

# **Intra-word consistency and accuracy in Finnish children aged 3–6 years**

**Anna-Leena Martikainen<sup>1</sup>, Tuula Savinainen-Makkonen<sup>1</sup>, Sari Kunnari<sup>1</sup>**

<sup>1</sup>University of Oulu, Finland

## **INTRA-WORD CONSISTENCY AND ACCURACY IN FINNISH CHILDREN AGED 3–6 YEARS**

### **Abstract**

This study examined intra-word consistency and accuracy in typically developing Finnish children and their relation to children's vocabulary size and phonological skills. A total of 80 typically developing Finnish children aged 3;0 to 6;11 were asked to name 20 words three separate times during a single assessment session. Responses were classified into four categories: 1) consistently correct productions, 2) consistently incorrect productions, 3) variable productions with hits (variable productions including at least one matched adult target), and 4) variable productions with no hits. The results revealed that 5- and 6-year-old children produced significantly more often consistently correct responses than younger children. However, even for the 3- and 4- year old children the most frequent response type was consistently correct production. Between these two youngest age groups (3 and 4), the only significant difference was in consistently incorrect responses, which the 3-year-olds produced more often than the older children. There was a significant negative correlation between consistently incorrectly produced words and children's phonological skills, but no other relationships were found. The results indicate that when assessing children with speech sound disorder (SSD), Finnish clinicians need to take into account the fact that even 3-year-old typically developing children generally produce words correctly, either consistently or inconsistently.

### **Keywords**

child, intra-word consistency, accuracy, typical speech development, Finnish

## Introduction

Speech variability is a well-documented phenomenon in typically developing young children's speech (Ferguson & Farwell, 1975; Macrae, 2013; Sosa, 2015; Sosa & Stoel-Gammon, 2006, 2012). It can be identified at the phonemic, i.e. segmental level, when a child pronounces the same phoneme multiple ways based on word position (Betz & Stoel-Gammon, 2005; Shriberg, Aram, & Kwiatkowski, 1997). For example, /s/ may be produced accurately in word-final position (as in *pois* [pɔ̃is], 'away') but may be substituted by /t/ in word-initial position (as in *suu* [t̪u:], 'mouth'). Variability of speech can also be observed at the whole-word level (Dodd, Holm, Crosbie, & McCormack, 2005; Grunwell, 1992; Holm, Crosbie, & Dodd, 2007), when a child produces a word differently within a short time period, e.g. *ham:astahna*, 'toothpaste' as [pɑ̃m:ɑ̃t:ɑ̃t:ɑ̃], [vɑ̃m:ɑ̃t:ɑ̃t:ɑ̃] and [hɑ̃n:ɑ̃(..)t̪ɑ̃:ɑ̃]. This kind of lexical level consistency and accuracy has not been previously studied in Finnish and is thus the focus of the current study.

Since the heterogeneity among children with speech sound disorder (SSD) has been increasingly recognized, several classification systems and sub-types with characterizing speech features have been proposed (e.g. Dodd, 1995; Shriberg, 2004; Stackhouse & Wells, 1997). The segmental and the lexical level variability has been reported in children with phonological/articulation disorders (Macrae, Tyler, & Lewis, 2014), childhood apraxia of speech (CAS, e.g. American Speech-Language-Hearing Association [ASHA], 2007), inconsistent speech disorder (ISD, Dodd et al., 2005), and cochlear implants (Faes & Gillis, 2018). Some researchers have presented that high variability in speech production would be a diagnostic indicator especially in CAS (ASHA, 2007) and ISD (Dodd et al., 2005), while others have questioned the role of intra-word variability in the differential diagnosis (Macrae & Sosa, 2015; Sosa, 2015). Furthermore, it has been presented that variability of speech could be a signal of pervasive speech-processing difficulties (Williams & Stackhouse, 2000) and may influence the outcome of treatment (Forrest, Dinnsen, & Elbert, 1997; Forrest & Elbert, 2001; Forrest, Elbert, & Dinnsen, 2000; but see also Tyler, Williams, & Lewis, 2006). Therefore, a clinician's thorough

knowledge of a typical progression of speech consistency and accuracy can have a critical role in diagnosing, timing and planning interventions, and predicting a prognosis of speech disorders.

After the period of variable speech productions in toddler ages (Ferguson & Farwell, 1975; Macrae, 2013; Sosa, 2015; Sosa & Stoel-Gammon, 2006, 2012), the consistency of speech is reported increasing with age among English-speaking children (Burt, Holm, & Dodd, 1999; de Castro & Wertzner, 2011; Holm et al., 2007; Kenney & Prather, 1986; Macrae, 2013; Sosa, 2015; Williams & Stackhouse, 2000). In the study of Holm et al. (2007), children aged from 3;0 to 6;11 named 25 pictures three times during one assessment session. Their responses were classified into four categories: 1) *consistently correct productions*, 2) *consistently incorrect productions*, 3) *variable productions with hits* (variable productions including at least one matched adult target), and 4) *variable productions with no hits* (Grunwell, 1992). In all age groups, the children produced more consistently correct responses than any other response types. The mean variability score decreased with age, being 12.5% for 3-year-olds, 6% for 4-year-olds, less than 4% for 5-year-olds and 1% for 6-year-old children. Sosa (2015), who used a similar study procedure in her research of children aged from 2;6 to 3;11, found that although variability decreased with age, average variability for 3-year-olds was still high. Among 3-year-olds, the most frequent response pattern was variable productions without hits (37%), followed by variable productions with hits (24%), consistently correct productions (24%), and finally consistently incorrect productions (13%). Macrae and Sosa (2015) found in their study that age (from 2;6 to 4;2) correlated significantly with intra-word inconsistency but did not explain variance of variability after expressive vocabulary was taken into consideration. The authors concluded that a higher level of speech processing (e.g. underlying representations for words) may be involved in inconsistency of speech.

