JSLHR

Research Article

Intraword Variability in French-Speaking Monolingual and Bilingual Children

Margaret M. Kehoe[a] and Emilie Cretton[a]

Purpose: This study examines intraword variability in 40 typically developing French-speaking monolingual and bilingual children, aged 2;6–4;8 (years;months). Specifically, it measures rate of intraword variability and investigates which factors best account for it. They include child-specific ones such as age, expressive vocabulary, gender, bilingual status, and speech sound production ability, and word-specific factors, such as phonological complexity (including number of syllables), phonological neighborhood density (PND), and word frequency.

Method: A variability test was developed, consisting of 25 words, which differed in terms of phonological complexity, PND, and word frequency. Children produced three exemplars of each word during a single session, and productions of words were coded as variable or not variable. In addition, children were administered an expressive vocabulary test and two tests tapping speech motor ability (oral motor assessment and diadochokinetic test). Speech sound ability was also assessed by measuring percent consonants correct on all words produced by the children during the session. Data were entered into a binomial logistic regression.

Results: Average intraword variability was 29% across all children. Several factors were found to predict intraword variability including age, gender, bilingual status, speech sound production ability, phonological complexity, and PND.

Conclusions: Intraword variability was found to be lower in French than what has been reported in English, consistent with phonological differences between French and English. Our findings support those of other investigators in indicating that the factors influencing intraword variability are multiple and reflect sources at various levels in the speech processing system.

Intraword variability, also known as token-to-token inconsistency or whole-word variability, refers to variable productions of the same target word. A classic example comes from a study by Ferguson and Farwell (1975), which reported that K., aged 1;4 (years;months), produced 10 different phonetic renditions of the word “pen” in a half-hour session. Intraword variability has been claimed to be a key characteristic of childhood apraxia of speech (CAS; American Speech-Language-Hearing Association, 2007) as well as inconsistent phonological disorder (Dodd et al., 2005); however, its efficacy as a diagnostic marker has been put into question by numerous studies showing it to be a common characteristic of typical speech development as well (Macrae & Sosa, 2015; Sosa, 2015). Intraword variability has also been shown to be influenced by several factors, both child and word related, including age, expressive vocabulary, and the lexical and phonological characteristics of the target word, also questioning whether intraword variability can be traced to a single source such as a speech-motor planning deficit.

The bulk of studies on intraword variability in typical speech development have focused on English-speaking children. The aim of the current study is to investigate intraword variability in typically developing French-speaking children. Given phonological differences between English and French, in particular, differences in syllable structure, rhythm, and stress, we posit that intraword variability may be present to a different degree in French- versus English-speaking children. We also extend previous studies by testing intraword variability in monolingual and bilingual children. Only one study that we are aware of has investigated intraword variability in bilingual children, finding no differences in rates of variability between bilingual Spanish–English and monolingual English children (Sosa & Bunta, 2019). We aim to determine whether the same results are obtained in French. Finally, we examine other potential factors that influence intraword variability including age, expressive vocabulary, gender, speech sound production ability, and the lexical and phonological characteristics of the target word, with the goal of determining which factors explain intraword

[a]Department of Psycholinguistics, Faculty of Psychology and Educational Sciences, University of Geneva, Switzerland

Correspondence to Margaret M. Kehoe: Margaret.Winkler-Kehoe@unige.ch

Editor-in-Chief: Bharath Chandrasekaran
Editor: Chao-Yang Lee

Received September 20, 2020
Revision received January 24, 2021
Accepted March 3, 2021
https://doi.org/10.1044/2021_JSLHR-20-00558

Disclosure: The authors have declared that no competing interests existed at the time of publication.

variability the best. In the following paragraphs, we summarize findings on intraword variability in typical speech and review factors that influence intraword variability. We complete the introduction with a statement of the research predictions.

Intraword Variability in Typical Speech

Intraword variability has been commonly associated with specific subtypes of speech sound disorders such as CAS, which is a deficit in the planification and programming of speech (American Speech-Language-Hearing Association, 2007), and deviant inconsistent phonological disorder, which is a deficit in phonological assembly and programming (Dodd, 1995, 2005). Despite the frequent observation of a high degree of word variability in children with CAS and inconsistent phonological disorder, there are findings in the literature that point to a less exclusive association between variability and these two subtypes of speech sound disorders (Betz & Stoel-Gammon, 2005; Murray et al., 2015).

A series of studies by Macrae, Sosa, and colleagues have found intraword variability rates to be extremely high in young monolingual English-speaking children (Macrae, 2013; Macrae & Sosa, 2015; Sosa, 2015; Sosa & Stoel-Gammon, 2012). We define intraword variability as variable productions of individual words from one token to the next within a fixed time period (i.e., typically a single session), whereby variability may consist of correct and incorrect productions (e.g., carotte /kaʁɔt/ "carrot" produced as [kaʁɔt] and [taʁɔt]) or different variants of incorrect productions (carotte produced as [taʁɔt] and [kaʁɔ]), the latter including substitutions, omissions, and additions of consonants and vowels. Macrae and Sosa (2015) report a mean intraword variability rate of 68% for a group (n = 43) of typically developing children, aged 2;6–4;2. Other studies with children, aged 1;9–5;5, report mean rates of variability ranging from 41% to 78% (Macrae, 2013; McLeod & Hewett, 2008; Sosa & Stoel-Gammon, 2012). All studies examine repeated productions of a given target word within a single session, although they vary as to whether the target words are elicited in a word naming task (Macrae & Sosa, 2015; Sosa, 2015) or a free-play session (Sosa & Stoel-Gammon, 2012). The only study that has found much lower rates of variability is Holm et al. (2007) who reported 13% variability in children aged 3;0–3;6. Differences between the results of this and other studies may relate to the way phonetic transcription was conducted: Holm et al. relied on online transcription whereas Sosa and colleagues employed consensus transcription (Macrae & Sosa, 2015; Sosa, 2015). However, methods of transcription cannot be the sole reason for the different results, because studies that have reported high rates of variability have not all employed consensus transcription (Sosa, 2015). To date, the discrepant findings between Holm et al. and other studies has not been fully resolved, although a recent study by Jones (2020) provides some partial resolve. He measured intraword variability in the spontaneous productions of five children, aged 0;11–4;0, extracted from the Providence Corpus. He recorded high rates of variability (e.g., 61%–72%) comparable to those reported by Macrae and Sosa (2015) when variability was determined according to strict criteria but much lower rates (e.g., 10%–17%) comparable to Holm et al. when variability was determined according to less strict criteria. Thus, it cannot be excluded that minor differences in what counts as variable may have led to different intraword variability rates across the various studies.

One other difference between the Holm et al. (2007) study and other studies was that they observed that the variability present in normally developing children was of a type that included an incorrect and correct production of the target word (referred to as variability with hits), whereas variability associated with speech sound disorders was characterized by multiple incorrect productions of a target word. Sosa (2015) coded variable productions in typically developing children, aged 2;6–3;11, and, nevertheless, was unable to confirm Holm et al.'s findings. The most common response type was variability, which did not include a correct production of the target word (i.e., variable with no hits). Thus, there remain divergent findings on intraword variability in typically developing children.

