and OD

The Friedman test revealed no significant changes per 100 words and per 100 syllables for the SLD, χ^2 (2) = 5.019, *p* = .081 and χ^2 (2) = 3.434, *p* = .180, respectively or for the OD, χ^2 (2) = 2.655, *p* = .265 and χ^2 (2) = 5.286, *p* = .071, respectively among the younger children followed up at ages 2, 3, and 4, (for descriptive statistics, see Table 5 and 6). Among the 6- to 9-year-old children, the Kruskal-Wallis tests suggested no between-group differences for the SLD per 100 words syllables χ^2 (3) = 3,225, *p* = .358, and per 100 syllables, χ^2 (3) = 3,556, *p* = .3141, but a significant between-group difference was found for OD both per 100 words, χ^2 (3) = 15,229, *p* = .002, and per 100 syllables, χ^2 (3) = 15,078, *p* = .002. Pairwise comparisons showed that the group of 6-year-old children had more OD per 100 words and per 100 syllables compared to the 7-year-olds, *U* = 9, *p* = .002 and

U = 10, *p* = .002, respectively, and to the 9-year-olds, *U* = 12, *p* =.004 and *U* = 8, *p* = .001, respectively.

The average amount of SLD ranged from 1.6 to 3.3/100 words across ages 2 to 4 years and from 1.0 to 1.9 across age groups 6 to 9 years (Table 5 and 6). In the younger age groups, 57% scored above the criterion of 3% of SLD per 100 words. When measured per 100 syllables, the average amount of SLD ranged from 1.0 to 1.8 across the ages 2 to 4 years, and from 0.5 to 1.0 across the age groups 6 to 9 years.

Discussion

The aim of the current preliminary study was to investigate age-related changes and differences in speech disfluencies in typically developing, Finnish-speaking 2- to 4- and 6- to 9-year-old children using a longitudinal and cross-sectional design, and to clarify whether the findings and norms of English or other languages can be applied to the Finnish language. This is highly important because at present it is mainly reference data for monolingual, English-speaking children that is available and used worldwide. Furthermore, these findings may also help clinicians to describe, define, and measure stuttering in Finnish-speaking children. The strengths of longitudinal designs over cross-sectional designs are well documented in developmental sciences. Cross-sectional data on children's speech collected at one point in time for different age groups can only explore the prevalence of disfluencies and hence differences between individuals. Longitudinal data, on the other hand, collected at two or more time points for the same group of children allow for detection of stability and change in the prevalence of disfluencies both within and between individuals.

We first hypothesised that there would be no significant difference in the total frequency of disfluencies in different age groups. The results for the longitudinal data on children followed-up from 2 to 4 years of age were in line with our hypotheses and verify the earlier cross-sectional studies assessing disfluencies per 100 syllables (Ambrose & Yairi, 1999) and per 100 words (Wexler &

Mysak, 1982) in English-speaking children. In accordance with the studies by Ambrose and Yairi (1999) and Wexler and Mysak (1982), Carlo and Watson (2003) did not find differences across age groups 3.5–4.0-year-old and 5.0–5.5-year-old Spanish-speaking children using a syllable-based metric. Only one, cross-sectional study, investigating Spanish-speaking children (Watson & Anderson, 2001), reported a difference in the total frequency across ages, i.e. a significantly higher total frequency of disfluencies both per 100 words and per 100 syllables in 3-year-old children when compared to 2-year-old children. A possible reason for this discrepancy with our and all other studies could be the different classification system used. Watson and Anderson included, for example, grammatical and nongrammatical pauses (i.e., silent pauses ≥1 second at grammatical or nongrammatical junctures) to their disfluency types. These disfluencies were not counted in our study or in the study by Ambrose and Yairi (1999). However, the different classification system cannot be the only reason for differences since it was also used by Carlo and Watson (2001) in older Spanish-speaking children, so perhaps language might also have a possible influence. Altogether, we suggest that the age, classification system, and the language that the child speaks may have an effect on children's speech disfluencies and should be taken into consideration when building normative data for normal disfluency.