From a motoric framework, variability of action has been hypothesized to play an essential role in the development of movement control (Hadders-Algra 2018; Thelen & Smith, 1994). It has been argued that variability can be a sign of a flexible neural system which enables learning of new behaviours, such

as producing new words (Smith, 2006; Smith & Goffman, 1998). It is well documented that articulator movements even in error-free speech, are slower and temporally more variable in children than in adults (e.g. Green, Moore, & Reilly, 2002; Nittrouer, 1993; Smith & Goffman, 1998). Implicitly, it could be assumed that highly variable speech motor functions may underlie the variation in consistency of speech productions. However, Goffman, Gerken, and Lucchesi (2007) found no association between segmental accuracy or variability and kinematic measures in typically developed children aged 4;2 to 4;10 and deduced that motor variability does not necessarily result in phonemic-level errors.

The findings of studies on the connection between segmental (articulatory/phonological) level measures and intra-word variability have been conflicting. Macrae and Sosa (2015) found no association between children's articulatory ability and intra-word consistency in children aged 2;6 to 4;2. On the other hand, phonetic and phonological complexity of a word have been reported to influence intra-word variability and accuracy, i.e. longer words with later-acquired phonemes and complex syllable structures may be difficult to say, and thus more variable (Leonard, Rowan, Morris, & Fey, 1982; Macrae, 2013; Sosa, 2015; Sosa & Stoel-Gammon, 2012).

Variability in speech has also been shown to vary as a function of vocabulary knowledge. A number of studies on the relationship between phonological/phonetic and lexical development have demonstrated that developmental gains in speech production and expressive communication function as activators for the acquisition of new speech sounds and phonological accuracy in early words (e.g. Davis, Van der Feest, & Yi, 2018; Kunnari, Savinainen-Makkonen, & Paavola, 2006; McCune & Vihman, 2001; Smith, McGregor, & Demille, 2006; see Stoel-Gammon, 2011 for review). One proposed explanation for this has been that underlining representations of words, which are presented to be holistic in nature in young children, become phonologically more fine-grained when the child's vocabulary grows (Edwards, Beckman, & Munson, 2004; Metsala & Walley, 1998). Thus, as Smith et al. (2006) stated "it is possible that children with larger lexicons are in some sense aware of and/or have had more experience producing

(or attempting to produce) a greater variety of segments, syllables, and word shapes. This increased exposure/experience might enhance their speech output capabilities...” (p. 370). Accordingly, it is supposed that young children and children with small vocabularies may have higher speech inconsistency due to incomplete phonological representations for words (Macrae, 2013; Macrae & Sosa, 2015; Sosa & Stoel-Gammon, 2012), although, no linear correlation has also been discovered (Sosa & Stoel-Gammon, 2006; Zanolini, Viterbori, & Saraceno, 2012). The view of holistic representations of words has been questioned (e.g. Swingley, 2005; Swingley & Aslin, 2000), however, and it has been argued that even young children’s phonological representations for familiar words can be more robust than is evidenced from perceptual analysis of production (Demuth & Song, 2012).

### ***Finnish phonological characteristics***

Since differences in language typologies may also have an effect on intra-word consistency and accuracy, it is necessary to consider some essential aspects of Finnish phonology. The Finnish phoneme inventory contains 13 consonants /p, t, d, k, m, n, ŋ, s, h, v, j, l, r/ and 8 vowels /ɑ, e, i, o, u, y, æ, œ/. In addition, the consonants /b, g, f, ʃ/ may occur in loanwords. The Finnish fricative system is simple, and in endogenous words there is voiced/voiceless opposition only between /t/ and /d/ (Karlsson, 1983). The phoneme /d/ is acquired late, probably due to its low frequency in Finnish (Toivainen, 1990). Motorically the most challenging sound to produce for Finns is an apico-alveolar trill /r/. All vowels and consonants, except for /d, h, j, v/ can be realised as short or long in the medial position of a word. In addition, vowels are used as a geminate in the initial and the final position of syllables. The quantity of a phoneme changes the meaning of the word, for example: [tuli] ‘fire’, [tuuli] ‘wind’, [tulli] ‘customs’. Word-initial and word-final consonant clusters are rare in Finnish words, but heterosyllabic clusters are common.

The most frequent syllable type in Finnish is CV (Häkkinen, 1978). Although the majority of words are bi-syllabic, long words with three or more syllables are also common (Karlsson, 1983). The number of monosyllabic words is small, only 0,1% according to the dictionary analysis (Karlsson, 1983, p. 217).

## **Current study**

One aim of the current study was to analyse intra-word consistency and accuracy in typically developing Finnish children aged 3;0 to 6;11 years. As this feature of speech production has not previously been studied in Finnish, a wide age range of participants was warranted to observe a typical trajectory of the phenomenon. The youngest children of the study were 3-year-olds, as children are often referred to a speech therapy assessment from that age onwards (Morgan et al., 2017), if their speech is unintelligible to family members or other communication partners. The second aim was to assess how phonological skills other than vocabulary knowledge are related to speech consistency and accuracy. Earlier studies have found a correlation between intra-word variability and word-level phonological and lexical factors, but children's phonological skills have been studied less. Furthermore, to our knowledge, research on intra-word consistency/variability and accuracy in typically developing children has been English-centric in nature. To deepen our understanding of the phenomenon, more research on typologically different languages is needed. Finnish is an ideal language for this purpose, as compared to English its sound inventory is small, words are long (bi- or multisyllabic), open syllables are typical, word-initial and word-final clusters are rare, and differences in phoneme quantity indicate differences in meaning.