Intraword Variability in Languages Other Than English

This study compares the rate of intraword variability in French-speaking children to what has previously been reported in English-speaking children. As such, it aligns itself with cross-linguistic studies that have examined whether the linguistic properties of the language to be learned influence patterns of development. For example, research has shown that Spanish-speaking children precede nouns by filler syllables or determiners earlier than German-speaking children due to prosodic structural differences between Spanish and German (Lleó & Demuth, 1999). Children produce multisyllabic words earlier when they are exposed to languages that contain higher numbers of multisyllabic words (e.g., Spanish and Finnish) than when they are exposed to languages containing fewer numbers (e.g., English; Lleó, 2006; Roark & Demuth, 2000; Savinainen-Makkonen, 2000). In segmental acquisition, children display earlier acquisition of dorsals relative to coronals in languages in which dorsals are frequent such as in Japanese and Greek (Beckman et al., 2003; Nicolaidis et al., 2003). We examine whether intraword variability is also conditioned by the linguistic characteristics of the ambient language.

Apart from English, intraword variability has been investigated in Brazilian Portuguese, Dutch, and Finnish (de Castro & Wertzner, 2011; Faes & Gillis, 2018; Martikainen et al., 2019, 2020). de Castro and Wertzner (2011) report findings on a speech inconsistency test with typically developing and phonologically disordered Brazilian Portuguese–speaking children. The children were aged 5–10 years, older than the children studied by Macrae, Sosa, and colleagues (Macrae, 2013; Macrae & Sosa, 2015; Sosa, 2015; Sosa &

Stoel-Gammon, 2012). The mean percentage of variability in the typically developing children was low, namely, 9.8%.

Faes and Gillis (2018)[1] reported variability rates (variable with hits + variable with no hits) in typically developing Dutch-speaking children. They observed values of 56%, 39%, and 36% for children aged 3, 4, and 5 years, respectively, values that approach the high ones reported for English-speaking children by Macrae and Sosa (2015). In contrast, Martikainen et al. (2019, 2020) report lower levels of variability in typically developing Finnish-speaking children. The median value of variable productions (variable with hits + variable with no hits) was 16% in children, aged 3 years, and 5% in children, aged 5–6 years, in Martikainen et al.'s (2019) study. It was 20% in children, aged 3;2–5;4, in Martikainen et al.'s (2020) study. Martikainen et al. (2019) posit that one of the reasons for the lower rates in Finnish compared to English relates to differences in the phonological characteristics of the two languages. Finnish has a smaller phoneme inventory and less complex syllable structure (i.e., more open syllables, fewer consonant clusters) than English. Thus, Finnish may present less challenges in terms of phonological complexity than English. In a similar vein, it could be posited that the higher rates of intraword variability observed by Faes and Gillis (2018) are consistent with the fact that Dutch, being a Germanic language, resembles English in its phonological characteristics, having relatively complex syllable structure.

Potential information on the influence of the target language on intraword variability also comes from Sosa and Bunta (2019), who measured intraword variability in bilingual Spanish–English children. They observed significantly lower mean rates of variability in Spanish compared to English (41% vs. 55%). Sosa and Bunta (2019) attributed the differences in variability to the fact that the children were more dominant in Spanish than English. It is possible, however, that phonological differences between Spanish versus English may also play a role in the different variability rates: Spanish has less complex syllable structure than English (Keffala et al., 2018). It also has a smaller vowel inventory, consisting of only five vowels compared to approximately 14 in English, and has syllable- as compared to stress-timed rhythm (Ramus et al., 1999; see below).

In the current study, we measure intraword variability in French. French differs from English in having a lower proportion of closed syllables than English. The percentage of closed syllables in written text is 24% according to Delattre and Olsen (1969) where it is 60% in English. French has syllable-timed rhythm compared to English, which has stress-timed rhythm (Ramus et al., 1999). Given that segmental accuracy is often reduced in unstressed versus stressed syllables in English (Schwartz & Goffman, 1995), the presence of increased prominence across all syllables of the word due to syllable-timed rhythm may facilitate segmental production. In addition, French differs from English in not having lexical but phrasal stress. Primary stress falls on the final syllable of the last lexical item in a phonological phrase (Dell, 1984), which, in a denomination task, would mean that the final syllable of the word receives stress. The more fixed stress pattern of French, as opposed to the more variable one of English, may also have a positive influence on consonant and vowel accuracy in young children.

Indeed, there are isolated reports that French-speaking monolinguals obtain superior phonological scores than English-speaking monolinguals when compared on similar tasks (Brosseau-Lapré & Rvachew, 2014; MacLeod et al., 2011). MacLeod et al. (2011) report earlier mastery of certain consonants (e.g., /v, z, r/) in French-speaking preschool children, aged 1;8–4;5, compared to English-speaking children. Brosseau-Lapré and Rvachew (2014) found that French-speaking phonologically disordered children, aged 4;0–5;11, made fewer consonant errors (including distortions) than English-speaking phonologically disordered children, although they did make more syllable-structure errors. In another study, Rvachew et al. (2014) documented earlier acquisition of word-initial consonant clusters in French-speaking children, aged 3;10–5;9, in comparison to English-speaking children. In summary, differences in the segmental and prosodic characteristics of French versus English appear to influence phonological acquisition patterns. We posit that the same factors may influence intraword variability. Specifically, we predict that we may observe lower rates of intraword variability in French- as opposed to English-speaking children.

Factors That Influence Intraword Variability

When examining factors that influence intraword variability, we separate them into those related to the child such as age, expressive vocabulary, gender, bilingual status, and speech sound production ability, and those related to the target word such as phonological complexity, phonological neighborhood density (PND), and word frequency (WF).

Child-Related Factors

Age. Several studies indicate that whole-word variability decreases with age. This has been reported by Macrae (2013) for children aged 1;9–3;1, Sosa (2015) for children aged 2;6–3;11, Holm et al. (2007) for children aged 3;0–6;11, and more recently by Jones (2020) for children tested longitudinally from age 0;11 to 4;0. Developmental effects on word variability have been reported in Dutch and Finnish as well (Faes & Gillis, 2018; Martikainen et al., 2019). This result is consistent with wide-scale findings indicating increasing articulatory and phonological precision with age (Bernthal, 2009, p. 179).

Expressive vocabulary. Another common finding is that expressive vocabulary level is correlated with word variability. Macrae (2013) observed a relationship between both age and expressive vocabulary on the one hand and whole-word variability on the other hand. Other studies have found a relationship between age and variability (Martikainen et al., 2019) or expressive vocabulary and variability (Macrae

[1]Faes and Gillis (2018) and also Sosa and Bunta (2019) investigate intraword variability in children with cochlear implants; however, we focus on their findings with normal-hearing children.

& Sosa, 2015; Sosa & Stoel-Gammon, 2012) but not necessarily both. For example, Macrae and Sosa (2015) found that age was not a significant predictor of whole-word variability when expressive vocabulary was included as predictor. Similarly, Sosa and Stoel-Gammon (2012) found a strong relationship (r = .81) between intraword variability and expressive vocabulary size as based on parent report in typically developing 2-year-olds; there was no significant difference in production variability for the two age groups studied, 2;0 and 2;6. The relationship between vocabulary and variability has been linked to children's underlying representations of words. Children who have smaller vocabularies may have less complete—more holistic and less well-specified—phonological representations (Menn & Matthei, 1992; Metsala & Walley, 1998), which in turn leads to variability in word production. In comparison to expressive vocabulary, receptive vocabulary appears to be less consistently associated with variability (Macrae & Sosa, 2015; Martikainen et al., 2019, 2020).

