

To cite this article: Anna-Leena Martikainen, Tuula Savinainen-Makkonen, Päivi Laukkanen-Nevala & Sari Kunnari (2019): Intra-word accuracy and consistency in Finnish-speaking children with speech sound disorder compared to their typically developing peers, *Clinical Linguistics & Phonetics*, DOI: 10.1080/02699206.2019.1696610

Intra-word accuracy and consistency in Finnish-speaking children with speech sound disorder compared to their typically developing peers

Anna-Leena Martikainen^{a*}, Tuula Savinainen-Makkonen^a, Päivi Laukkanen-Nevala^a and Sari Kunnari^a

^aLogopedics, University of Oulu, Oulu, Finland

*Anna-Leena Martikainen, University of Oulu, Faculty of Humanities, Logopedics, PO Box 1000, 90014 University of Oulu, Finland

Email: anna-leena.martikainen@student.oulu.fi

Intra-word accuracy and consistency in Finnish-speaking children with speech sound disorder compared to their typically developing peers

This study examined intra-word accuracy and consistency in 32 preschool-aged Finnish-speaking children with speech sound disorder (SSD) compared to their typically developing (TD) age- and gender-matched peers. Accuracy and consistency of speech production were assessed by a picture-naming task repeated three times in one assessment session. Responses were classified into four categories: 1) consistently correct, 2) consistently incorrect, 3) variable with hits (when a child's variable responses included at least one matched with the adult target), and 4) variable with no hits (when responses included at least two different response types without the matched adult target). In addition, relationships between intra-word accuracy and consistency and children's receptive vocabulary knowledge and articulatory ability based on spontaneous speech samples were investigated. The findings showed that the children with SSD produced significantly more often 'consistently incorrect' and 'variable with no hits' responses than the TD children. There was a significant negative correlation between 'variable with no hits' responses and receptive vocabulary knowledge and articulatory abilities among the children with SSD. As intra-word accuracy and consistency has not previously been studied in Finnish children with SSD, the findings highlighted the need for drafting guidelines for assessment and intervention by paying close attention to

high intra-word variability without correct word forms already from age three onwards.

Keywords: child; intra-word variability; speech sound disorder; typical speech development; Finnish

Introduction

Intra-word consistency of speech production, i.e. whether a child pronounces the same word similarly or differently when repeating it within a short period of time, has been the object of study in typically developing children (TD) (Macrae, 2013; Martikainen, Savinainen-Makkonen, & Kunnari, 2019; Sosa, 2015) and children with speech sound disorder (SSD) (e.g. Dodd, Holm, Crosbie, & McCormack, 2005; Dodd & Leahy, 1989; Iuzzini, 2012; Iuzzini-Seigel, Hogan, & Green, 2017; Marquardt, Jacks, & Davis, 2004). Knowledge of the typical trajectory of intra-word accuracy and consistency is essential in order that speech-language therapists can be aware of when variability of speech production is a relevant factor to be taken into account in assessing and diagnosing children with speech impairment. On the other hand, the occurrence and characteristics of variability in the speech of children with SSD are also important to investigate, in order to determine whether different subgroups and possible underlying causes can be identified.

Overall intra-word variability has been reported to decrease after the toddler years among TD children (Faes & Gillis, 2018; Holm, Crosbie, & Dodd, 2007; Martikainen et al, 2019; Sosa, 2015). However, variability scores found in the studies have ranged remarkably, possibly due to differences in the phonology and syllable and word structures of the target languages, as well as in the study procedures: among 3+

years the figures vary from 12% to 62%, 4+ years from 6% to 39%, and 5+ years from 4% to 36%. Full details of these assessment findings have been compiled in Table 1.

Table 1. A sample of assessment findings concerning overall speech variability (based on broad transcription) in typically developing children.

Assessment task	Language	Overall variability	References
Spontaneous speech	Dutch	3+ years 56% ^a	Faes & Gillis (2018)
		4+ years 39% ^a	
		5+ years 36% ^a	
Picture naming (IA)	English	3+ years 12% ^a	Holm et al. (2007)
		4+ years 6% ^a	
		5+ years 4% ^a	
		6+ years 1% ^a	
Picture naming (IA)	English	3+ years 62% ^a	Sosa (2015)
Picture naming (VM)	Finnish	3+ years 16% ^b	Martikainen et al. (2019)
		4+ years 10% ^b	
		5–6+ years 5% ^b	

Note. IA = Word Inconsistency Assessment. VM = A variability measure with a similar measurement procedure to IA but with 20 Finnish words. ^a Mean percentage. ^b Median percentage.

Considering different sub-groups of SSD presented in the literature (e.g. Dodd et al., 2005; Shriberg, 2004), children with phonological and/or articulation disorders have also been reported to have inconsistent speech production (Dodd et al., 2005; Iuzzini, 2012; Macrae, Tyler, & Lewis, 2014), although the more typically associated features

of phonological disorder are consistent non-developmental error patterns and, in the case of articulation disorder, consistent substitutions and/or distortions of the same sounds (Dodd et al., 2005). A high level of inconsistency is particularly associated with childhood apraxia of speech (CAS) (American Speech-Language-Hearing Association, [ASHA], 2007; Iuzzini, 2012; Iuzzini-Seigel et al., 2017; Marquardt et al., 2004) and inconsistent speech disorder (Dodd & Leahy, 1989; Dodd et al., 2005). For CAS, no pathognomonic signs have been found but a deficit in consistency of speech production is one of the features which have led to some agreement among researchers (ASHA, 2007; Iuzzini-Seigel et al., 2017), and is the most frequently used criterion in identifying CAS by speech-language therapists (Forrest, 2003; Malmenholt, Lohmander, & McAllister, 2017). For inconsistent speech disorder, intra-word variability is the main clinical characteristic used in a diagnosis, but contrary to CAS, no oro-motor implementation deficits exist (Bradford & Dodd, 1996).