## **Method**

### ***Participants***

A total of 80 children between the ages of 3;0 and 6;11 were recruited for the study. Recruitment from normal daycare centre populations continued until data were obtained from 10 boys and 10 girls in each

age group (four in all). Based on parental questionnaires and reports of the teachers of the daycare centres, the children's overall development had proceeded typically. All parents reported that they had no concerns about their child's hearing and the child had passed the annual screening test of hearing at the child's health clinic. The children had no reported neurological diseases. They were monolingual speakers of Finnish and exhibited normal speech development, as confirmed by a result of the Finnish Test for Phonology (Kunnari, Savinainen-Makkonen, & Saaristo-Helin, 2012) less than 1 SD below the mean (> 16 percentile). The Finnish Test for Phonology is a single-word picture naming task with normative data for children between aged 2;0 and 6;11. The total score of the test (maximum 127 points) is a composite of three sub-test scores measuring children's phonotactic skills: phoneme and syllable length, word length and phoneme combination. The normal vocabulary skills were confirmed by scores of less than 1 SD below the mean on the Expressive One-Word Picture Vocabulary Test – 4 (EOWPVT, Martin & Brownell, 2010a) and on the Receptive One-Word picture Vocabulary Test – 4 (ROWPVT, Martin & Brownell, 2010b). Typical speech understanding was confirmed by scores of less than 1 SD below the average according to the Finnish standardised version of the Reynell Developmental Language Scales III (Edwards et al., 1997). In addition, oral-mechanism examination revealed normal oral structure and oral-motor function for all children. The descriptive statistics for participants and summary data for inclusion criteria are presented in table 1.



**Table 1.** Participant mean demographic and summary data for inclusion criteria.

Age group	n	Mean age	RDLs III	ROWPVT	EOWPVT	Phonology test
3+ years						
all	20	3;5 (0;3)	109.7 (4.8)	56.4 (11.3)	50.7 (6.6)	110.1 (10.8)
male	10	3;6 (0;3)	109.3 (5.5)	59.1 (12.0)	50.9 (8.1)	109.0 (11.1)
female	10	3;5 (0;3)	110.0 (4.1)	53.6 (10.6)	50.4 (5.0)	111.1 (11.1)
4+ years						
all	20	4;5 (0;3)	102.5 (8.9)	81.7 (13.5)	66.4 (9.9)	120.8 (7.1)
male	10	4;6 (0;4)	100.9 (9.1)	84.8 (12.4)	69.6 (12.0)	123.2 (4.8)
female	10	4;4 (0;2)	104.1 (9.0)	78.6 (14.4)	63.1 (6.4)	118.4 (8.3)
5+ years						
all	20	5;4 (0;3)	102.7 (10.2)	108.8 (24.4)	81.1 (10.7)	124.7 (2.6)
male	10	5;4 (0;3)	103.8 (10.6)	104.8 (24.2)	81.7 (12.2)	125.3 (1.8)
female	10	5;4 (0;3)	101.6 (10.1)	112.8 (25.2)	80.5 (9.6)	124.1 (3.1)
6+ years						
all	20	6;6 (0;3)	108.2 (11.0)	138.1 (18.4)	90.8 (10.9)	125.3 (1.3)
male	10	6;5 (0;4)	106.5 (13.3)	132.2 (16.0)	89.3 (11.2)	125.1 (1.7)
female	10	6;6 (0;3)	110.1 (8.2)	143.9 (18.7)	92.2 (11.0)	125.5 (0.7)

Ages are given as years; months. Standard deviations are in parentheses. RDLs III = The Reynell Developmental Language Scales III, Comprehension Scales, standard score (M = 100, SD = 15); ROWPVT = Receptive One-Word Picture Vocabulary Test – 4, raw score out of 190; EOWPVT = Expressive One-Word Picture Vocabulary Test – 4, raw score out of 190; Phonology Test = The Test for Phonology, raw score out of 127.

The study plan was approved by the Research Ethics Committee of Northern Savo Hospital District. Permission for the study was also given by the local head of Child Services. All parents signed an informed consent.

### *Procedure*

At local children's daycare centres which agreed to participate in the research, the staff was informed of the inclusion criteria of the study and asked to distribute the information letter to the parents of children they assumed to fulfill the criteria. Parents who gave permission for their child to take part in the study returned the informed consent to the researcher by mail or via the staff of the daycare centre. Each child participated in three to four assessment sessions. The children were tested at their own daycare centre, except for one who was tested at her home due to timetable issues. All sessions were audio- and video-recorded using a high-quality digital voice recorder (Olympus LS-11, Linear PCM recorder) and a digital video recorder (Canon Legria HF G25) with a microphone (Sennheiser ME 4) at a distance of approximately 30 centimeters from the mouth. All testing was carried out by the same certified experienced speech and language therapist (the first author).

Consistency and accuracy of production was studied with twenty 2–5 syllabic words (see Appendix 1). Seventeen words selected from the Test for Phonology (Kunnari et al., 2012) were supplemented with three extra words. Target words with /d/ were not included due to their marginal role in the Finnish consonant paradigm. In addition, monosyllabic words were excluded, as they are rare in Finnish and presumably too insensitive to measure speaking abilities in children of 3–6-years, as Finnish children are reported to be capable of producing disyllabic words even during the period of the first 50 words (e.g. Kunnari, 2000, 2002; Saaristo-Helin, 2009). Children were asked to name the target words three separate times during the same assessment session.

Phonological skills were studied from spontaneous speech samples for about 15 minutes obtained in conversation between a child and the examiner about child's hobbies and favourite toys and games. A target set of words was 90 non-questionable word types. Differently inflected words were regarded as being different word types.