Gender. Holm et al. (2007) reported gender differences in intraword variability. Girls produced more consistent correct responses and fewer variable incorrect responses than boys. In contrast, Sosa and Stoel-Gammon (2012) as well as Martikainen et al. (2019) did not find significant differences in production variability between boys versus girls.

Bilingual status. Sosa and Bunta (2019) compared intraword variability in monolingual English and bilingual Spanish–English children. Based on the findings of previous literature, which indicate that bilingual children are not strongly disadvantaged in the area of phonological development compared to monolingual (Fabiano-Smith & Goldstein, 2010; Goldstein & Bunta, 2012; Goldstein et al., 2005; Grech & Dodd, 2008), the authors did not predict differences between monolingual and bilingual children. Indeed, their predictions were confirmed: Bilingualism did not affect overall rates of whole-word variability.

Nevertheless, alternative scenarios concerning the influence of bilingualism on intraword variability could be envisaged. Bilinguals need to represent greater numbers of phonetic categories within the same phonetic space (Byers-Heinlein & Fennell, 2014), which may lead to more specified phonological representations and consequently reduced variability relative to monolinguals. Alternatively, bilinguals need to deal with two different phonetic inventories, which may have different phonetic variants for the same phoneme (e.g., dental /t/ in French vs. alveolar /t/ in English). Execution of these fine phonetic differences may lead to greater variability in word production compared to monolinguals. Further studies on intraword variability in bilingual children would help to determine whether these alternative possibilities are attested.

Speech sound production ability. Macrae and Sosa (2015) examined the influence of children's speech sound abilities on word variability. They did not find a significant relationship between the two. They queried, however, whether the test they used to measure speech sound abilities, namely, a standardized test of articulation (the Goldman-Fristoe Test of Articulation–Second Edition; Goldman & Fristoe, 2000), was sensitive enough to measure speech abilities in children with typical phonological development. More recently, Martikainen et al. (2019) evaluated the relationship between intraword variability and speech sound ability using percent consonants correct (PCC) based on broad phonetic transcription of spontaneous speech and observed similar results to Macrae and Sosa (2015): Phonological skills did not correlate with variability measures. In a later study, they calculated PCC on the basis of narrow phonetic transcription and found a relationship between variability (i.e., variable with no hits) and speech sound ability (Martikainen et al., 2020). They argue that narrow transcription that includes detection of speech sound distortions may better reflect the motoric difficulties that are often associated with variability. One of the features of the current study will be to measure speech sound production ability using tasks that tap both the phonological and speech motor aspects of speech production.

Word-Related Factors

Phonological complexity. Studies have shown that high complexity target words are produced less consistently than low complexity target words (Macrae, 2013; Sosa, 2015; Sosa & Stoel-Gammon, 2012). Sosa and Stoel-Gammon (2012) measured phonological complexity by determining the average age of acquisition of each consonant or cluster in a given target word. They predicted that words containing later acquired consonants would be produced more variably than words with early acquired consonants, a finding that was confirmed using a multiple regression model. Following Sosa and Stoel-Gammon (2012), Macrae (2013) designed a test of word variability in which 10 words were selected to contain only early developing sounds (stops, nasals, glides) and 10 words were selected to contain late developing sounds (fricatives, affricates, and liquids) and consonant clusters. Macrae (2013) found phonological complexity to have a strong positive effect on word variability.

Word length is also an aspect of phonological complexity. Sosa (2015) investigated the influence of word length on word variability in children aged 2;6–3;11. She reported that one-syllable words were the most stably produced, whereas four-syllable words were the most variable: 27 out of 33 children displayed 100% variability in their productions of four-syllable words. Faes and Gillis (2018) also showed that variability increased with increasing word length in Dutch-speaking children.

Quantitative measures of phonological complexity have been developed including Stoel-Gammon's (2010) Word Complexity Measure and Jakielski's (2000) Index of Phonetic Complexity (IPC).[2] Both measures assign complexity points for parameters known to pose difficulty for children in phonological production. They include presence of consonant

[2]Jakielski (2000) used the term *phonetic complexity*; however, we prefer to use phonological complexity as the parameters are more phonological than phonetic. Nevertheless, we code [ʁ] as a fricative rather than its phonological manner of articulation, liquid.

variegation, consonant clusters, dorsal place of articulation (PoA; vs. coronal or labial), and fricative and liquid manner of articulation (vs. stop, nasal, and glide). In the current study, we employ Jakielski's (2000) IPC to operationalize phonological complexity.

PND and WF. Sosa and Stoel-Gammon (2012) examined the influence of PND and WF on whole-word variability. PND refers to the number of phonological neighbors of a word whereby a phonological neighbor is a word that differs from another word by substitution, deletion, or addition of a sound in any word position (Luce & Pisoni, 1998). Words that contain many phonological neighbors are said to belong to dense neighborhoods, whereas those that contain few neighbors belong to sparse neighborhoods. WF refers to the frequency of occurrence of words, that is, the number of times a word has been produced. Sosa and Stoel-Gammon (2012) found that both PND and WF influenced word variability. Words that had high-density neighborhoods and were highly frequent were less variable than words that had low-density neighborhoods and were less frequent. Their findings were recently confirmed by Jones (2020) when examining intraword variability in the spontaneous speech productions of five English-speaking children. He documented a robust association between PND and WF of the target word on the one hand and variability on the other.

It can be assumed that both factors influence variability by their impact on phonological representations. Words with many phonological neighbors are likely to have more detailed phonological representations than words with fewer neighbors in order to prevent confusion with similar sounding words. Frequent exposures to a target word contribute to lexical strengthening leading to more stable and robust phonological representations. These findings, in conjunction with those showing an influence of vocabulary size on variability, are consistent with models of phonological development that suggest that phonological representations become more segmental and abstract as children add words to their lexicons.

Other word-related factors that have been examined in studies on word variability include phonotactic probability (i.e., relative frequency of individual sounds and sound sequences in syllables and words) and age of acquisition (i.e., length of time a word has been in a child's vocabulary); however, they will not be a focus of the current study since they have not been found to consistently influence word variability (Macrae, 2013; Sosa & Stoel-Gammon, 2012).

Summary and Aims of Study

A review of the literature reveals that intraword variability is present to a high degree in typically developing English-speaking children, aged 2;0–4;6. It appears to be strongly present in Dutch- but less present in Finnish-speaking children, but, as yet, we know little about intraword variability across other languages.

Studies show that word variability is conditioned by child- and word-specific factors. Word variability decreases as children get older and as their expressive vocabulary size increases. To date, there is no evidence that bilingualism or speech sound production ability influences intraword variability, and findings are inconsistent with respect to gender differences. Studies focusing on word-specific factors have observed that words with low phonological complexity, high-density neighborhoods, and high frequency are less variable than words with high phonological complexity, low-density neighborhoods, and low frequency.

In the current study, we investigate intraword variability in French-speaking monolingual and bilingual children, aged 2;6–4;8. Our first aim is to determine whether intraword variability is present to a high degree in French phonological development as it is in English. Our second aim is to examine factors that influence intraword variability. Child-specific factors include age, expressive vocabulary knowledge, gender, bilingual status, and speech sound production ability (phonological and speech motor ability), and word-specific factors include phonological complexity, PND, and WF.

Based on our literature review, we predict that intraword variability will

- be lower in French than it has been reported to be in English because of differences in syllable structure and prosody between the two languages;
- decrease with increasing age and expressive vocabulary level;
- not be related to bilingual status or speech sound production ability; and
- be greater for words with high phonological complexity, low PND, and low WF.