Our cross-sectional results on age groups 6–9 years revealed a significant difference in the total frequency of disfluencies and this in contrast to our hypotheses. The total frequency of disfluencies was higher in 6-year-old than in 7-year-old and 9-year-old children. The 6-year-olds had more disfluencies than any other age as shown by the descriptive statistics. This was especially true when assessed per 100 words. The total frequency of disfluencies in the present study (9.2% per 100 words), however, equals the finding by Wexler and Mysak (1982) reporting 9.1% disfluencies per 100 words among 6-year-old children but not the results by Haynes and Hood (1977). In the latter study, the total frequency of disfluencies was 7.2% per 100 words among 6-year-olds, which is much lower than that found in the present study. Concerning the total frequency of disfluencies among 8-year-old children, the frequency of total disfluencies was more similar between the present study and the study by Haynes and Hood (5.8 and 6.8% per 100 words, respectively). Wexler and Mysak (1982) did not find a significant difference when they compared 4-year-old and 6-year-old children (9.1% per 100 words in both age groups). Likewise, Haynes and Hood did not find any significant difference between the age groups of 4-, 6- and 8-year-olds. Moreover, in the present study the total frequency of disfluencies was almost similar among 7- and 8-year-old children. The results of our study thus suggest that there is a peak at the age of 6 years in the total frequency of disfluencies, which was not found in other studies. This finding was not due to significant disfluencies of one or two children. There is a possibility that some other factor than age has an effect on the results, for example, the limited sample size in the current study. Therefore, the finding of our study needs to be interpreted with caution, and further studies are needed to verify this finding.

According to our second hypothesis, we expected to find age-related differences in the type of disfluencies. Against our hypothesis, no significant changes were found among the children followed-up from 2 to 4 years of age, whereas the results for 6- to 9-year-old children were in line with the hypothesis. Even though our findings did not support our hypothesis among the younger children, they accord with the cross-sectional findings by Ambrose and Yairi (1999) investigating the same age range as the current study. In contrast to our finding, differences in the types of disfluencies were earlier found in both English- and Spanish-speaking young children (Watson & Anderson, 2001; Wexler & Mysak, 1982). Wexler and Mysak's cross-sectional study included 2-, 4-, and 6-year-old English-speaking and Watson and Anderson's cross-sectional study 2- and 3-year-old Spanish-speaking children. In addition to the longitudinal design of the current study, one other reason for the different results between studies may be the different classification system used to define disfluencies. Wexler and Mysak had seven categories, Ambrose and Yairi had six categories, and Watson and Anderson fourteen categories. Watson and Anderson (2001) defined, for example, grammatical and ungrammatical pauses that were not defined in other studies. The classification used in the present study was closest to the classification used by Ambrose and Yairi (1999) which may partly explain the similar results between these two studies and is a reminder of the importance of classification when defining disfluencies for normative data.