Lexical level inconsistency has been reported in studies examining English language acquisition in children with SSD. Iuzzini (2012) assessed lexical level inconsistency in children with CAS, phonological disorder, and typical speech development (aged 3;0–5;11) by the Word Inconsistency Assessment (IA) (Dodd, Hua, Crosbie, Holm, & Ozanne, 2006), in which a child is asked to name 25 pictures of words three times within the assessment session. Generally, the typically developing children were the most consistent, the children with CAS the most inconsistent and scores of the children with phonological disorder were between them. Based on narrowly transcribed data (distortions, substitutions, deletions, and glottal stops were considered as errors), Iuzzini determined a diagnostic cutoff score of less than 48% for phonological disorder and higher than 52% for CAS. The diagnosis of inconsistent speech disorder has also been based on the result of the Word

Inconsistency Assessment: When at least 40% of words are produced variably based on broadly transcribed responses (i.e. without phonetic details), a child is considered to have an inconsistency disorder (Dodd et al., 2006). Macrae et al. (2014) studied 18 children with SSD without an existing diagnosis of CAS (aged 3;6–5;5) and calculated their lexical level variability from a story retell when substitutions and omissions of consonants were considered as errors. Mean variability for the participants was 41%.

In the research literature, speech perception (Sayyahi, Soleymani, Akbari, Bijankhan, & Dolatshahi, 2017), phonological planning (Dodd et al., 2005; Holm, Farrier, & Dodd, 2008), and motoric implementation of speech (Forrest, Dinnsen, & Elbert, 1997; Iuzzini, 2012; Iuzzini-Seigel et al., 2017) have been proposed to result in variability in speech production. Concerning typical speech perception, variability in input, for example hearing words spoken by different talkers, is stated to be necessary for a child to establish phoneme categories and their boundaries of the ambient language, and thus robust phonological representations of words (Richtsmeier, Gerken, Goffman, & Hogan, 2009). However, a considerable deficit in auditory discrimination, which is rather well demonstrated at least in some children with SSD (e.g. Froud & Khamis-Dakwar, 2012; Hearnshaw, Baker, & Munro, 2018; Maassen, Groenen, & Crul, 2003; Munson, Edwards, & Beckman, 2005; Nijland, 2009; Rvachew & Jamieson, 1989; Sayyahi et al., 2017), may lead to inability to establish stable categorical representations for native phonemes and thus to variability of speech production (Sayyahi et al., 2017).

A deficit in assembling a phonological plan for production of a target utterance has also been hypothesized to be connected with inconsistent speech (Dodd et al., 2005; Holm et al., 2008). That is, a child may have good phonological representations, but selecting and sequencing phonemes for the target word or sentence are impaired.

Consequently, a motor programme for articulation set up from that plan “may have articulatory parameters that are too broad” (Dodd et al., 2005, p. 67).

From a motoric framework, some level of instability and motor flexibility is needed so that a child’s speech production system can modulate, for example, to anatomical changes of oral structures and to a different phonetical context of words (Whiteside, Dobbin, & Henry, 2003). Hypothetically speaking, highly inconsistent movements of speech organs during speech development may, however, hinder the establishment of a relationship between articulatory patterns and their auditory and somatosensory targets for later speech production. Compared to typically developing children, variability in articulatory movements based on kinematic measures, presumably reflecting a planning and/or programming level of speech production, has been reported to be prominent in children with CAS (e.g. Grigos, Moss, & Lu, 2015; Iuzzini, 2012; Nijland, Maassen, & van der Meulen, 2003; Nijland et al., 2003; Terband, Maassen, van Lieshout, & Nijland, 2011, Vuolo & Goffman, 2016). It has been stated, however, that from the findings hitherto reported it is not possible to conclude straightforwardly the existing relationship between variable articulatory movements and segmental level inconsistency (Vuolo & Goffman, 2016). Moreover, at phonemic level measures, no relationships have been found in typically developing preschool-aged children between either overall articulatory/phonological skills or phonological complexity of the target word and intra-word variability (Faes & Gillis, 2018; Macrae & Sosa, 2015; Martikainen et al., 2019). To the best of our knowledge, this factor has not previously been studied among children with SSD.

Another possible explanation for inconsistency in speech production, derived from a core deficit in speech motor skills, may be a flow-on effect to further higher-level phonological processes (Forrest et al., 1997; Iuzzini, 2012). If sounds are

repeatedly produced inconsistently, a child receives different perceptual patterns, which may have a deteriorating effect on categorical representations of phonemes of the ambient language (Forrest, Elbert, & Dinnsen, 2000). Terband, Maassen, Guenther, and Brumberg, (2014) provided one more viewpoint when they found with their computer simulations that simultaneous deficits in speech motor programming and auditory processing produced symptoms typical of CAS, including token-to-token variability, among others.

Speech inconsistency has also been approached via children's vocabulary knowledge. Among typically developing preschool-aged children the results have been mixed: Martikainen et al. (2019) found no relationship between expressive or receptive vocabulary and speech variability among children aged 3;0 to 6;11 whereas Macrae and Sosa (2015) found that expressive vocabulary only predicted the variance in rates of inconsistency among children aged 2;6 to 4;2. For children with inconsistent speech disorder, Dodd et al. (2005) reported a lower receptive and expressive vocabulary knowledge as compared to controls, but in a study of Holm et al. (2008) only a receptive vocabulary was administered and no difference was observed. Furthermore, Macrae et al. (2014) found a negative association between word inconsistency and receptive vocabulary in children with SSD. Like Macrae and Sosa (2015), Macrae et al. (2014) suggested that the finding was related to the children's lexical representations, which were less stable due to small vocabulary according to the lexical restructuring model (Metsala & Walleys, 1998). This model states that underlying representations of words gradually become phonologically more detailed when an expansion of the child's lexicon gives rise to the need to distinguish words from each other. Some researchers have, however, challenged this impression

and proposed that even young children's presentations may be phonologically robust (Demuth, 2011; Swingley, 2005).