We do not make strong predictions concerning gender effects on intraword variability given conflicting findings. We also leave open the possibility that there may be differences in intraword variability in monolingual and bilingual children as suggested by the alternative possibilities presented above.

Method

Participants

Participants include 40 typically developing French-speaking monolingual and bilingual children, aged 2;6–4;8,[3] recruited in kindergartens in the cantons of Geneva and Valais, Switzerland. Participants were not required to achieve a minimum score in order to be included in the study; however, out of the 45 children originally tested, we made an exception for two children who were excluded due to extremely low scores on the expressive vocabulary test (less than 2 *SD*s below the mean). Two children were also excluded due to failure to complete several of the tests, and

[3]Our intention was to test children, aged 2;6–4;6; however, a male bilingual child, aged 4;8, was inadvertently tested, and because he did not pattern differently from the other children in his age group, he was included in the study.

one child was excluded because of inaudible audio recording. Table 1 displays the distribution of children according to age range, bilingual status, and gender. Overall, there were 15 girls and 25 boys; 21 monolinguals and 19 bilinguals. The average age of the children was 3;7 (monolinguals: 3;6; bilinguals: 3;8). All parents of children included in the study completed an informed consent form, which was approved by the University of Geneva's ethics committee.

Bilingual status was determined by a parent questionnaire (loosely based on the Parents of Bilingual Children Questionnaire; Tuller, 2015), in which parents indicated whether their child spoke another language at least 30% of the time in addition to French. They were required to indicate which language the child spoke at home and with whom, and at what age the child had acquired French. Parents were also required to judge the language usage of French and the other language on a scale from 1 to 5. In addition, they indicated whether they had any concerns about their child's speech and language development.

Information provided in the questionnaire revealed that all children (monolingual and bilingual) had normal hearing, were in good health, and were developing normally. The bilingual children had all acquired French before the age of 3 years and, thus, could be considered simultaneous bilinguals. Fourteen of the 19 bilinguals were dominant in French. The remaining were dominant in the home language (n = 2; Polish and Spanish) or were balanced bilinguals (n = 3). The principal languages spoken by the bilinguals were Italian (five children), Albanese (four children), Swiss–German (three children), and Portuguese (three children). Socioeconomic status was determined by asking parents how many years of supplementary schooling they did (beyond obligatory schooling). Appendixes A and B provide further details of the monolingual and bilingual children including exact age; languages spoken by the mother, father, and caretaker; dominance; and number of years of supplementary schooling by the mother.

Test Materials

Variability Test

The variability test is based on the Inconsistency Assessment of Dodd (1995), which involves asking a child to name 25 pictures on three separate occasions in a single testing session. We selected 25 words from the *l'Inventaire Français du Développement Communicatif chez le nourrisson: mots et phrases* (Kern & Gayraud, 2010), the French version of the MacArthur Communicative Development Inventories (Fenson, 1993), to ensure that all words were familiar to children as young as age 2;6. Words were then selected to consist of one to four syllables, to be imageable, and to contain a range of phonological neighbors, word frequencies, and syllable structures. The database, Lexique 3 (New et al., 2007), was used to provide PND and WF information. In addition, we employed the IPC (Jakielski, 2000) to determine the phonological complexity of each word. A point was assigned to each word if it contained a dorsal consonant (e.g., clé [**k**le] "key"); a fricative or liquid (e.g., fleur [**fl**œʁ] "flower"; étoile [etwa**l**] "star"); a final consonant (e.g., robe [ʁɔ**b**] "dress"); three or more syllables (e.g., pantalon [**pãtalɔ̃**] "pants/trousers"); two or more consonants with different PoAs (e.g., robe [ʁɔ**b**] "dress," which has dorsal and labial PoAs); a tautosyllabic cluster, a cluster that occurs within a syllable (e.g., fleur [**fl**œʁ] "flower"); or a heterosyllabic cluster, a cluster that is split across two syllables (e.g., tracteur [tʁa**kt**œːʁ] "tractor"). The IPC also assigns points to rhotic vowels, but since rhotic vowels do not occur in French, this category was excluded. Appendix C presents the 25 stimulus words of the variability test along with their WF, number of phonological neighbors, syllabic structure, and IPC value. Appendix D provides the phonetic form and the meaning of the variability test words in English.

The variability test was administered in the form of a memory game (i.e., the child has to find a pair of the same picture by remembering where the picture is situated within an array of pictures). The child was first asked to name all the pictures so as to ensure that the child was familiar with the word and also to obtain the first production of the stimulus words. The child was then required to play the memory game. Because of the number of pictures (n = 25) and the young age of the children, the task was administered in two goes. A memory game was played with 13 pairs at the beginning of the session and then with 12 pairs later on. In between, other tests were administered (see General Procedure section). During the memory game, the child was required to name the picture that he/she selected. The experimenter paid careful attention to elicit at least three productions of each stimulus word during the session. Due to varying degrees of cooperation, some children did not produce multiple productions of all 25 words. Missing data were present on 10 children, although, in the majority of cases (n = 8 children), only one to three items out of 25 words were not produced 3 times. In the case of two children, aged 2;7 and 2;8, however, eight words were not produced 3 times. Rather than excluding these children altogether, variability was determined on the basis of the number of items produced 3 times. In those cases in which there were more than three productions of the stimulus word, the first three tokens were taken with the proviso that they were intelligible productions.

In the majority of cases, children responded with a single word or article plus noun (e.g., une fleur "a flower") but occasionally they produced phrases as well (e.g., c'est une joli fleur "It's a pretty flower"). We included the stimulus word regardless of whether it was produced in isolation or within a short phrase. However, the word was always

Table 1. Number of participants according to age (years;months), bilingual status, and gender.

Age group	Monolingual		Bilingual		Total
	B	G	B	G	
2;6–2;11	1	4	2	2	9
3;0–3;5	7	1	2	1	11
3;6–3;11	0	2	4	1	7
4;0–4;8	5	1	4	3	13

Note. B = boy; G = girl.

situated in prominent phrase-final position. Most of the time, productions were spontaneous but on a few isolated occasions, it was necessary to include imitated productions. Kehoe and Havy (2019) documented no differences in percent consonant accuracy between spontaneous and imitated productions with a similar database of young French bilinguals; however, McLeod and Masso (2019) found greater percent accuracy in imitated productions for consonants produced in medial and final positions. We are unaware of studies that have examined whether the presence of a verbal model influences whole-word variability.

In terms of coding, we used a simple binary distinction. Word productions were coded as variable when at least two of the three productions were different. We took into account both vowel (e.g., fleur /flœʁ/ "flower" produced as [flaʁ] and [flœʁ]) and consonant differences (e.g., fleur produced as [flœʁ], [fjœʁ] or [flœ]) in determining whether variability was present or not. We did not take distortions (e.g., escalier /ɛskalje/ "stairs" produced as [ɛskalje] and [ɛs̪kalje]) into account; however, they were very infrequent (two examples) in the database (see also Brosseau-Lapré & Rvachew, 2014, who report low percentages of distortions in the speech of French- vs. English-speaking children). Examples of variable and nonvariable productions are given in Table 2.