Earlier studies have shown decrease in revisions and prolognations in fluently speaking person. Yairi & Clifton (1972) compared pre-school-aged children, high school seniors and geriatric persons and found that revisions decreased from preschool group to geriatric group. Similarly, prolongations were found to decrease from the age of 3.5- to 5-year-old children (DeJoy & Gregory, 1985), and from the age of 2 years to the age of 6 years (Wexler & Mysak, 1982). In line with this, among 6- to 9-year-old children, the frequency of revisions decreased from 6 to 7 years of age both per 100 words and per 100 syllables. Revisions are defined as a correction of word choice, grammatical or phonological errors, or adding or deleting lexical information. The achievement demands for 6-year-olds in the Kindergarten classrooms the year before Finnish children enters primary school may remarkably increase stress among children, which might appear as changes in speech fluency. It is possible that this is implied by revisions like re-formulation of speech as shown during the semistructured conversations with the examiner. During the first years of primary school, children's social, emotional, and cognitive readiness for school increases and they gradually learn to cope with the increased achievement demands, which might appear in speech fluency as a decrease in revisions. Similarly, the frequency of prolongations per 100 syllables decreased from 7 to 8 years of age, whereas for part word repetitions, the frequency increased from 7 to 9 years of age per 100 words but not per 100 syllables, which has not been reported earlier. Prolongation of sounds is often used in Finnish communicational situations purposely to keep the turn while thinking of the next word or during sentence formulation (e.g., I thiiink that…). Even though this communicative act was left out of the disfluency counts, there is a possibility that sometimes it was not clear whether the prolongation was a disfluency or a communicative act and was measured as a prolongation. Concerning the category of part word repetition, it is sometimes very similar to revision, word or sentence formulation, especially at the end of a word (e.g., I want mil/milk). Thus, there is a possibility that some revisions were categorised as part word repetitions even after extensive theoretical teaching, practicing with the help of video samples categorized together with detailed written examples of transcriptions as well as online consultation. Einarsdóttir and Ingham (2005) raised the question of the reliability of classifications. In the current study the data for 6- and 7-year-old and correspondingly 8- and 9-year-old children were classified and analysed by the same person coming from the same university unit with the same educational background. Cordes and Ingham (1995) showed that researchers working in the same location displayed high agreement with each other in classifications. Thus, the differences in the results of our study may not be only due to different classification used by the examiners. Moreover, Wingate (1987) raised the question as to whether classification used to diagnose stuttering is appropriate for the classification of normal disfluency, and should, for example, communicative differences between cultures and languages be assessed to reach a real understanding of normal disfluencies. Moreover, Leclercq et al. (2018) mentioned cultural and communicative differences. Even though these differences are not in the scope of the present study, the results such as ways of keeping turns by prolonging sounds between turns in communicative situations may be real findings on disfluencies in Finnish fluent speakers.

Our third hypothesis was that there are no changes in the frequency of SLD whereas the frequency of OD will increase by age. This hypothesis was supported by our findings showing no changes in the total amount of SLD among the younger or older groups of children. Similarly to our study, Ambrose and Yairi (1999) did not find a difference for SLD between 2- to 5-year-old children. Moreover, no increase for OD was found among the younger children followed up at age 2, 3, and 4 years. This finding was against our hypothesis but in line with Ambrose and Yairi’s cross-sectional study (1999), reporting no changes for OD in children from 2 to 5 years. However, in accordance with the hypothesis we found a significant age group difference across older children, especially due to the high frequency of OD in 6-year-old children both per 100 words and per 100 syllables (7.2%

and 3.8%, respectively). To our knowledge, no other studies have reported changes for OD across the ages of 6–9 years, and therefore it is difficult to compare our finding to earlier studies. Some studies report the frequency and changes for OD in younger age groups (Ambrose & Yairi, 1999; Pellowski & Conture, 2002). The frequency of OD among the 6-year-olds in the present study was similar to that found by Leclercq et al. (2018) in French-speaking 4-year-old children (i.e., 7.9% per 100 words) but higher than found in 3- to 6-year-old English-speaking children per 100 words (Pellowski & Conture, 2002; Tumanova et al., 2014). In Ambrose and Yairi (1999), the mean frequency of OD per 100 syllables was found to range from 3.8 to 4.5 across the age groups of 2–4 years. In our study, the frequency of OD per 100 syllables ranged from 0.93 to 2.44 across the age groups of 2–4 and 7–9 years, the mean OD 3.8% per 100 syllables being the highest in 6-year-old children. In our 6-year-old group, the frequency of OD per 100 syllables was near the frequencies reported earlier in younger age groups per 100 syllables (Ambrose & Yairi, 1999; Carlo & Watson, 2003) whereas in 2- to 4- and 7- to 9-year-old children the frequency of OD was lower than or similar to many other studies (Ambrose & Yairi, 1999; Carlo & Watson, 2003; Natke et al., 2006). The high frequency of OD in 6-year-olds is mainly due to the high frequency of revisions and filled pauses, which are often linked to hesitations. The reason for this high frequency was discussed earlier, and we assume that filled pauses are linked to the same kind of reasons.