Some phonological aspects of Finnish

Since language-specific characteristics may influence the occurrence of intra-word accuracy and consistency in the early phases of speech development, it is worthwhile to review some basic facts of the target language under investigation. The Finnish consonant inventory /p, t, d, k, n, m, ŋ, l, r, s, j, v, h/ has been reported to be comprehensive (Kunnari, Savinainen-Makkonen, & Paavola, 2006; Savinainen-Makkonen, Kaikkonen, Saaristo-Helin, & Kovasiipi-Nieminen, 2009), excluding /r/ and /d/ (Saaristo-Helin, 2009), and the vowel inventory, /ɑ, e, i, o, u, y, æ, œ/, almost completely mastered (Kunnari, Savinainen-Makkonen, & Saaristo-Helin, 2012) at the age of 3 on. Contrary to word-initial and word-final consonant clusters, heterosyllabic clusters are common in Finnish words, and children are reported to manage with them rather accurately at age 3+ (Saaristo-Helin, 2009). Monosyllabic words are rare in Finnish, whereas bisyllabic words are the most frequent (Karlsson, 1983), and Finnish children are reported to be able to produce them during the period of the first fifty words (e.g. Kunnari, 2002; Saaristo-Helin, 2009; Savinainen-Makkonen, 2000). Long words with three or more syllables are common (Karlsson, 1983), and in addition the suffixes attached to word stems lengthen many shorter words further (e.g. *auto* 'a car', *autossa* 'in a car', *autossani* 'in my car'). Typically developing children can manage these longer words without truncation at 3+ years (Saaristo-Helin, 2009).

Current study

Most of the research into intra-word variability has been carried out among English-speaking children. Studies on typologically different languages could, however, strengthen our comprehension of the phenomenon. Furthermore, the findings of earlier studies among TD children raise the significance of language-specific investigations, as the trajectory of the development appears to vary between languages. In addition, previous research on possible associations between intra-word variability and child-specific characteristics has been sparse and the findings partly contradictory.

Understanding factors correlated with inconsistent speech production could provide some indication of appropriate intervention methods for children with SSD. The specific questions of the current study were: (a) how accurate and consistent is the speech production of Finnish children with SSD, compared to age- and gender-matched TD children? and (b) are receptive vocabulary knowledge and articulatory ability associated with children's speech accuracy and consistency?

Method

Participants

A total of 64 children, aged from 3;2 to 5;4, participated this study: 32 children with SSD (14 girls and 18 boys) and 32 typically developing peers who were matched for age (within 2 months) and gender. The participants were recruited in several ways, including an announcement in a local newspaper and on the notice-boards of local child health clinics, and an information letter to local child care centres. Of the 45 children with SSD proposed to the study by their parents, 13 were excluded due to

weak receptive language skills and/or inability to name pictures of the phonology test. The TD children were a subset of children who participated in the study described in Martikainen et al. (2019).

For both study groups the criteria for inclusion were normal receptive language skills as evidenced by a standard score of 85 or higher on the Reynell Developmental Language Scales (Kortesmaa, Heimonen, Merikoski, Warma, & Varpela, 2001), and normal receptive vocabulary as evidenced by a standard score no more than 1 SD below the mean on the Receptive One-Word Picture Vocabulary Test – 4 (ROWPVT–4, Martin & Brownell, 2010). The results of Mann-Whitney U -tests confirmed that the study groups did not differ in receptive language skills ($p > .05$).

The other inclusion criteria were normal oro-motor skills, no reported neurological diseases, and use of Finnish as the primary language. Furthermore, all the participants were required to have normal non-verbal intelligence, which was determined for the children with SSD by a standard score of 85 or higher on the Leiter International Performance scale–3rd Edition (Roid, Miller, Pomplun, & Koch, 2013) and for the TD children by a parent and daycare teacher report. Additionally, all the SSD participants passed bi-lateral hearing screenings for pure tone at 500, 1000, 2000 and 4000 Hz presented at 20 dB HL. For the TD participants, all parents reported that their child had passed annual hearing screening tests made at a local child health clinic. Finally, inclusion criteria for the SSD group included a score under the 16th percentile on the Finnish Test for Phonology (Kunnari et al., 2012) and for the TD group at or above the 16th percentile. All the SSD participants were either followed up by their local child health clinic or were waiting for a referral or an invitation to an assessment by a speech-language therapist for suspected SSD. No attempt was made to classify children to the different types of SSD because of the lack of appropriate

diagnostic criteria for Finnish-language children. In both groups, case history information was obtained by asking the parents to complete a questionnaire. A summary of demographics and test scores by groups is presented in Table 2.

Table 2. Participants' mean ages, and the mean and standard deviations for the tests which were used as inclusion criteria. Values are presented separately for children with speech sound disorder (SSD) and for children with typical development (TD).

Study group	Mean age (months)	RDLS III	ROWPVT	Phonology test
SSD	52.53 (7.5)	102.56 (9.1)	79.63 (22.5)	3.09 (4.2)
TD	52.16 (7.4)	103.81 (9.2)	77.66 (18.7)	65.69 (23.3)

Standard deviations are in parentheses. SSD = children with speech sound disorder; TD = typically developing children; RDLS III = The Reynell developmental Language Scales III, Comprehension Scales, standard score (M = 100, SD = 15); ROWPVT = Receptive One-Word Picture Vocabulary Test, raw score out of 190; Phonology Test = percentile of the Test for Phonology.

The study protocol was approved by the Research Ethics Committee of the Northern Savo Hospital District, and the local Head of the Child Services gave permission for the study. Informed consent was provided for each participant by a parent.

Procedure

The children with SSD were mainly assessed at a private clinic of the first author, but eight participants were tested at their day care centres due to logistical reasons. The TD children were assessed at their day care centres, apart from one child who was tested at her home due to timetable issues. All data collection sessions were conducted in dyadic interactions between the examiner and a child, and video- and audio-recorded using a video recorder (Canon Legria HF G25) with a microphone (Sennheiser ME 4) and a digital voice recorder (Olympus LS-11, Linear PCM recorder).

To examine the accuracy and consistency of production, children were asked to name twenty pictures three times during a single assessment session. The target words were chosen to elicit a production of 2–5 syllabic words composed of sound and syllable structures frequently used by children in a given age range (Appendix 1). Only those items which a child named three times were included in the analysis. Of the children with SSD, 25 produced responses for all 20 target words, five children produced responses for 19, one child for 18 and one for 17 target words. Of the TD children, 26 produced responses for all 20 target words and six for 19 target words. The missing responses resulted from the children's naming problems. The response was considered incorrect if there was either a phonemic or a phonetic error (i.e. distortion, except aspiration) in the production of consonants or vowels. The participants' responses were categorized into one of four classes according to response type: consistently correct (when each of the three responses matched with the adult target), consistently incorrect (when all the three responses were similar but did not match the adult target), variable with hits (when variable responses included at least

one matched with the adult target), and variable with no hits (when the responses included at least two different response types without the matched adult target) (Grunwell, 1992). The proportion of each response type was calculated from the responses, which were included in the analysis. The ‘Variable with no hits’ rate, for example, was counted by dividing the number of variably incorrectly produced words by the total number of target words which a child named three times during the assessment and multiplying the result by 100. For example, if a child could name 18 pictures three times in the intra-word variability measure and out of these 18 target words, six were produced variably without the matched adult target, the ‘variable with no hits’ percentage was calculated as follows:

$(6 [= \text{number of variably produced target words}] / 18 [= \text{total number of target words included in the analysis}]) \times 100 = 33\%$.