Vocabulary Test

Expressive vocabulary in French was tested using the subtest "Dénomination Phonologie/Lexique" of the test battery Evaluation du développement du langage oral chez l'enfant de 2 ans 3 mois à 6 ans 3 mois (EVALO2-6; Coquet et al., 2009). Children were required to name a series of 40 items (32 nouns and eight verbs). If the children spontaneously named the picture, they received a score of 2. If the children named the picture after having received a phonological cue (the first phoneme of the word), they received a score of 1. We utilized the initial test score (without phonological cue) as we considered it to be a more valid measure of the children's vocabulary knowledge. We also employed the raw rather than the standardized score, given that the normative data for the EVALO2–6 do not take bilingualism into consideration. The total possible score was 80.

Tests of Speech Sound Production Ability

To measure children's speech sound production ability, we collected data from three sources. First, we calculated PCC on all word productions elicited during the test session. This included the words spoken in the variability and vocabulary tests and any other spontaneous productions elicited in conversation. Second, we tested children's articulatory gestures via an oral motor assessment (OMA). Third, we assessed children's ability to sequence sounds via an oral diadochokinetic (DDK) task. The first measure targeted children's phonological skills and the second and third targeted children's speech motor/articulatory skills. We are aware, nevertheless, that it is difficult to tap, in an isolated fashion, the separate components of speech sound ability. A measure of PCC will also reflect speech motor skills, as will an OMA and DDK task reflect phonological abilities to some extent. The inclusion of tests tapping speech motor control stems from the fact that high intraword variability is often reported in children with speech motor planning deficits. We were interested in knowing whether intraword variability would be sensitive to speech motor control differences even in normally developing children.

OMA. The OMA consisted of the subtest "Praxies Bucco-facial et Linguales" from the test battery EVALO2–6 (Coquet et al., 2009). The child was required to realize 18 articulatory gestures and phonetic sequences via imitation (e.g., elevate the tongue tip towards the nose; puff up the two cheeks). The child received a maximum score out of 18.

DDK. The DDK task was based on the "Epreuves de séries diadococinésiques" developed for French-speaking children by Martinez Perez et al. (2015). In the current study, children were required to repeat the two sequences /pa-ta-ka/ and /bə-də-gə/ 6 times. To make the task accessible for children as young as age 2;6, we presented the child with a visual image of six caterpillars, each composed of three sections (Masson, 2017). The experimenter pointed to the different sections of each caterpillar to assist the child in realizing the correct number of sequences. The time taken to produce six repetitions of each sequence was calculated, and the mean time for the two sequences was determined. Three children did not cooperate in the DDK task.

General Procedure

Children took part in a single session of approximately 30 min conducted in a quiet room in the children's kindergarten. The children interacted with one native French-speaking experimenter (the second author) and a French-speaking assistant from the kindergarten. The tasks were given in a general order: 1. presentation of variability test items (first production of stimulus words); 2. DDK or OMA; 3. variability test in the form of a memory game (second and third productions of Items 1–13); 4. vocabulary test; 5. DDK or OMA; and 6. variability test in form of memory

Table 2. Examples of variable and nonvariable productions.

Stimulus words	Production no. 1	Production no. 2	Production no. 3	Variable/nonvariable
chien	sjɛ̃	sjɛ̃	sjɛ̃	Not variable
carotte	kaʁɔt	kaʁɔt	kaʁat	Variable
éléphant	elefɑ̃	elefɑ̃	ejefɑ̃	Variable
escalier	ɛʃkalje	ɛskalje	ekalje	Variable
zèbre	zɛʁbʁ	zɛʁb	zɛʁ	Variable

game (second and third productions of Items 14–25). Although it might have been preferable to completely counterbalance the order of tests, this general order was found to be most conducive to obtaining the children's cooperation.

Data Transcription

Children's speech was recorded with a portable digital tape recorder (MARANTZ, TASCAM DR-2d) and unidirectional condenser microphone placed on a table in front of the children. Using Phon, a software program designed for the analysis of phonological data (Rose & MacWhinney, 2014; Rose et al., 2006), each child's WAV format file was segmented, and stimulus words were identified and transcribed. A French-speaking graduate student (second author), who had experience in phonetic transcription, performed the analyses. She transcribed the data using broad phonemic transcription. The transcribed data were transferred to Excel. Calculations of PCC were computed automatically for each child in Phon.

Reliability

Six participants (15% of the sample) were retranscribed by a second transcriber (one of three undergraduate students who had experience in phonetic transcription) using the Blind Transcription function of the Phon program. Mean point-to-point agreement for consonants was 93.83% (range of 90.46%–96.38% for the six children), and mean agreement for vowels was 94.42% (range of 90.32–97.77%), which indicates good interrater agreement.

Data Coding

Data were analyzed using mixed-effects logistic regression, which allowed us to model production accuracy on the basis of binomial data. The analyses were performed using R statistical software (R Development Core Team, 2015) and the lme4 package (Bates et al., 2015) for mixed-effects models. To evaluate the contribution of each predictor in the model, we performed pairwise model comparisons between a saturated and a more restricted model. The saturated model included all main effects, whereas the restricted model omitted the predictor under consideration. Comparisons were made using log likelihood ratio tests, which yield a chi-squared statistic.

The dependent variable was variability coded as 0 or 1 (not variable or variable) for each word in the variability test. Child-related predictor variables included age (coded in months), vocabulary (raw score on vocabulary test), gender (male, female), bilingual status (monolingual, bilingual), PCC (percentage score), oral motor ability (raw score on test), and DDK score (mean time in milliseconds). Word-related predictor variables included phonological complexity (based on IPC), number of syllables, PND, and WF. WF was log transformed due to the skewed nature of the raw frequency values (Brysbaert et al., 2018). We also included the control variable, years of supplementary education of the mother, which was an indicator of socioeconomic status. To determine what factors best influenced whole-word variability, we entered all variables and examined which variables were significant. The random part of the model included random intercepts for participants and items. The model was fitted using maximum likelihood estimation.

Results

Table 3 presents descriptive statistics on the tests conducted in the study. The mean variability across the 40 children in the study was 29%. The mean vocabulary score was 48 out of a maximum of 80. The mean PCC in the sample was 92%. The children scored approximately 14 out of 18 on the OMA and repeated a DDK sequence (e.g., /pataka/) 6 times in an average time of 2.1 s. Table 4 presents the same results for monolingual and bilingual children separately. We examined whether there were any differences between the monolingual and bilingual children on the basis of a series of *t* tests. Monolingual and bilingual children did not differ in terms of their age, $t(38) = 0.46$, $p = .62$; PCC, $t(38) = .14$, $p = .91$; OMA, $t(38) = 0.62$, $p = .57$; nor DDK results, $t(34) = 0.07$, $p = .95$, but they did differ in terms of their vocabulary scores, $t(38) = 2.71$, $p = .01$. Bilinguals had smaller vocabulary scores than monolinguals.

We also examined the correlations among the child-related variables and among the word-related variables. These results are presented in Tables 5 and 6. For the child-related variables, all predictor variables were significantly correlated with variability with the exception of the DDK score. Age and vocabulary were also significantly correlated with PCC. For the word-related variables, all variables were significantly correlated with variability with the exception of WF. In addition, WF was significantly correlated with PND. The correlations were of a moderate degree and did not present a risk of multicollinearity.

Figures 1 and 2 present the mean intraword variability for the two categorical variables, gender and bilingual status. Figure 1 indicates that there were no apparent differences between the mean variability of males and females, and Figure 2 shows that monolinguals and bilinguals obtained similar variability scores with a tendency for bilinguals to display greater variability than monolinguals. Independent

Table 3. Descriptive statistics on tests conducted including variability, vocabulary, oral motor assessment, oral DDK task, and PCC for the entire sample (*N* = 40).