In our fourth hypothesis, we assumed that a syllable-based metric would be more sensitive than the word-based metric in counting SLD and OD. The often-used criterion for normal disfluency is less than 3 SLD per 100 words (e.g., Boey et al, 2007; Conture, 2001; Tumanova et al., 2014) or per 100 syllables (e.g., Ambrose & Yairi, 1999; Natke et al. 2006; Yairi & Ambrose, 2005). In the younger age group (from 2 to 4 years) the number of SLD in our study ranged from 1.5 to 3.3 per 100 words and was lower in the older age groups. SLD was high in 3-year-old children (i.e., 3.3% per 100 words) especially when compared to the study by Pellowski and Conture (2002) reporting 1.1% SLD per 100 words in English-speaking 3- to 4-year-old children. In French-speaking 4-year-old children, Leclercq et al. (2018) found 2.6% SLD per 100 words, which is higher than that found in English-speaking children but does not exceed the limit of 3%. In children aged 6 to 9 years, the frequency of SLD per 100 words was in the same range as found earlier in younger children per 100 words (Pellowski & Conture, 2002; Tumanova et al., 2014). Of particular (clinical) relevance is that the frequencies of SLD per 100 words as a clinical marker of stuttering in Finnish-speaking children may be problematic because 57% of the children in the younger group scored above this criterion. In other words, using the norm of 3 SLD per 100 words as a clinical marker of stuttering would result in false positive diagnoses of stuttering in the Finnish language, specifically at early ages among 2- to 4-year olds. This finding adds to similar conclusions of previous studies in another language, i.e. French (e.g., Leclercq, 2018) and in bilingual children (Eggers et al., 2019).

When a syllable-based metric was used, the frequency of SLD per 100 syllables was very similar at ages 2, 3, and 4 years in the present study and the study by Ambrose and Yairi (1999) (1.3% vs. 1.2%, 1.8% vs. 1.6%, and 0.9% vs. 0.9% per 100 syllables, respectively). Moreover, the frequency of SLD across older age groups (i.e., 6–9 years) was similar to the frequency for the younger age group in this study, ranging from 0.51 to 1.03 SLD per 100 syllables. These frequencies are equal to or are even lower than in many other earlier studies among younger children (Ambrose & Yairi, 1999; Carlo & Watson, 2003, Natke et al., 2006). Our finding shows that the frequency of SLD does not change in Finnish-speaking children from early childhood to primary school age, against common belief. Furthermore, the frequency of SLD in all groups is less than 3% per 100 syllables. Therefore, it seems that especially for young children, the syllable-based metric would be a more reliable clinical marker of stuttering than the word-based metric, as also reported earlier (Yairi, 1997; Natke et al., 2006). To be able to follow up the development of the child, the same metric should be used throughout all age levels. Furthermore, it is clinically easier to reach a sufficiently large sample size for diagnosing purposes using the syllable-based metric in Finnish language where the words are long and children's first words are already 2-syllable words as indicated by our data, the average sample size in our 2-year-old children being 84 words whereas the number of syllables was 164.

Several authors agree that excessive tension, as frequently seen in blocks, is not typical of normally fluent speech and the more this occurs, the higher the likelihood is of speech being perceived as abnormal or stuttered (e.g., Einarsdóttir, 2009; Ward, 2006). Several studies have reported no blocks (Boey et al., 2007; Juste & Andrade, 2010) or only very low numbers of blocks in CWNS (Pellowski & Conture, 2002). In accordance with this, only one block was defined in one 3-year-old child in the current study.

One of the caveats of this study is the size of the groups. Although the group sizes were small, their age ranges were similar, making it possible to report and compare data on a yearly basis. Furthermore, for the youngest groups, the same cohort of children was longitudinally investigated at the ages of two, three, and four. This longitudinal design probably depicts more reliably and validly quantitative and qualitative changes over time. Moreover, speech samples were collected in two varied speaking situations from each child, which is supposed to give an accurate assessment of a child's overall speech production (Riley, 1994).