Furthermore, both consistent response types (i.e. consistently correct and consistently incorrect) and variable response types (i.e. variable with hits and variable with no hits) were summed up to obtain overall word consistency and variability rates.

Spontaneous speech samples were used to calculate the percentage of consonants correct (hereafter sPCC) and the percentage of vowels correct (hereafter sPVC) (Shriberg, Austin, Lewis, McSweeney, & Wilson, 1997). The samples were elicited in conversation for an average of 15 minutes (SD 2.3 min) about a child’s favourite toys, games, and hobbies. The data was collected from the beginning of a recording until 90 word types were obtained (mean 90, range 86–90), such that differently inflected words were considered as different word types.

To analyse the association between vocabulary knowledge and intra-word accuracy and consistency, the raw scores of ROWPVT (Martin & Brownell, 2010)

were used to represent the level of receptive vocabulary. The test requires a child to point to one of four illustrations labelled by the examiner.

Analysis

Speech samples were narrowly transcribed by the same speech-language therapist using the video recordings, and the audio recording was utilized to confirm any transcription, if necessary. Both vowels and consonants were transcribed. Another experienced and certified speech-language therapist carried out reliability analysis from a randomly selected 10% of the samples of the Finnish test for Phonology, the spontaneous speech sample and the intra-word variable measure. Average inter-rater reliability was 88% (range 86–89) for consonants, and 94% (range 90–96) for vowels. Intra-rater reliability was assessed on a randomly selected 10% of the data and yielded 93% (range 86–95) reliability for consonants and 97% (range 93–100) for vowels. To examine the relationship between receptive vocabulary and intra-word accuracy and consistency, the raw scores of ROWPVT were used to represent size of the receptive vocabulary.

Statistical tests were computed with SPSS for windows (25.0, SPSS Inc.) and RStudio (version 1.1.463 by lme4 package version 1.1-21). Non-parametric statistical analysis for comparisons was used, as the data did not show normal distribution. Because of this, instead of mean and standard deviation, descriptive statistics with median and interquartile range (Q1–Q3) were reported to present rates of consistency and accuracy both in the SSD and the TD groups. To evaluate word accuracy and consistency between the study groups, the odds ratios (OR) with 95 % confidence interval (CI) were estimated by logistic mixed models having target words and individuals as mixed effects. Possible relationships between the child-specific variables, i.e. receptive vocabulary size, sPCC, sPVC and median percentage of

accuracy and consistency were studied with Spearman's correlation coefficients. Thereafter, as accuracy and consistency of speech has been reported to increase with age in TD children, the specificity of the relationship between articulatory ability, receptive vocabulary and measures of speech accuracy and consistency was investigated by age-adjusted Spearman's correlation (Conover, 1999).

Results

Intra-word accuracy and consistency

The descriptive statistics for the median and interquartile range (IQR; Q1–Q3) of each response type for the study groups are presented in Figure 1. The median percentage of overall word consistency for the TD children was 80% (IQR 66–85, range 47–100), and for the children with SSD 55% (IQR 40–73, range 15–90). The median percentages of overall word variability were 20% (IQR 15–34, range 0–53) and 45% (IQR 27–61, range 10–85), respectively. After controlling for the variation in the target words and in the individuals, the OR for the TD children to produce consistent responses more frequently was 3.1 (95% CI 2.0–4.8; $p < .001$). Of the consistent responses, the children with SSD more probably produced 'consistently incorrect' response types than the TD children (OR = 170; 95% CI 47–613; $p < .001$). They also more probably produced 'variable with no hits' response types (OR = 56; 95% CI 14–218; $p < .001$) when only variable responses were considered.

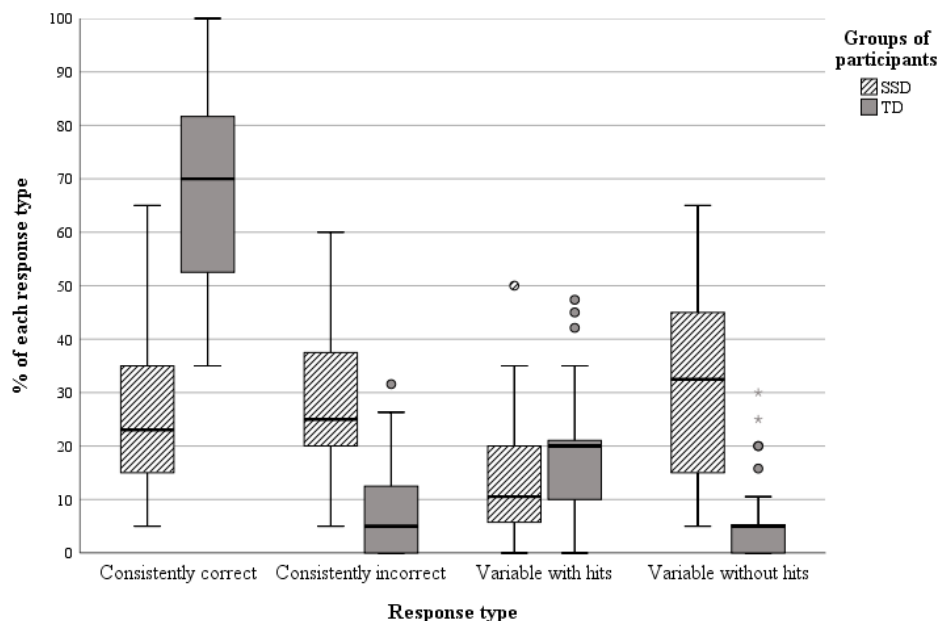


Figure 1. Median and interquartile range of percentage of each response type for the children with speech sound disorder (SSD) and typically developing children (TD), calculated from three repetitions of the target words.