Test	*M*	*SD*	Range
Variability[a]	29.08	16.95	8–75
Vocabulary	48.35	12.70	20–68
PCC[b]	92.71	7.57	65.08–99.07
OMA	13.98	2.81	5–18
DDK	2.11	0.90	0.9–4.6

Note. OMA = oral motor assessment; DDK = diadochokinetic.

[a]Variability is presented as percentage score; Vocabulary is a raw score out of 80; PCC is presented as percentage score; OMA is a raw score out of 18; DDK is time in seconds. [b]PCC = percent consonants correct.

Table 4. Descriptive statistics on tests conducted including variability, vocabulary, oral motor assessment, oral DDK task, and PCC for monolingual (n = 21) and bilingual children (n = 19) separately.

	Monolinguals			Bilinguals		
Test	***M***	***SD***	**Range**	***M***	***SD***	**Range**
Variability[a]	26.44	14.53	8–68	32	19.26	8–75
Vocabulary	53.05	12.16	26–68	43.16	11.46	20–62
PCC[b]	92.97	8.95	65–99	92.43	5.93	83–99
OMA	13.62	2.50	7–17	14.37	3.13	5–18
DDK	2.12	0.90	0.9–4.6	2.09	0.92	1.1–4.3

Note. OMA = oral motor assessment; DDK = diadochokinetic.
[a]Variability is presented as percentage score; Vocabulary is a raw score out of 80; PCC is presented as percentage score; OMA is a raw score out of 18; DDK is time in seconds; [b]PCC = percent consonants correct.

t tests did not indicate any differences in variability between males and females, and monolingual and bilinguals based on mean scores per participant (gender: $t(38) = 0.19$, $p = .42$; bilingualism: $t(38) = 1.04$, $p = .31$). However, it must be realized that these tests do not take into account the uneven numbers of males and females, and monolinguals and bilinguals at different ages nor the influence of other factors (e.g., socioeconomic status, vocabulary). A more rigorous test of the influence of gender and bilingual status will be provided in the logistic regression model.

We also present variability according to syllable number. Figure 3 shows that variability increased with increasing syllable number, although the clearest difference was between one-, two, and three-syllable words on the one hand and four-syllable words on the other. Finally, we display intraword variability as a function of age group to provide additional information on the influence of age on whole-word variability. As can be seen in Figure 4, the youngest age group (2;6–2;11) had the highest intraword variability (i.e., 45%), the next two age groups (3;0–3;5 and 3;6–3;11) had similar levels of intraword variability (i.e., 27%–29%), and the oldest age group (4;0–4;8) had the lowest level of variability (i.e., 19%).

To determine what factors best influenced children's intraword variability, we ran a mixed-effects logistic regression

Table 5. Correlation coefficients between the child-related variables.

Variable	Age	Vocab	PCC[a]	OMA	DDK
Variability	−.58[b]	−.52[b]	−.65[b]	−.46[b]	.27
Age	—	.26	.49[b]	.35[b]	−.17
Vocab	—	—	.48[b]	.23	−.21
PCC	—	—	—	.18	−.08
OMA	—	—	—	—	−.13
DDK	—	—	—	—	—

Note. OMA = oral motor assessment; DDK = diadochokinetic.
[a]PCC = percent consonants correct.
[b]$p < .05$.

Table 6. Correlation coefficients between the word-related variables.

Variable	Complexity	PND[a]	WF
Variability	.64[b]	−.60[b]	−.20
Complexity	—	−.36	−.38
PND	—	—	.53[b]
WF	—	—	—

Note. WF = (log-transformed) word frequency.
[a]PND = phonological neighborhood density.
[b]$p < .01$.

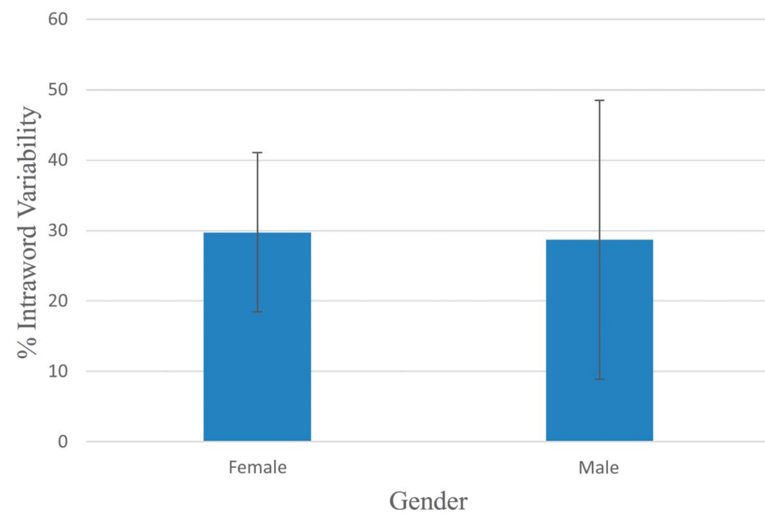
models entering all predictor variables as well as the control variable, supplementary years of schooling by mother. There were 968 individual items (40 children × 25 words – missing data) included in the analysis. Whether the word was variable or not served as the dependent variable. We employed information criteria such as the Akaike information criterion and Bayesian information criterion to determine the best fitting model. In initial analyses, the variable DDK was not significant. Because it resulted in the removal of three children (three children did not complete the DDK tests), we reran the analysis excluding DDK and subsequently including all 40 children in the analyses. In the resultant model, several factors were found to significantly improve model fit to data. The following child-related factors were significant: age ($\beta = -.42$, $\chi^2(1) = 12.08$, $p < .001$), gender ($\beta = .44$, $\chi^2(1) = 4.95$, $p = .03$), bilingual status ($\beta = -.60$, $\chi^2(1) = 7.59$, $p = .006$), PCC ($\beta = -.41$, $\chi^2(1) = 13.80$, $p < .001$), and scores on the OMA ($\beta = -.22$, $\chi^2(1) = 5.18$, $p = .02$). The following word-related factors were significant: word complexity ($\beta = .43$, $\chi^2(1) = 12.82$, $p < .001$), PND ($\beta = -.52$, $\chi^2(1) = 9.89$, $p = .002$), and WF ($\beta = .25$, $\chi^2(1) = 4.12$, $p = .04$). In addition, years of supplementary schooling by the mother was a significant factor in the model ($\beta = -.32$, $\chi^2(1) = 7.89$, $p = .005$). Expressive vocabulary ($\beta = .06$, $\chi^2(1) = .255$, $p = .61$) and syllable number ($\beta = -.05$, $\chi^2(1) = .12$, $p = .73$) did not emerge as significant factors. A summary of the statistical model is provided in Appendix E.

To summarize the findings: As age, consonant accuracy, and oral motor abilities of the children increased, intraword variability decreased. As word complexity increased, intraword variability increased, and as PND increased, intraword variability decreased. The influence of WF on intraword variability was positive meaning that as frequency increased, intraword variability increased, which is an effect opposite to the one reported in the literature (see later discussion). Bilinguals obtained higher variability scores than monolinguals, and males obtained higher variability scores than females.