A second caveat is the size of the speech samples. For the 2-year-olds, the sample size was below 100 words. This may be partly due to the data collection. Data was collected during a semistructured play and book reading situation in a laboratory. Many children seemed to be shy of the unknown place and produced little speech. In addition, book reading with some children resulted in naming of objects and further to one-word utterances. However, in the Finnish language one-word utterance often include several syllables and morphemes, and therefore the number of words may be remarkably lower than the number of syllables. Accordingly, although 50 utterances did not always yield 100 words, the criterion of 100 syllables was reached and, what is more, the results at the youngest age level corroborated the findings of Ambrose and Yairi (1999). In all other ages, the number of disfluencies were counted per 100 words or syllables making the data in different age groups comparable to each other as well as to earlier studies using the same metric. The results of this preliminary study may suggest that, keeping the caveats in mind, the results bring new data on normal speech disfluency especially in children learning to speak Finnish.

Conclusions

To the best of our knowledge, no recent studies have reported on the development of speech disfluencies from two years of age up to school age. The results of the current study showed that the total frequency and type of speech disfluencies in Finnish 2- to 4-year-old children was similar, while some age-related changes were encountered across the age groups of 6–9-years. Even though both SLD and OD were encountered in all groups, the frequency of SLD was remarkably smaller than the frequency of OD. In addition, only one block in one child was found. A clinically relevant notion was that the often-used criterion for normal disfluency, i.e., <3 SLD per 100 syllables, was applicable in all age groups but the word-based metric was not. Thus, our preliminary findings suggests that the English guideline for normal disfluency is not totally valid for the Finnish language. Moreover, a syllable-based metric should be used, revealing more exactly the differences among and between different age groups.

References

Adams, M. R. (1977). A clinical strategy for differentiating the normally nonfluent child and the incipient stutterer. *Journal of Fluency Disorders*, *2*(2), 141–148. https://doi.org/10.1016/0094-730X(77)90017-1

Ambrose, N.G., & Yairi, E. (1999). Normative disfluency data for early childhood stuttering. *Journal of Speech-Language and Hearing Research*, *42*(4), 895–909. https://doi.org/10.1044/jslhr.4204.895

Bloodstein, O., & Ratner, N. (2008). *A handbook on stuttering* (6th ed.). Thomson Delmar Learning.

Boersma, P., & Weenink, D. (2012). *Praat: Doing phonetics by computer (version 5.4)* [Computer software]. Institute of Phonetic Sciences.

Boey, R., Wuyts, F., Van de Heyning, P., Bodt, M., & Heylen, L. (2007). Characteristics of stuttering-like disfluencies in Dutch-speaking children. *Journal of Fluency Disorders*, *32*(4), 310–329. https://doi.org/10.1016/j.jfludis.2007.07.003

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. https://doi.org/10.1016/S0094-730X(03)00004-4

Conture, E. G. (2001). *Stuttering: Its nature, diagnosis, and treatment*. Allyn and Bacon.

Cooper, E., & Cooper, C. (1985). *Cooper personalized fluency control therapy handbook* (Rev. ed.). DLM Teaching Resources.

Cordes, A. K., & Ingham, R. J. (1995a). Judgments of stuttered and nonstuttered intervals by recognized authorities in stuttering research. *Journal of Speech and Hearing Research, 38*(1), 33–41. https://doi.org/10.1044/jshr.3801.33

Curlee, R. (1993). Identification and management of beginning stuttering. In R. Curlee (Ed.), *Stuttering and related disorders of fluency* (pp. 1–22). Thieme Medical Publishers.

Darley, F. L., & Moll, K. L. (1960). Reliability of language measures and size of language samples. *Journal of Speech and Hearing Research*, *3*(2), 166–173. https://doi.org/10.1044/jshr.0302.166

DeJoy, D. A., & Gregory, H. H. (1985). The relationship between age and frequency of disfluency in preschool children. *Journal of Fluency Disorders*, *10*(2), 107–112. https://doi.org/10.1016/0094-730X(85)90019-1

Eichorn, N., & Fabus, R. (2012). Assessment of fluency disorders. In C. Stein-Rubin and R. Fabus (Eds.), *Clinical assessment and professional report writing in speech-language pathology* (pp. 347–398). Delmar/Cengage Learning.