## *Analysis*

The broad transcriptions of speech samples were made by the first author from the video recordings and confirmed from audio records if necessary. To determine inter-judge reliability, a randomly selected 10% of the samples of the Finnish test for Phonology, the spontaneous speech and the intra-word variability measure were independently re-transcribed by another speech and language therapist who was an experienced transcriber. Phoneme-by-phoneme interrater reliability for consonants ranged from 93% to 98%, with a mean of 96%, and for vowels from 97% to 100%, with a mean of 98%. Intrarater reliability was assessed by re-transcribing randomly selected 10% of the speech samples after a period of several months. The intrarater reliability for consonants ranged from 94% to 99%, with a mean of 97%, and for vowels from 98% to 100%, with a mean of 99%.

To test intra-word consistency and accuracy, only those words for which participants provided a spontaneous response three times were included in the analysis. Of the 80 children, 63 produced responses for all 20 target words, 15 children produced responses for 19 and two children for 18 target words. These missing responses were due to children's naming problems. Responses were encoded as variable if there were any phonemic differences in the production of consonants or vowels. Consistent responses were classified (Grunwell, 1992) as 1) consistently correct (all three productions were the same and matched the adult target), and 2) consistently incorrect (all three productions were the same but did not match the adult target). Variable productions were classified into two categories according to the response type: 1) variable with hits (variable productions including at least one matched adult target), and 2) variable with no hits (variable productions, none of them matching the adult target). The proportion of each response type was calculated from the responses, which were included in the analysis. Consistently correct rates, for example, were calculated by dividing the number of consistently correctly produced words by the number of all words produced three times, and multiplying the result by 100.

Phonological skills were estimated from spontaneous speech samples by using a percentage of consonants correct-revised (hereafter sPCC-R) and a percentage of vowels correct-revised (hereafter sPVC-R) (Shriberg & Kwiatkowski, 1982; Shriberg & Austin, 1997), where distorted consonants and vowels were scored as correct. Analysis were prepared from the first 90 word types (mean 89.7, range 80–90) of each sample.

To analyse the relationship between vocabulary knowledge and intra-word variability, two measures of vocabulary were used. The raw scores of EOWPVT were used to represent the size of expressive vocabulary, and ROWPVT scores were used to represent the size of receptive vocabulary.

All statistical tests were conducted with SPSS for windows (25.0, SPSS Inc.). A Shapiro-Wilk test was used to test the normality assumption for each variable, and it appeared that not all variables evidenced normal distribution. Thus, descriptive statistics with median and interquartile range (Q1–Q3) were reported to exhibit rates of accuracy and consistency within each age group. A non-parametric Kruskal-Wallis test with a Dunn's test as a post hoc test were used to test pairwise comparisons of word consistency between and within the age groups. There were no differences in the results of the 5- and 6-year-olds, and therefore a combined group of 5- and 6-year-olds was used in the comparisons with the younger age groups. Furthermore, no differences were found between gender in each age group, so no further analyses were conducted for this factor. Possible correlations between child-specific characteristics, including vocabulary size and sPCC-R and sPVC-R reflecting a child's phonological skills, and median percent of accuracy and consistency were studied with Spearman's rho correlation coefficients. Thereafter, as the participants' age ranged from 3;0 to 6;11 and the consistency of speech productions has been reported to increase with age, a partial correlation controlling for the age of the children was conducted in order to examine the size of the unique portion of variance of vocabulary knowledge and phonological skills in intra-word consistency and accuracy.

## Results

### *Intra-word consistency and accuracy*

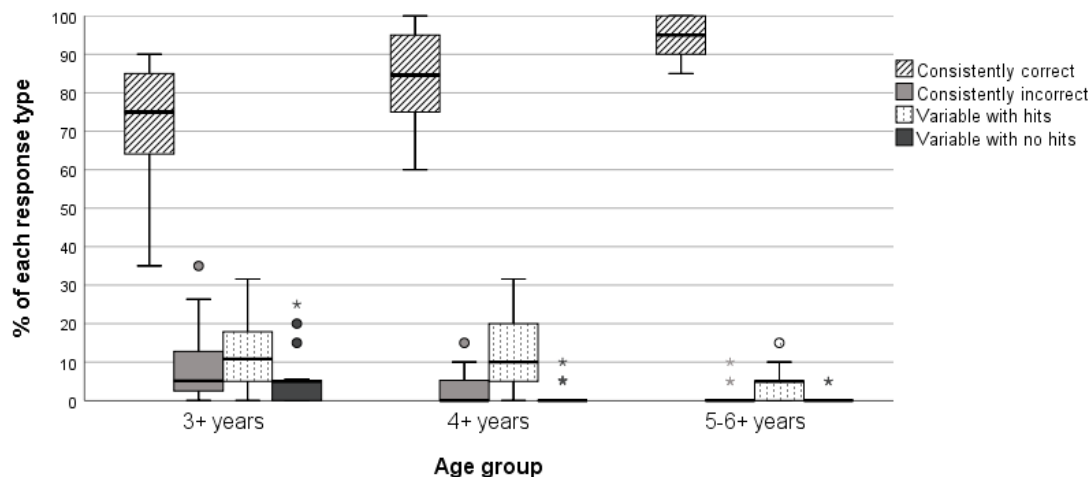


Figure 1. Median and interquartile range of percentage of consistently correct, consistently incorrect, variable with hits, and variable with no hits response types for each age group. The points and the asterisks represent outliers of the data.