Relationship between the child-specific characteristics and accuracy and consistency of speech

Correlations between receptive vocabulary, articulatory ability and different response types are presented by study group in Table 3. Both among the children with SSD and TD, receptive vocabulary skills were positively correlated with consistent correct responses and negatively correlated with variable responses without hits. Furthermore, in the TD children, receptive vocabulary knowledge was negatively correlated with consistently incorrect responses. Accuracy of consonants in spontaneous speech (sPCC) was correlated with all response types in both study groups, except for variable with hits responses in the TD children. Accuracy of vowels in spontaneous speech (sPVC) was positively correlated with ‘consistently correct’ and negatively with

‘variable with no hits’ responses in the SSD children.

Table 3. Spearman’s correlation coefficients between the child-specific characteristics (i.e. receptive vocabulary knowledge and articulatory ability) and accuracy and consistency of speech in children with SSD and TD.

	Consistently correct		Consistently incorrect		Variable with hits		Variable with no hits	
	SSD	TD	SSD	TD	SSD	TD	SSD	TD
RVT	0.61***	0.50**	-0.09	-0.47**	0.29	-0.23	-0.68***	-0.42*
sPCC	0.54**	0.76***	-0.39*	-0.82***	0.48**	-0.10	-0.48**	-0.64***
sPVC	0.70***	0.19	0.11	-0.04	-0.12	0.04	-0.53**	-0.08

SSD = children with speech sound disorder; TD = typically developing children; RVT = Receptive One-Word Picture Vocabulary Test; sPCC = spontaneous speech PCC; sPVC = spontaneous speech PVC.

* $p < .05$, ** $p < .01$, *** $p < .001$.

To explore whether these associations could be explained by the participants’ age, a partial analysis was conducted. As shown in Table 4, after removing age-related variance, the higher the receptive vocabulary scores a child with SSD had, the more ‘consistently correct’ and the less ‘variable with no hits’ responses s/he produced. In the TD children, there was no correlation between vocabulary scores and intra-word accuracy and consistency. sPCC was correlated with all response types in the children with SSD. Thus, more accurately produced consonants in spontaneous speech were related to more frequently produced accurate word forms and less frequently produced incorrect word forms in the speech variability measure. In the TD group, sPCC was positively correlated with ‘consistently correct’ responses and negatively with

‘consistently incorrect’ and ‘variable with no hits’ responses. In the children with SSD, more accurately produced vowels in spontaneous speech were correlated with more ‘consistently correct’ responses and less ‘variable with no hits responses’ in the variability measure. No other statistically significant relationships were observed.

Table 4. Spearman’s partial correlations, controlling for age of the participants, between the child-specific characteristics (i.e. receptive vocabulary knowledge and articulatory ability) and accuracy and consistency of speech in children with SSD and TD.

	Consistently correct		Consistently incorrect		Variable with hits		Variable with no hits	
	SSD	TD	SSD	TD	SSD	TD	SSD	TD
RVT	0.49**	0.04	-0.11	-0.21	0.09	0.03	-0.53**	-0.09
sPCC	0.48**	0.62***	-0.40*	-0.76***	0.41*	0.13	-0.39*	-0.50**
sPVC	0.62***	-0.11	0.13	0.18	-0.33	0.20	-0.40*	0.16

SSD = children with speech sound disorder; TD = typically developing children; RVT = Receptive One-Word Picture Vocabulary Test; sPCC = spontaneous speech PCC; sPVC = spontaneous speech PVC.

* $p < .05$, ** $p < .01$, *** $p < .001$.

Discussion

In this study, intra-word consistency and accuracy was studied among Finnish-speaking SSD and TD children, aged 3;2–5;4 years, by classifying their responses into

four categories: consistently correct, consistently incorrect, variable with hits and variable with no hits. Furthermore, overall word variability and the correlations between each of the response types, and the children's receptive vocabulary knowledge and articulatory ability were calculated.

Intra-word consistency and accuracy

In our study, among the TD children the proportion of consistently correctly produced words was high, as could have been expected from the earlier findings that at age 3 onwards both the consonant and vowel inventories are almost fully mastered by Finnish children (Kunnari et al., 2006; Kunnari et al., 2012; Savinainen-Makkonen et al., 2009; Saaristo-Helin, 2009). The children with SSD produced less consistently correct and more consistently incorrect responses. Concerning variable responses, the TD children produced more 'variable with hits' response types, whereas the children with SSD produced more 'variable with no hits' responses, which emphasizes the importance of separating the types of variable responses from each other. The median of overall variability scores among the children with TD increased to 20% due to rather frequently occurring variable responses with a matched adult target. It can thus be concluded that, within a given age range, variable responses including a correct word form still belong to typical speech development, indicating maturing phonetical/phonological skills. Variable responses without correct forms, on the contrary, are rare among TD children (5%) and are therefore important to take into consideration when assessing children's speech production skills.

When comparing the overall variability figures of the TD children to the earlier findings, it is important to notice that in the current study overall variability was not calculated in separate age groups but among children whose age ranged from 3;2 to

5;4 years. Furthermore, the variability score of our study (20%) was based on a narrow transcription, whereas in previous studies a broad transcription has been used. This may account for the lower median scores in the study of Martikainen et al. (2019), in which values ranged from 16% at age 3+ to 5% at age 5–6+. In Dutch, variability scores are higher than in Finnish, ranging from 56% at age 3+ to 36% at age 5+ (Faes & Gillis, 2018). In English, Holm et al. (2007) reported somewhat lower scores than in the present study but Sosa (2015) provided notably higher levels of variability for children aged 3+ (62%). Additional studies would clarify the picture of a typical developmental trajectory, but even from the findings obtained so far we can draw the conclusion that language-specific studies are needed in order to provide reliable knowledge for speech-language therapists to utilize when assessing children with SSD.

The median percentage of overall word variability among the SSD children was 45%, which is parallel with the study of Macrae et al. (2014). Thus, it seems that speech variability manifests itself similarly despite differences between the target languages. However, the remarkable difference between the studies relates to the employed transcription: Macrae et al. transcribed the consonants at a phonemic level, whereas we analyzed both the consonants and vowels with a narrow transcription. It is possible that the result of Macrae et al. would have been higher if distortions of speech sounds had been included in the analysis.