Eggers, K., & Elen, R. (2018). Spraakonvloeiendheden bij personen die niet stotteren: Invloed van geslacht en leeftijd. *[Speech disfluencies in persons who do not stutter: Influence of gender and age]. Logopedie [Logopedics].*

Eggers, K., Van Eerdenbrugh, S. & Byrd, C. (2019). Speech disfluencies in bilingual Yiddish-Dutch speaking children. *Clinical Linguistics & Phonetics*, *34*(6), 576-592.

Einarsdóttir, J. (2009). *The identification and measurement of stuttering in preschool children* [Unpublished doctoral dissertation]. University of Reykjavik, Iceland.

Einarsdóttir, J., & Ingham, R. (2005). Have disfluency-type measures contributed to the understanding and treatment of developmental stuttering? *American Journal of Speech-Language Pathology*, *14*(4), 260–273. https://doi.org/10.1044/1058-0360(2005/026)

Groth-Marnat, G. (1997). *Handbook of psychological assessment.* Wiley.

Hakulinen, A. (2000). *Suomen kielen rakenne ja kehitys. 5. muuttumaton painos. [The structure and development on Finnish language]*. Department of Finnish, University of Helsinki, Finland.

Haynes, W., & Hood, S. (1977). Language and disfluency variables in normal speaking children from discrete chronological age groups. *Journal of Fluency Disorders, 2*(1)*,* 57–74. https://doi.org/10.1016/0094-730X(77)90010-9

Hutchins, T. L., Brannick, M., Bryant, J. B., & Silliman, E. R. (2005). Methods for controlling amount of talk: Difficulties, considerations and recommendations. *First Language*, *25*(3), 347-363. https://doi.org/10.1177/0142723705056376

Johnson, W., Boehmler, R., Dahlstrom, W., Darley, F., Goodstein, L., Kools, J., Neeley, J., Prather, W. Sherman, D., Thurman, C., Trotter, W., Williams, D., & Young, M. (1959). *The onset of stuttering*. University of Minnesota Press.

Juste, F., & Andrade, C. (2010). Speech disfluency types of fluent and stuttering individuals: Age effects. *Folia Phoniatrica & Logopaedica, 63*(2), 57–64. https://doi-org.ezproxy.utu.fi/10.1159/000319913

Kunnari, S. (1998). Syllable number in early words. In K. Heinänen & M. Lehtihalmes (Eds.), *Proceedings of the Seventh Nordic Child Language Symposium* (pp. 85–88). University of Oulu, Finland.

Kunnari, S. (2000). *Characteristics of early lexical and phonological development in children acquiring Finnish* [Unpublished doctoral dissertation]. University of Oulu, Finland.

Kunnari, S., Savinainen-Makkonen, T., & Saaristo-Helin, K: (2012). Fonologiatesti. Lasten äänteellisen kehityksen arviointimenetelmä [*Phonological Test. An assessment method for children´s phonological development]*. Niilo Mäki Institute, Finland.

Laalo, K. (2011). *Lapsen varhaiskielioppi ja miniparadigmat* [Early grammar and minimi paradigms in children]. Suomalaisen Kirjallisuuden Seura.

Leclercq, A.-L., Suaire, P., & Moyse, A. (2018). Beyond stuttering: Speech disfluencies in normally fluent French-speaking children at age 4. *Clinical Linguistics and Phonetics, 32*(2), 166-179. https://doi.org/10.1080/02699206.2017.1344878

Natke, U., Sandrieser, P., Pietrowsky, R., & Kalveram, K. (2006). Disfluency data of German preschool children who stutter and comparison children. *Journal of Fluency Disorders*, *31*(3), 165–176. https://doi.org/10.1016/j.jfludis.2006.04.002

Pellowski, M., & Conture, E. (2002). Characteristics of speech disfluency and stuttering behaviors in 3- and 4-year-old children. *Journal of Speech-Language and Hearing Research*, *45*(1), 20–34. https://doi.org/10.1044/1092-4388(2002/002)

Riley, G. (1994). *Stuttering Severity Instrument of Children and Adults. 3.* CC Publications.