Discussion

This study examined intraword variability in French-speaking children. We investigated whether French-speaking children obtain similar levels of intraword variability to English-speaking children and what factors best account for it. We found a mean value of 29% for intraword variability

Figure 1. Mean percentage intraword variability according to gender. Error bars refer to standard deviations.



in this sample, which is a figure lower than has been reported for English-speaking children of a similar age. We observed that several factors influence word variability, including age, gender, bilingual status, the child's speech production abilities, and the lexical and phonological characteristics of the target word. In the following paragraphs, we summarize the findings and compare them to what we know about intraword variability in English-speaking children.

Rate of Intraword Variability in French-Speaking Children

At the outset of the study, we predicted that intraword variability would be lower in French- compared to English-speaking children. French has a lower proportion of closed syllables than English and has syllable-timed rhythm, which may facilitate the production of segments in unstressed syllables. It also has a more "fixed" stress pattern than English, which may also have an influence on intraword variability. Studies that have tested intraword variability in English-speaking children of a similar age report variability rates of 41%–78% with a commonly reported value of 68% (Macrae & Sosa, 2015). Instead, we documented a lower percentage of variability in the French-speaking children. Two young children obtained very high variability scores (65% and 75%), but these scores did not characterize the majority of children in the sample. Our prediction, thus, appears to be confirmed, namely, that the different

Figure 2. Mean percentage intraword variability according to bilingual status. Error bars refer to standard deviations. bi = bilingual; mon = monolingual.

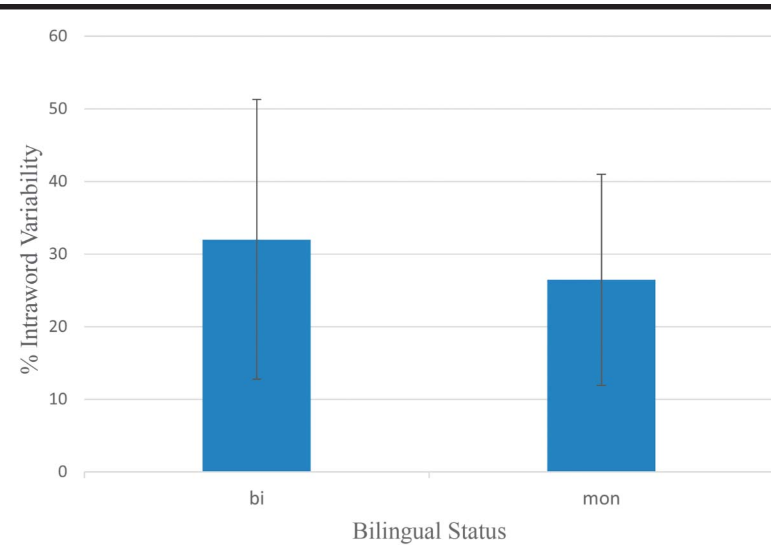
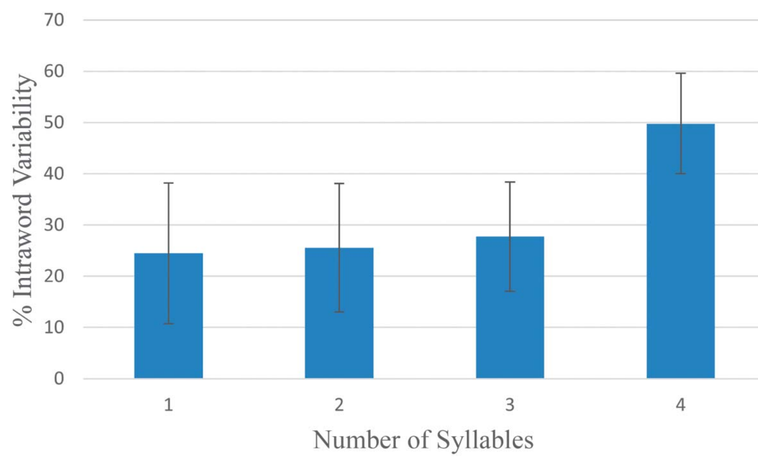


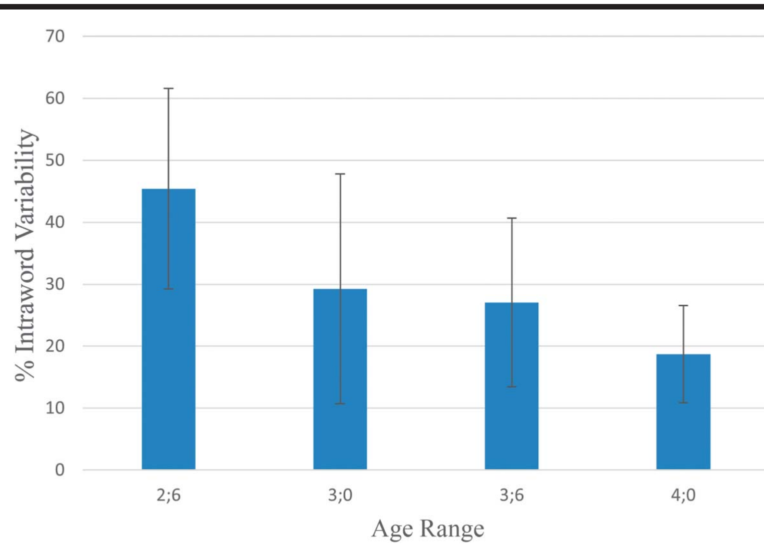
Figure 3. Mean percentage intraword variability according to number of syllables. Error bars refer to standard deviations.



phonological characteristics of French relative to English influence intraword variability rates. These findings are consistent with those of Martikainen et al. (2019, 2020), who also explain the lower rates of intraword variability in Finnish compared to English by the smaller consonant inventory and reduced syllable structure complexity of Finnish. These findings are also consistent with numerous other studies of a cross-linguistic nature revealing that language-specific phonological differences influence phonological acquisition patterns. These observations have been made for children's prosodic (Lleó, 2006; Lleó & Demuth, 1999; Roark & Demuth, 2000) and segmental patterns (Beckman et al., 2003; Nicolaidis et al., 2003). We have also found similar effects for intraword variability.

Nevertheless, for our prediction to be fully confirmed, we would need to investigate intraword variability in a single study, which includes English- and French-speaking children tested under the same methodological conditions. We would also need to widen the range of languages tested to include phonologically diverse languages, which vary on phonological parameters such as syllable structure complexity, phoneme inventory size, and rhythm. Such studies would have theoretical implications in that they would inform us on which phonological parameters are most likely to generate variability, thus enlarging our understanding of these phonological parameters. They would also have clinical implications in that they would help us predict whether children speaking a given target language will exhibit high degrees of intraword variability or not.

Figure 4. Mean percentage intraword variability according to age range (2;6–2;11; 3;0–3;5; 3;6–3;11; 4;0–4;8). Error bars refer to standard deviations.



Factors That Influence Intraword Variability

Our study revealed that several factors influence intraword variability.

Child-Related Factors

Age. Our findings are consistent with numerous reports indicating that, as children get older, intraword variability decreases (Holm et al., 2007; Jones, 2020; Macrae, 2013; Sosa, 2015). Intraword variability was 45% in the youngest group of children compared to 19% in the oldest group. Many factors may be responsible for reduced variability with age. As indicated in Table 5, there were significant correlations between age on the one hand and PCC and oral motor ability on the other. As age increased, children's consonant accuracy and scores on the OMA increased. Since both factors were also related to intraword variability, it could be assumed that children's increasing accuracy in speech sound production contributed to a reduction in intraword variability across age. Other factors apart from speech sound production ability (e.g., developing cognitive abilities, expanding vocabulary) may also be responsible for reduced variability with age.