It is obvious that the participants' age range was one reason for the wide variation in the children's production, but among the children with SSD it may also reflect differently manifesting speech disorders. Iuzzini (2012) proposed that lexical level inconsistency scores (i.e. overall variability) based on a narrow transcription greater than 52% would point to a diagnosis of CAS and less than 48% to

phonological disorder. On the other hand, the diagnosis of inconsistent speech disorder has been founded on inconsistency scores of over 40% in the Word Inconsistency Assessment (based on a broad transcription) (Dodd et al., 2006). In the present study, the participants were not classified according to their speech features. Given overall variability ranging from 10–85% among the participants with SSD, we could, however, assume that some of the children may have consistent speech errors with less variation, as in children with phonological and/or articulation disorders (Dodd et al., 2005), and other children may have more variable productions, as in children with CAS (e.g. ASHA, 2007) or inconsistent speech disorder (Bradford & Dodd, 1996). Further study of lexical inconsistency with a larger sample of Finnish children with SSD is thus clearly warranted in order to be able to provide data sufficiently to discover whether distinct subgroups could be formed.

Relationships between consistency and accuracy of speech and the receptive vocabulary knowledge and articulatory abilities

Among the participants with TD, there was no relationship between receptive vocabulary knowledge and intra-word accuracy and consistency when the effect of age was removed, which is in line with the results of Martikainen et al. (2019) and Macrae and Sosa (2015). It has been presented that vocabulary knowledge reflects a state of underlying presentations of the words (Metsala & Walleys, 1998) which, in turn, is related to consistency of speech production (Macrae et al., 2014). Judging from that, the findings of the current study argue for the presence of well-defined lexical representations from the age of 3 onwards among typically developing Finnish children. It may be that in Finnish language, with long words and thus possible lower neighbourhood density i.e. numbers of phonologically similar words in the lexicon

(Stoel-Gammon, 2011), and simple syllable structures, lexical representations are robust enough for accurate and consistent speech production even at an earlier age than in English (see also Martikainen et al., 2019; Munson, Edwards, & Beckman, 2011).

In the SSD group, by contrast, the vocabulary knowledge was statistically significantly correlated between accuracy and consistency of speech. In other words, the wider receptive vocabulary a child had, the more consistently correct responses s/he produced and vice versa, the smaller receptive vocabulary a child had, the more variable responses without matched adult forms s/he produced. These findings are contrary to those of Holm et al. (2008) but parallel with those of Dodd et al. (2005) and Macrae et al. (2014), pointing to the hypothesis of the existence of underspecified and unstable representations for words, which may result in inconsistent speech production.

When age-related variance was removed, sPCC as a measure of the articulatory ability was significantly correlated with all response types in both the study groups except ‘variable with hits’ responses among the TD children. These findings indicate that those children who had more consonant errors in spontaneous speech also produced the words in the variability measure more inconsistently, without matched adult targets. To the best of our knowledge this is the first time this association has been reported among children with SSD. Concerning the relationship between the TD children’s articulatory ability and inconsistency of speech, the results are opposed to the studies of Macrae and Sosa (2015) and Martikainen et al. (2019), which may be explained by the level of transcription used. In both previous studies the data was analyzed at the phonemic level, whereas in the current study phonetic transcription was used. A motoric-based account for inconsistency of speech hypothesized that an

impairment in speech motor planning and/or programming is connected to the phenomenon (e.g. Grigos et al. 2015; Terband et al., 2011, Vuolo & Goffman, 2016). If this is the case, then using the phonemic level (broad) transcription, and thus excluding speech sound distortions from the analysis, may substantially affect the results, as distortions are proposed to be associated with impairments at the level of speech motor planning (Van der Merwe & Steyn, 2018).

Errors in vowel production separated the study groups, since sPVC was statistically significantly correlated with ‘consistently correct’ and ‘variable with no hits’ responses in the children with SSD but not in the TD children. The findings indicate that the more a child with SSD produced correct vowels, the more likely s/he was to produce consistently correct responses. On the other hand, if the child had a lot of vowel errors in spontaneous speech, s/he more probably produced variable responses without the adult matched target in the variability measure. This is parallel with the findings of Holm et al. (2008), who found that children with inconsistent speech disorder had significantly more vowel errors than children with delayed phonological development, children with consistent atypical speech errors and TD. Furthermore, some researchers have included vowel distortions in the diagnostic criteria of CAS, (e.g. Shriberg, Potter, & Strand, 2011). Thus, vowel errors occurring among the participants of our study might specifically be related to inconsistent speech production, demonstrating the importance of involving the vowels in the analysis.

To summarize, our findings showed that the group of children with SSD produced words highly inconsistently, without correct forms, which was related both to their receptive vocabulary knowledge and to their articulatory abilities. It may be that some of them primarily have a deficit in speech perception, which leads to unspecified phonological representations and emerges as inconsistent speech

production. Based on our study protocol, however, we cannot rule out the role of an impaired speech discrimination. No direct connection between impaired articulatory movements and phonemic/phonetic level errors has been demonstrated (Vuolo & Goffman, 2016), but our results derived from phonetically transcribed spontaneous speech samples did give some indication of the role of the children's articulatory abilities in speech consistency. A measure which would be similar for each of the participants and may reflect more directly the motor subsystems of speech, such as a diadochokinetic task, might highlight this connection even more clearly. Finally, as an indication of the variety of developmental systems, speech inconsistency may be accounted for by a pure speech motor impairment which is reflected indirectly in speech production by phonological representations, as Forrest et al. (1997) and Iuzzini (2012) suggested, or by a concurrent impairment in auditory processing and speech motor skills, as Terband et al. (2014) proposed to be case in children with CAS.

Conclusions

The findings of our study provide evidence that at an age when children are typically referred to an assessment due to an impairment in speech development, Finnish speech-language therapists should take into account whether a child produces highly variable speech without correct word forms. Furthermore, both the child's articulatory abilities and receptive vocabulary knowledge, and possibly also speech perception skills, should be investigated in order to provide suitable guidelines for the choice of effective intervention methods.

Declaration of Conflicting Interest

The authors report no conflicts of interest.