In all age groups, the most frequent response type was ‘consistently correct’: the median was 75% among children aged 3 + years and it increased with age to 95% (figure 1). The median of ‘consistently incorrect’ response types was about 5% in the youngest age group, and 0% among older children. Both response types appeared to be statistically significantly affected by the age of the child ( $H(2) = 37.31, p < 0.001$ ;  $H(2) = 29.90, p < 0.001$ , respectively). Post-hoc analysis revealed that the oldest children significantly more often produced consistently correct responses than the younger children (children aged 3 + and 5–6 + years,  $p < 0.001$ ; 4+ and 5–6 + years,  $p = 0.001$ ). There was no difference between two youngest age groups. When considering consistently incorrect responses, children aged 3+ years produced them statistically significantly more often than the older children (children aged 4 + years,  $p < 0.05$ ; 5–6 + years,  $p < 0.001$ ). There was no difference between children aged 4 + and 5–6 + years.

The median for the ‘variable with hits’ response type was almost the same in the two youngest age groups, about 10%, and it decreased to 5% for older children. The ‘variable with no hits’ rate was 5% among children aged 3 + years, whereas older children produced them seldom, if at all. As ‘variable with no hits’ responses displayed floor effects basically in all the age groups, no statistical analyses were conducted. ‘Variable with hits’ response types were significantly affected by age ( $H(2) = 18.17, p < 0.001$ ). There were no differences in ‘variable with hits’ responses between children aged 3 + and 4 + years, but the oldest children produced them significantly less frequently than the younger ones (3 + and 5–6 + years,  $p = 0.001$ ; 4 + and 5–6 + years,  $p = 0.003$ ).

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

The results of correlational analyses between the child-specific characteristics (i.e. age, vocabulary knowledge and phonological skills) and accuracy and consistency of speech are presented in table 2. There were positive relationships between consistently correct responses and a child’s vocabulary size and sPCC-R. Furthermore, with the exception of sPCV-R, vocabulary knowledge and phonological skills correlated significantly negatively with the rest of the response types, i.e. consistently incorrect, variable with hits and variable with no hits. When controlling for participants’ age, there were no significant correlations between variables, except for a negative correlation between sPCC-R and consistently incorrect response types (partial  $r_s = -.50, p < .001$ ). This result indicated that the more accurately a child produced consonants in spontaneous speech, the less he/she produced consistently incorrect word forms in the variability measure.

**Table 2.** Spearman's rho correlation coefficients and partial correlations (in parentheses), controlling for age of the participants, between the child-specific characteristics and accuracy and variability of speech.

Child-specific variables	consistently correct	consistently incorrect	variable with hits	variable with no hits
ROWPVT	0.62 *** (0.08)	-0.57 *** (-0.08)	-0.41 *** (-0.01)	-0.43 *** (-0.04)
EOWPVT	0.61 *** (0.09)	-0.57 *** (-0.13)	-0.39 *** (0.002)	-0.41 *** (-0.02)
sPCC-R	0.57 *** (0.21)	-0.70 *** (-0.50) ***	-0.29 * (0.04)	-0.39 *** (-0.11)
sPVC-R	0.18 (-0.07)	-0.17 (0.05)	-0.07 (0.10)	-0.25 * (-0.11)

ROWPVT = Receptive One-Word Picture Vocabulary Test; EOWPVT = Expressive One-Word Picture Vocabulary Test; sPCC-R = spontaneous speech PCC-R; sPVC-R = spontaneous speech PVC-R.

\* $p < 0.05$ , \*\*\* $p < 0.001$ .

## Discussion

Intra-word consistency and accuracy was examined in typically developing Finnish-speaking children aged 3;0 to 6;11 years. Their response types were divided into four mutually exclusive categories: consistently correct, consistently incorrect, variable with hits, and variable with no hits. Furthermore, possible correlations between word accuracy and consistency, and children's vocabulary knowledge and phonological skills were addressed.

### *Intra-word consistency and accuracy in different age groups*

In the current study, highly consistent and phonemically accurate speech production skills were achieved at the age of 5, at which age skills were similar to those of the 6-year-olds. Furthermore, even among younger age groups, the responses were rather consistent. The median of consistently correct responses was higher among the 4-year-old children (85%) than among the 3-year-olds (75%), but the difference was not statistically significant. The 3-year-olds produced consistently incorrect responses significantly more often than the older children, although even among the 3-year-olds this response type was rare (the median of consistently incorrect responses was about 5%). The results are parallel with those of Holm et al. (2007), although Sosa (2015) found remarkably fewer consistently correct responses (24%) and somewhat more consistently incorrect responses (13%) among the 3-year-olds than in the current study.

In our study, proportions of variable responses with hits and without hits for 3-year-olds were about the same as in the study of Holm et al., but considerably lower than in Sosa's study. In Sosa's study the proportion of words produced variably with hits was 24% and without hits 38%, which was the most frequent response type. Similarly, Macrae and Sosa (2015) reported high mean inconsistency scores of 70% for children aged an average of 3;6 years. In our study, children aged 3 and 4 years produced similarly variable responses with hits, but 5- and 6-year-old children produced them significantly less frequently than the younger ones.