Savinainen-Makkonen, T. (2000). Learning long words – A typological perspective. *Language and Speech*, *43*(2), 205–222. https://doi.org/10.1177/00238309000430020401

Silvén, M., Ahtola, A., & Niemi, P. (2003). Early words, multiword utterances and maternal reading strategies as predictors of mastering word inflection in Finnish. *Journal of Child Language, 30(2),* 253–279. https://doi.org/10.1017/S0305000902005548

Stes, R. (2004). *Vloeiendheidsstoornissen: Rationale eclectische sociaal-cognitieve gedragstherapeutische benadering – Deel III: Normaal (on)vloeiende spreekcomponenten.* Lezing postgraduaat stottertherapie, Antwerpen.

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. https://doi.org/10.1016/j.jcomdis.2014.01.003

Ward, D. (2006). *Stuttering and cluttering: Frameworks for understanding and treatment.* Psychology Press.

Watson, J., & Anderson, R. (2001). Disfluencies of 2- and 3-year-old Spanish-speaking children from Puerto Rico. *Contemporary Issues in Communication Science and Disorders*, *28,* 140–150.

Wechsler, D. (2005). *Wechsler Intelligence Scale for Children*, (WISC-III). [Finnish version Heiskari, P., Jakobsen, B. Marila, A. (Eds)]. Hogrefe Psykologien Kustannus.

Wexler, K. (1982). Developmental disfluency in 2-, 4-, and 6-year-old boys in neutral and stress situations. *Journal of Speech-Language and Hearing Research*, *25*(2), 229–234. https://doi.org/10.1044/jshr.2502.229

Wexler, K. B., & Mysak, E. D. (1982). Disfluency characteristics of 2-, 4-, and 6-year-old males. *Journal of Fluency Disorders, 7*(1*),* 37–46. https://doi.org/10.1016/0094-730X(82)90036-5

Williams, D. E., Silverman, F. H., & Kools, J. A. (1968). Disfluency behaviour of elementary school stutterers and nonstutterers: The adaptation effect. *Journal of Speech and Hearing Research*, *11*(3), 622–630. https://doi.org/10.1044/jshr.1103.622

Wingate, M. (1987). Fluency and disfluency: Illusion and identification. *Journal of Fluency Disorders, 12*(2), 79–101. https://doi.org/10.1016/0094-730X(87)90015-5

Yairi, E. (1981). Disfluencies of normally speaking two-year-old children. *Journal of Speech and Hearing Research*, *24*(4), 490–495. https://doi.org/10.1044/jshr.2404.490

Yairi, E. (1982). Longitudinal studies of disfluencies in two-year-old children. *Journal of Speech and Hearing Research*, *25*(1), 155–160. https://doi.org/10.1044/jshr.2501.155

Yairi, E. (1997). Disfluency characteristics of childhood stuttering. In R. Curlee & G. Siegel (Eds.), *Nature and treatment of stuttering* (pp. 49–78). Allyn & Bacon.

Yairi, E., & Ambrose, N. G. (2005). *Early childhood stuttering: For clinicians by clinicians*. PRO-ED, Inc.

Yairi, E., & Ambrose, N. G. (2013). Epidemiology of stuttering: 21st century advances. *Journal of Fluency Disorders*, *38*(2). 66–87. https://doi.org/10.1016/j.jfludis.2012.11.002

Yairi, E., & Clifton, N. (1972). Disfluent speech behavior of preschool children, high school seniors, and geriatric persons. *Journal of Speech and Hearing Research, 15*(4), 714–719. https://doi.org/10.1044/jshr.1504.714

Yairi, E., & Seery, C. (2011). *Stuttering: Foundations and Clinical Applications.* Pearson Education.