References

- American Speech-Language-Hearing Association. (2007). *Childhood apraxia of speech* [Position statement]. Retrieved from <http://www.asha.org/policy/PS2007-00277/>
- Bradford, A., & Dodd, B. (1996). Do all speech-disordered children have motor deficits? *Clinical Linguistics & Phonetics*, *10*, 77–101.
- Conover, W. J. (1999). *Practical nonparametric statistics* (3rd ed.). New York, NY: John Wiley & Sons.
- Demuth, K. (2011). Interactions between lexical and phonological development: Cross-linguistic and contextual considerations – a commentary on Stoel-Gammon's 'Relationships between lexical and phonological development in young children'. *Journal of Child Language*, *38*, 69–74.
- Dodd, B., Holm, A., Crosbie, S., & McCormack, P. (2005). Differential diagnosis of phonological disorders. In B. Dodd (Ed.), *Differential diagnosis and treatment of children with speech disorders* (pp. 44–70). London: Whurr.
- Dodd, B., Hua, Z., Crosbie, S., Holm, A., & Ozanne, A. (2006). *Diagnostic Evaluation of Articulation and Phonology*. San Antonio, TX: Pearson.
- Dodd, B., & Leahy, J. (1989). Phonological disorders and mental handicap. In M. Beveridge, G. Conti-Ramsden, & I. Leudar (Eds.), *Language communication in mentally handicapped children* (pp. 33–56). London: Chapman & Hall.
- Faes, J., & Gillis, S. (2018). Intraword variability in children with cochlear implants: The long-term development up to 5 years of age and a comparison with children with normal hearing. *American Journal of Speech-Language Pathology*, *27*, 706–720.

- Forrest, K., Dinnsen, D. A., & Elbert, M. (1997). Impact of substitution patterns on phonological learning by misarticulating children. *Clinical Linguistics & Phonetics, 11*, 63–76.
- Forrest, K. (2003). Diagnostic criteria of developmental apraxia of speech used by clinical speech-language pathologists. *American Journal of Speech-Language Pathology, 12*, 376–380.
- Forrest, K., Elbert, M., & Dinnsen, D. A. (2000). The effect of substitution patterns on phonological treatment outcomes. *Clinical Linguistics & Phonetics, 14*, 519–531.
- Froud, K., & Khamis-Dakwar, R. (2012). Mismatch negativity responses in children with a diagnosis of childhood apraxia of speech (CAS). *American Journal of Speech-Language Pathology, 21*, 302–312.
- Grigos, M. I., Moss, A., & Lu, Y. (2015). Oral articulatory control in childhood apraxia of speech. *Journal of Speech, Language, & Hearing Research, 58*, 1103–1118.
- Grunwell, P. (1992). Assessment of child phonology in the clinical context. In C. A. Ferguson, L. Menn, & C. Stoel-Gammon (Eds.), *Phonological development: Models, research, implications* (pp. 457–483). Timonium, MD: York Press.
- Hearnshaw, S., Baker, E., & Munro, N. (2018). The speech perception skills of children with and without speech sound disorder. *Journal of Communication Disorders, 71*, 61–71.
- Holm, A., Crosbie, S., & Dodd, B. (2007). Differentiating normal variability from inconsistency in children's speech: Normative data. *International Journal of Language & Communication Disorders, 42*, 467–486.

- Holm, A., Farrier, F., & Dodd, B. (2008). Phonological awareness, reading accuracy and spelling ability of children with inconsistent phonological disorder. *International Journal of Language & Communication Disorders, 43*, 300–322.
- Iuzzini, J. (2012). *Inconsistency of speech in children with childhood apraxia of speech, phonological disorders, and typical speech*. (Unpublished doctoral dissertation). Indiana University, Bloomington. Retrieved from <https://www.researchgate.net/publication/258693515>
- Iuzzini-Seigel, J., Hogan, T. P., & Green, J. R. (2017). Speech inconsistency in children with childhood apraxia of speech, language impairment, and speech delay: Depends on the stimuli. *Journal of Speech, Language and Hearing Research, 60*, 1–17.
- Karlsson, F. (1983). *Suomen kielen äänne- ja muotorakenne* [Finnish phonology and morphology]. Porvoo, Helsinki, Juva: WSOY.
- Kortesmaa, M., Heimonen, K., Merikoski, H., Warma, M-L., & Varpela, V. (2001). *Reynellin kielellisen kehityksen testi* [Reynell developmental language scales]. Helsinki: Psykologien Kustannus.
- Kunnari, S. (2002). Word length in syllables: Evidence from early word production in Finnish. *First Language, 22*, 119–135.
- Kunnari, S., Savinainen-Makkonen, T., & Paavola, L. (2006). Kaksivuotiaiden suomalaislasten konsonantti-inventaarit [Consonant inventories of two-year-old Finnish-speaking children]. *Puhe ja kieli, 26*, 71–79.
- Kunnari, S., Savinainen-Makkonen, T., & Saaristo-Helin, K. (2012). *Fonologiatesti. Lasten äännteellisen kehityksen arviointimenetelmä* [Test for Phonology]. Jyväskylä: Niilo Mäki Instituutti.

- Maassen, B., Groenen, P., & Crul, T. (2003). Auditory and phonetic perception of vowels in children with apraxic speech disorders. *Clinical Linguistics & Phonetics*, *17*, 447–467.
- Macrae, T. (2013). Lexical and child-related factors in word variability and accuracy in infants. *Clinical Linguistics & Phonetics*, *27*, 497–507.
- Macrae, T., & Sosa, A. V. (2015). Predictors of token-to-token inconsistency in preschool children with typical speech-language development. *Clinical Linguistics & Phonetics*, *29*, 922–937.
- Macrae, T., Tyler, A. A., & Lewis, K. E. (2014). Lexical and phonological variability in preschool children with speech sound disorder. *American Journal of Speech-Language Pathology*, *23*, 27–35.
- Malmenholt, A., Lohmander, A., & McAllister, A. (2017). Childhood apraxia of speech: A survey of praxis and typical speech characteristics. *Logopedics Phoniatrics Vocology*, *42*, 84–92.
- Marquardt, T. P., Jacks, A., & Davis, B. L. (2004). Token-to-token variability in developmental apraxia of speech: Three longitudinal case studies. *Clinical Linguistics & Phonetics*, *18*, 127–144.
- Martikainen, A., Savinainen-Makkonen, T., & Kunnari, S. (2019). Intra-word consistency and accuracy in Finnish children aged 3–6 years. *Clinical Linguistics & Phonetics*, *33*, 815–830.
- Martin, N. A., & Brownell, R. (2010b). *Receptive One-Word Picture Vocabulary Test – 4*, Finnish version: S. Kunnari & T. Välimaa; In validation. Novato, CA: Academic Therapy Publications.
- Metsala, J. L., & Walleys, A. C. (1998). Spoken vocabulary growth and the segmental restructuring of lexical representations: Precursors to phonemic awareness and