Holm et al., Macrae, and Sosa all studied English-speaking children, and due to differences in the target languages (e.g. Kunnari, 2000, 2002; Saaristo-Helin, Savinainen-Makkonen, & Kunnari, 2006; Savinainen-Makkonen, 2000; Vihman & Velleman, 2000), it is clear that direct comparison of the results is not possible. The relatively small consonant inventory of Finnish may allow more variation in the pronunciation of the sounds and yet still be distinguished from each other. This may reduce the effect on intra-word variability rates. In addition, the rareness of word-initial and word-final consonant clusters and late acquired consonants (such as fricatives), which are supposed to be more challenging to



pronounce and use, might turn up as more correctly and consistently produced words if compared to children learning other languages. On the other hand, as Finnish words are typically bi- or multisyllabic, and thus phonetically and phonologically more demanding to produce, they may pose more challenges for young children. However, even 3-year-olds are reported to manage them without truncation (Saaristo-Helin, 2009). Thus, Finnish could be deemed less challenging to acquire than English, with its larger consonant inventory and more late-acquired consonants and word-initial and -final consonant clusters. This could explain, at least partly, the discrepancy between our results and the findings of the studies of Macrae and Sosa (Macrae & Sosa, 2015; Sosa, 2015). Drawing this conclusion is complicated by the fact that regardless of the above-mentioned differences from English, the findings of the present study are parallel with those of Holm et al. (2007), reflecting rather low intra-word variability scores even among the youngest age group. In both studies, most of the variable responses in children aged 3–4 years also contained correct forms of the words, i.e. variable responses with matched adult targets, signaling maturing phonetic/phonological skills (Holm et al., 2007).

Phonetic and phonological complexity of the target words used in the variability assessment influences the results. Sosa (2015) proposed that the high variability rate in her study might be explained by the used Inconsistency Assessment (Dodd, 1995, p. 270), in which most target words included late-acquired consonants and, the share of multisyllabic words was high. However, except for one interchanged word, the target words were the same in the studies of both Sosa and Holm et al. (2007) study. So, based on these findings, it is not possible to draw the conclusion that differences in languages and hence differences in sound and syllabic structures of the target words can explain the incompatibility of the results. In our study, the obvious floor effect for ‘variable with no hits’ and the almost equally apparent one for ‘consistently incorrect’ responses reflect the fact that the phonological structures of the target words were rather easily mastered by the participants. From the clinical point of view, however, it is important that the chosen set of target words represents children’s typical expressive vocabulary and phonological skills at a given age range, as was apparently the case for the target words of the current study.

Another potential reason for discrepancy in findings could be differences in the transcription procedure. As in the current study, both Sosa (2015) and Holm et al. (2007) analysed speech samples using broad phonetic transcription including both consonants and vowels. To determine inter-rater agreement, a share of re-analysed data was varied in the studies (100% in Sosa; 7,3% in Holm et al.; 10% in the current study), but inter-judge reliability was high, ranging from 93% to 100% in all cases. Bearing in mind the fact that the focus in all these studies was on typically developing children and their elicited single word productions, it could be assumed that the main reason for discrepancy between the results of the current study and that of Holm et al. (2007) compared to Sosa (2015) might not be based on the transcription. However, the use of an acoustic or kinematic analysis method would be needed to bring more clarification to this matter.

### ***Relation between the child-specific characteristics and intra-word accuracy and consistency***

There was no connection between vocabulary size and intra-word accuracy and consistency when participants' age was controlled. This finding is consistent with that of Sosa and Stoel-Gammon (2006), but is in contrast to previous works, which have found significant correlation between expressive vocabulary and the amount of intra-word variability (Macrae, 2013; Macrae & Sosa, 2015; Sosa & Stoel-Gammon, 2012). In previous studies, most of the children have been younger (aged 1;9–4;2) than in our study, which might explain the mixed findings. To the best of our knowledge, the association has not been studied among older children. However, accuracy of non-word repetitions in typically developed children aged 3;0 and above have been reported to be more dependent on the size of expressive vocabulary than on children's articulatory ability (Edwards et al., 2004). Edwards et al. concluded that their results support the hypothesis that the more a child acquires and produces words, the more robustly generalised his/her phonological knowledge will be, and the more accurately the child will produce words. However, as Stoel-Gammon (2011) states, the production of non-words and known words reflect different processes, so the findings cannot be seen as directly comparable.

Munson, Edwards and Beckman (2011) presented one further explanation for the different results, hypothesizing that quality of phonological representations may be language-specific. That is, when a child is acquiring a language such as Finnish, with long words and simple syllable structures, there is not necessarily any need for as highly specified representations as for example in English, with its shorter words composed of more complex syllable structures. Languages with a high proportion of long words tend to have fewer phonologically similar words in the lexicon (Stoel-Gammon, 2011), and hence there are probably less demands on differentiating them from one another. Furthermore, one can assume that less time is needed to build sufficiently accurate phonological representations when there are less confusable words to deal with. Cross-language studies would be needed to provide more precise answers, but based on the current results, the nature of phonological representations for known Finnish words appears to be so fine grained, as early as the age of 3 onward, that no relation to word variability exists.

In our study, phonological skills did not correlate with variable response types after controlling for age. This result is consistent with the work of Macrae and Sosa (2015), who concluded that one potential explanation for their findings could be that the sensitivity of the used articulation test (GFTA) may not be sufficient to examine the articulatory abilities of children with typical speech-language development. This hypothesis is in alignment with observations that phonetic and phonological complexity of a word influences intra-word variability and accuracy, but only among younger children (Macrae, 2013; Sosa & Stoel-Gammon, 2012). In the current study, PCC-R and PVC-R calculated from spontaneous speech samples were used to reflect a child's phonological skills. In some previous studies, it has been brought out that in a conversation, used syllable and word forms may be easier and phoneme distributions more limited than in single-word samples (Masterson, Bernhardt, & Hofheinz, 2005; Morrison & Shriberg, 1992). We did not analyse the children's spontaneous speech samples and thus cannot compare structures in a child's spontaneous speech to targeted words in the intra-word variability measure. However, the low proportions of variable and consistently incorrect responses and high raw scores in the Test for Phonology (table 1) speak in favour of the participants' good phonological skills. It appears

that typically developed Finnish 3–6-year-old children’s basic phonological skills are so stable for known words that no correlation to word variability appears. The only significant correlation was found between sPCC-R and consistently incorrect response type, i.e. when phonemic accuracy of consonant production increased in spontaneous speech, the number of consistently incorrect words decreased, which fits this conclusion well.