- early reading ability. In J. L. Metsala & L. C. Ehri (Eds.), *Word recognition in beginning literacy* (pp. 89–120). Mahwah, NJ: Erlbaum.
- Munson, B., Edwards, J., & Beckman, M. E. (2005). Relationships between nonword repetition accuracy and other measures of linguistic development in children with phonological disorders. *Journal of Speech, Language, & Hearing Research, 48*, 61–78.
- Munson, B., Edwards, J., & Beckman, M. E. (2011). Phonological representations in language acquisition: Climbing the ladder of abstraction. In A.C. Cohn, C. Fougerson, & M. K. Huffman (Eds.), *The handbook of laboratory phonology* (pp. 288–309). Oxford: University Press.
- Nijland, L. (2009). Speech perception in children with speech output disorders. *Clinical Linguistics & Phonetics, 23*, 222–239.
- Nijland, L., Maassen, B., & van der Meulen, S. (2003). Evidence of motor programming deficits in children diagnosed with DAS. *Journal of Speech, Language, and Hearing Research, 46*, 437–450.
- Nijland, L., Maassen, B., van der Meulen, S., Gabreels, F., Kraaimaat, F. W., & Schreuder, R. (2003). Planning of syllables in children with developmental apraxia of speech. *Clinical Linguistics & Phonetics, 17*, 1–24.
- Richtsmeier, P. T., Gerken, L., Goffman, L., & Hogan, T. (2009). Statistical frequency in perception affects children's lexical production. *Cognition, 111*, 372–377.
- Roid, G. H., Miller, L. J., Pomplun, M., & Koch, C. (2013). *Leiter International Performance Scale - Third Edition*. Wood Dale, IL: Stoelting.
- Rvachew, S., & Jamieson, D., G. (1989). Perception of voiceless fricatives by children with a functional articulation disorder. *Journal of Speech and Hearing Disorders, 54*, 193–208.

- Saaristo-Helin, K. (2009). Measuring phonological development: A follow-up study of five children acquiring Finnish. *Language and Speech*, *52*, 55–77.
- Savinainen-Makkonen, T. (2000). Learning long words - A typological perspective. *Language & Speech*, *43*, 205–225.
- Savinainen-Makkonen, T., Kaikkonen, H., Saaristo-Helin, K., & Kovasiipi-Nieminen, P. (2009). 2½-vuotiaiden konsonanttiyhtymät [Medial consonant clusters in 2½-year-old children]. *Puhe ja kieli*, *29*, 165–181.
- Sayyahi, F., Soleymani, Z., Akbari, M., Bijankhan, M., & Dolatshahi, B. (2017). Effect of gap detection threshold on consistency of speech in children with speech sound disorder. *Research in Developmental Disabilities*, *61*, 151–157.
- Shriberg, L. D. (2004). Childhood speech sound disorders: From postbehaviorism to the postgenomic era. In R. Paul & P. Flipsen, Jr. (Eds.), *Speech sound disorders in children: in honor of Lawrence D. Shriberg* (pp. 1–33). San Diego, CA: Plural Publishing.
- Shriberg, L. D., Austin, D., Lewis, B. A., McSweeney J. L., & Wilson D. L. (1997). The percentage of consonants correct (PCC) metric: Extensions and reliability data. *Journal of Speech, Language, & Hearing Research*, *40*, 708–722.
- Shriberg, L. D., Potter, N. L., & Strand, E. A. (2011). Prevalence and phenotype of childhood apraxia of speech in youth with galactosemia. *Journal of Speech, Language, and Hearing Research*, *54*, 487–519.
- Sosa, A. V. (2015). Intraword variability in typical speech development. *American Journal of Speech-Language Pathology*, *24*, 24–35.
- Stoel-Gammon, C. (2011). Relationships between lexical and phonological development in young children. *Journal of Child Language*, *38*, 1–34.

- Swingle, D. (2005). 11-month-olds' knowledge of how familiar words sound. *Developmental Science*, 8, 432–443.
- Terband, H., Maassen, B., Guenther, F. H., & Brumberg, J. (2014). Auditory–motor interactions in pediatric motor speech disorders: Neurocomputational modeling of disordered development. *Journal of Communication Disorders*, 47, 17–33.
- Terband, H., Maassen, B., van Lieshout, P., & Nijland, L. (2011). Stability and composition of functional synergies for speech movements in children with developmental speech disorders. *Journal of Communication Disorders*, 44, 59–74.
- Van der Merwe, A., & Steyn, M. (2018). Model-driven treatment of childhood apraxia of speech: Positive effects of the speech motor learning approach. *American Journal of Speech-Language Pathology*, 27, 37–51.
- Vuolo, J., & Goffman, L. (2016). An exploratory study of the influence of load and practice on segmental and articulatory variability in children with speech sound disorders. *Clinical Linguistics & Phonetics*, 31, 331–350.
- Whiteside, S. P., Dobbin, R., & Henry, L. (2003). Patterns of variability in voice onset time: a developmental study of motor speech skills in humans. *Neuroscience Letters*, 347, 29–32.

Appendix 1. Item overview of the intra-word variability measure

/mu:mi/	Moomin character
/vaip:a/	napkin
/poika/	boy
/housut/	trousers
/sam:ak:o/	frog
/ju:sto/	cheese
/pulk:a/	sled
/sakset/	scissors
/mansik:a/	strawberry
/kuk:a/	flower
/tietokone/	computer
/nap:i/	button
/muna/	egg
/orava/	squirrel
/lentokone/	airplane
/ham:astahna/	toothpaste
/kilpikon:a/	turtle
/helikopteri/	helicopter
/itke:/	to cry
/vetoketju/	zip