## **Conclusion**

In the present study of typically developing Finnish-speaking children aged 3;0 to 6;11 years, the relatively high consistency and accuracy scores indicated that even the youngest children had well developed phonological abilities. For Finnish clinicians, this is important information when they assess children with SSD and make decisions on referral to speech therapy. In the future, more research on Finnish-acquiring children with SSD is needed in order to provide more language-specific assessment tools for differential diagnosis of speech impairment.

## **Declaration of Conflicting Interests**

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/>
- Betz, S. K., & Stoel-Gammon, C. (2005). Measuring articulatory error consistency in children with developmental apraxia of speech. *Clinical Linguistics & Phonetics*, *19*, 53–66.

- Burt, L., Holm, A., & Dodd, B. (1999). Phonological awareness skills of 4-year-old British children: An assessment and developmental data. *International Journal of Language & Communication Disorders, 34*, 311–335.
- Davis, B., Van der Feest, S., & Yi, H. (2018). Speech sound characteristics of early words: influence of phonological factors across vocabulary development. *Journal of Child Language, 45*, 673–702.
- de Castro M. M., & Wertzner, H. F. (2011). Speech inconsistency index in Brazilian Portuguese-speaking children. *Folia Phoniatica et Logopaedica, 63*, 237–241.
- Demuth, K., & Song, J. Y. (2012). How phonological representations develop during first-language acquisition. In Cohn, A. C., Fougeron, C., & Huffman, M. K. (Eds.), *The Oxford handbook of laboratory phonology* (pp. 397–406). Oxford: University Press.
- Dodd, B. (Ed.). (1995). *Differential diagnosis and treatment of children with speech disorders*. West Sussex: Whurr Publishers Ltd.
- 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). West Sussex: Whurr Publishers Ltd.
- Edwards, J., Beckman, M. E., & Munson, B. (2004). The interaction between vocabulary size and phonotactic probability effects on children's production accuracy and fluency in nonword repetition. *Journal of Speech, Language, and Hearing Research, 47*, 421–436.
- Edwards, S., Fletcher, P., Garman, M., Hughes, A., Letts, C., & Sinka, I. (1997). *Reynell developmental language scales III*. The University of Reading edition. Berkshire: NFER Nelson (Translation and standardization of the Finnish version: Psykologien kustannus Oy, 2001).
- 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.

- Ferguson, C. A., & Farwell, C. B. (1975). Words and sounds in early language acquisition. *Language*, *51*, 419–439.
- 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., & Elbert, M. (2001). Treatment for phonologically disordered children with variable substitution patterns. *Clinical Linguistics & Phonetics*, *15*, 41–45.
- Forrest, K., Elbert, M., & Dinnsen, D. A. (2000). The effect of substitution patterns on phonological treatment outcomes. *Clinical Linguistics & Phonetics*, *14*, 519–531.
- Goffman, L., Gerken, L., & Lucchesi, J. (2007). Relations between segmental and motor variability in prosodically complex nonword sequences. *Journal of Speech, Language & Hearing Research*, *50*, 444–458.
- Green, J. R., Moore, C. A., & Reilly, K. J. (2002). The sequential development of jaw and lip control for speech. *Journal of Speech, Language, and Hearing Research*, *45*, 66–79.
- 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, Maryland: York Press.
- Hadders-Algra, M. (2018). Early human motor development: From variation to the ability to vary and adapt. *Neuroscience & Biobehavioral Reviews*, *90*, 411–427.
- 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.
- Häkkinen, K. (1978). Eräistä suomen kielen rakenteen luonteenomaisista piirteistä ja niiden taustasta [Some phonological phenomena in Finnish: Background and characteristics]. An unpublished Licentiate thesis. Department of Finnish and General Linguistics, University of Turku.

- Karlsson, F. (1983). *Suomen kielen äänne- ja muotorakenne* [Finnish phonology and morphology]. Porvoo, Helsinki, Juva: WSOY.
- Kenney, K. W., & Prather, E. M. (1986). Articulation development in preschool children: Consistency of productions. *Journal of Speech and Hearing Research, 29*, 29–36.
- Kunnari, S. (2000). *Characteristics of early lexical and phonological development in children acquiring Finnish* (Doctoral Dissertation). Acta Universitatis Ouluensis, B34. Oulu University Oulu Press: Oulu.
- Kunnari, S. (2002). Word length in syllables: Evidence from early word production in Finnish. *First Language, 22*, 119–135.
- Kunnari, S., Savinainen-Makkonen, T., & Saaristo-Helin, K. (2012). *Fonologiatesti. Lasten äänteellisen kehityksen arviointimenetelmä* [Test for Phonology]. Jyväskylä: Niilo Mäki Instituutti.
- 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.
- Leonard, L. B., Rowan, L. E., Morris, B., & Fey, M. E. (1982). Intra-word phonological variability in young children. *Journal of Child Language, 9*, 55–69.
- 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.
- Martin, N. A., & Brownell, R. (2010a). *Expressive One-Word Picture Vocabulary Test – 4* (Finnish version: Kunnari, S. & Välimaa, T.; In validation). Novato, CA: Academic Therapy Publications.

- Martin, N. A., & Brownell, R. (2010b). *Receptive One-Word Picture Vocabulary Test – 4* (Finnish version: Kunnari, S. & Välimaa, T.; In validation). Novato, CA: Academic Therapy Publications.
- Masterson, J. J., Bernhardt, B. H., & Hofheinz, M. K. (2005). A comparison of single words and conversational speech in phonological evaluation. *American Journal of Speech-Language Pathology, 14*, 229–241.
- McCune, L., & Vihman, M. M. (2001). Early phonetic and lexical development: A productivity approach. *Journal of Speech, Language, and Hearing Research, 44*, 670–684.
- Metsala, J. L., & Walley, A. C. (1998). Spoken vocabulary growth and the segmental restructuring of lexical representations: Precursors to phonemic awareness and early reading ability. In L. J. Metsala & L. C. Ehri (Eds.), *Word recognition in beginning literacy* (pp. 89–120). Mahwah, NJ: Erlbaum.
- Morgan, A., Ttofari Eecen, K., Pezic, A., Brommeyer, K., Mei, C., Eadie, P., ... Dodd, B. (2017). Who to refer for speech therapy at 4 years of age versus who to “Watch and wait”? *The Journal of Pediatrics, 185*, 200–204.e1.
- Morrison, J. A., & Shriberg, L. D. (1992). Articulation testing versus conversational speech sampling. *Journal of Speech & Hearing Research, 35*, 259–273.
- Munson, B., Edwards, J., & Beckman, M. E. (2011). Phonological representations in language acquisition: Climbing the ladder of abstraction. In A.C. Cohn, C. Fougeron, & M. K. Huffman (Eds.), *The handbook of laboratory phonology* (pp. 288–309). Oxford: University Press.
- Nittrouer, S. (1993). The emergence of mature gestural patterns is not uniform: Evidence from an acoustic study. *Journal of Speech & Hearing Research, 36*, 959–972.
- Saaristo-Helin, K. (2009). Measuring phonological development: A follow-up study of five children acquiring Finnish. *Language and Speech, 52*, 55–77.
- Saaristo-Helin, K., Savinainen-Makkonen, T., & Kunnari, S. (2006). The phonological mean length of utterance: Methodological challenges from a crosslinguistic perspective. *Journal of Child Language, 33*, 179–190.



- Savinainen-Makkonen, T. (2000). Word-initial consonant omissions - a developmental process in children learning Finnish. *First Language, 20*, 161–185.
- Shriberg, L. D. (2004). Childhood speech sound disorders: From postbehaviorism to the postgenomic era. In Paul, R., & Flipsen, P. Jr. (Eds.), *Speech sound disorders in children: in honor of Lawrence D. Shriberg* (pp. 1–33). Plural publishing: San Diego.
- Shriberg, L. D., Aram, D. M., & Kwiatkowski, J. (1997). Developmental apraxia of speech: II. Toward a diagnostic marker. *Journal of Speech, Language, and Hearing Research, 40*, 286–312.
- Shriberg, L. D., & Austin, D. (1997). The percentage of consonants correct (PCC) metric: Extensions and reliability data. *Journal of Speech, Language & Hearing Research, 40*, 708–722.
- Shriberg, L. D., & Kwiatkowski, J. (1982). Phonological disorders III: A procedure for assessing severity of involvement. *Journal of Speech and Hearing Disorders, 47*, 256–270.
- Smith, A. (2006). Speech motor development: Integrating muscles, movements, and linguistic units. *Journal of Communication Disorders, 39*, 331–349.
- Smith, A., & Goffman, L. (1998). Stability and patterning of speech movement sequences in children and adults. *Journal of Speech, Language & Hearing Research, 41*, 18–30.
- Smith, B. L., McGregor, K. K., & Demille, D. (2006). Phonological development in lexically precocious 2-year-olds. *Applied Psycholinguistics, 27*, 355–375.
- Sosa, A. V. (2015). Intraword variability in typical speech development. *American Journal of Speech-Language Pathology, 24*, 24–35.
- Sosa, A. V., & Stoel-Gammon, C. (2006). Patterns of intra-word phonological variability during the second year of life. *Journal of Child Language, 33*, 31–50.
- Sosa, A. V., & Stoel-Gammon, C. (2012). Lexical and phonological effects in early word production. *Journal of Speech, Language & Hearing Research, 55*, 596–608.
- Stackhouse, J., & Wells, B. (1997). *Children's speech and literacy difficulties*. London: Whurr.

- Stoel-Gammon, C. (2011). Relationships between lexical and phonological development in young children. *Journal of Child Language*, 38, 1–34.
- Swingley, D. (2005). 11-month-olds' knowledge of how familiar words sound. *Developmental Science*, 8, 432–443.
- Swingley, D., & Aslin, R. N. (2000). Spoken word recognition and lexical representation in very young children. *Cognition*, 76, 147–166.
- Thelen, E., & Smith, L. B. (1994). *A dynamic systems approach to the development of cognition and action*. Cambridge, Massachusetts: The MIT Press.
- Toivainen, J. (1990). *Acquisition of Finnish as a first language: General and particular themes*. Turku: Publications of the department of Finnish and general linguistics of the University of Turku 35.
- Tyler, A. A., Williams, M. J., & Lewis, K. E. (2006). Error consistency and the evaluation of treatment outcomes. *Clinical Linguistics & Phonetics*, 20, 411–422.
- Vihman, M. M., & Velleman, S. L. (2000). The construction of a first phonology. *Phonetica*, 57, 255–266.
- Williams, P., & Stackhouse, J. (2000). Rate, accuracy and consistency: Diadochokinetic performance of young, normally developing children. *Clinical Linguistics & Phonetics*, 14, 267–293.
- Zanobini, M., Viterbori, P., & Saraceno, F. (2012). Phonology and language development in Italian children: An analysis of production and accuracy. *Journal of Speech, Language & Hearing Research*, 55, 16–31.

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

/mu:mi/	Moomin character
/vaip:a/	napkin
/poika/	boy
/housut/	trousers
/sam:ak:io/	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
/uetoketju/	zip