Expressive vocabulary. Our analyses indicated that expressive vocabulary score was significantly correlated to intraword variability (see Table 5) as has been reported by other researchers (Sosa & Stoel-Gammon, 2012), but when entered into the statistical model, alongside other variables, it did not prove to be a significant predictor. This is in contrast to the findings of Macrae (2013) and Macrae and Sosa (2015) who found vocabulary to be a significant predictor of variability. One of the possible reasons why vocabulary did not predict intraword variability is because its validity was reduced due to the inclusion of bilingual children. Many studies indicate that bilingual children score lower than monolingual children on vocabulary tests when tested in one language (Hoff et al., 2012), a finding that was confirmed in the current study: The vocabulary scores of the bilingual group were lower than the monolingual group. It may be the case that a vocabulary score that takes into account the bilingual's other language(s) may have better reflected the relationship of vocabulary to intraword variability.

Gender. Our findings revealed that gender improved model fit to data. When all factors were controlled, boys obtained higher variability scores than girls. In a simple *t*-test analysis when all factors were not controlled, there were no gender differences in intraword variability. As noted in the literature review, there are conflicting findings on whether gender influences variability scores, which, in part, could be due to different statistical methods among studies and the effects of uncontrolled factors. When gender differences are reported, it is boys who are the most susceptible to having poorer speech sound production abilities (McLeod, 2009) and this study was no exception in finding greater intraword variability in boys than girls.

Bilingual status. Based on the findings of Sosa and Bunta (2019), who did not find differences in intraword variability rates between bilingual Spanish–English and monolingual English-speaking children, we predicted that bilingualism would not influence children's whole-word variability in French. In general, studies comparing the phonological development of monolingual and bilingual children have not shown major differences between the two groups on global phonological abilities such as PCC, percent vowels correct, whole-word proximity, or consonant inventory size (Fabiano-Smith & Goldstein, 2010; Goldstein & Bunta, 2012; Goldstein et al., 2005; Grech & Dodd, 2008). Nevertheless, we considered the alternative possibility that bilingualism may have an impact on intraword variability, either leading to reduced variability relative to monolinguals because bilinguals have more tightly specified phonological representations or increased variability relative to monolinguals because bilinguals need to realize sounds across two phonetic inventories, some sounds being phonetically similar but not exactly the same. This may be a productive challenge for them.

Indeed, in a simple *t*-test analysis, bilinguals did not differ from monolinguals in intraword variability. However, in our regression model, while controlling for factors such as vocabulary level, gender, years of maternal education, and speech sound ability, bilingual children were found to have higher levels of intraword variability than monolingual children. This result is consistent with reports indicating that bilinguals may be more variable than monolinguals in their phonological development (Core & Scarpelli, 2015; Hambly et al., 2013). Gildersleeve-Neumann et al. (2008) reported higher variability in consonant errors in bilingual English–Spanish children compared to monolingual English and predominantly English children; Kehoe and Havy (2019) observed greater variability in vowel accuracy in bilingual French-speaking compared to monolingual children. Although these findings report variability on a group level, it is likely that this variability is present at the individual level as well. In bilingual language development, one language may exert an influence upon the other, a phenomenon known as cross-linguistic interaction. We hypothesize that cross-linguistic interaction in the form of phonetic transfer effects (e.g., French /r/ variably produced as French [ʁ] or Spanish [r]) may underlie some of the increased variability observed in the bilinguals' speech (see also Hambly et al., 2013).

Speech sound production ability. An important finding of the study was that speech sound ability influenced intraword variability. Previous studies that have employed an articulation test (Macrae & Sosa, 2015) or PCC based on broad phonetic transcription of a spontaneous language sample (Martikainen et al., 2019) have not found speech sound ability to correlate with variability. In this study, we selected three different tests to tap children's speech sound ability: one test tapping children's phonological precision and the other two tapping children's articulatory or speech motor control. Our results indicated that children's phonological precision, as measured by PCC, and children's articulatory control, as measured in an OMA, predicted intraword variability. The DDK results did not correlate with whole-word variability nor did they correlate with age or the OMA results.

We hypothesize that the most likely explanation as to why DDK did not pattern with the OMA results is methodological. The DDK task was more complicated for the children than the OMA. Many children needed encouragement to produce repeated productions of the DDK sequences, often stopping and restarting during the test, thus compromising the time-based measure. Every attempt was made to adapt the task for children as young as age 2;6 (e.g., pointing to an image of a caterpillar); however, more work would be needed to develop a valid protocol. It was also the only test in the battery in which three children did not comply with the task. In contrast, children had no difficulty executing gestures in the OMA.

Why is speech sound ability related to intraword variability? It seems logical that if children are highly accurate in their speech sound ability, they are unlikely to be variable. This was observed in the current study: Children who obtained high PCC scores had low variability scores. However, if children are highly inaccurate, they could be either variable (i.e., variable with hits or no hits) or not variable (i.e., consistently incorrect). Studies that have coded children's variable response types have found the response type of consistently incorrect to be infrequent (Martikainen et al., 2019; Sosa, 2015). The most frequent response type is variable with no hits, suggesting that if children are inaccurate, they are likely to be variable. Many of the factors that contribute to inaccuracy in speech sound ability, some of which were tested in the current study such as age and vocabulary level, also underlie variability. We conclude that the failure to find a significant correlation in the past between speech sound ability and variability may relate to the type of tests used, which were not sensitive enough to tap into speech sound ability in typically developing children.

Word-Related Factors

With respect to word-related factors, three factors were associated with intraword variability: phonological complexity, PND, and WF.

Phonological complexity. Our findings support those of Macrae (2013) and Sosa and Stoel-Gammon (2012) in showing that the phonological complexity of the target word strongly influences intraword variability. They determined phonological complexity by averaging across age of acquisition of individual consonants, whereas we employed a metric that assigned points to different aspects of phonological complexity known to pose difficulty for children in early production. They include segmental complexity (e.g., presence of dorsals, fricatives, and liquids), POA variegation, syllable structure complexity (i.e., presence of codas, tautosyllabic, and heterosyllabic clusters), and presence of long words. It could be posited that such a metric may be an even more effective way to measure phonological complexity because it takes more than segmental complexity into consideration. Our results indicated that the three most variable words were long (ordinateur: mean variability = 57.89%; hélicoptère: 52.63%) and had complex syllable structure (zèbre: 43.59%). The three least variable words were short and had less complex syllable structure (clé: 7.69%; poisson: 7.69%; papillon: 10%).

Word length as a parameter is incorporated into the measure of phonological complexity. We also examined the separate effect of number of syllables. As seen in Figure 3, children have higher rates of intraword variability in four syllables compared to other words. Nevertheless, when entered into our statistical model, number of syllables was not a significant predictor of word variability.

PND and WF. Similar to Sosa and Stoel-Gammon (2012) and Jones (2020), we observed that PND of the target word influenced intraword variability. Words having a high number of phonological neighbors such as robe, fleur, and carotte had low levels of variability (i.e., 15.79%, 13.15%, and 12.82%, respectively). Words with few phonological neighbors such as crocodile, toboggan, and arrosoir had higher levels of variability (i.e., 30.77%, 32.55%, and 30.77%, respectively). We assume that the effect of PND on variability reflects the strength of phonological representations. In order to avoid confusion in perception and production with more similar-sound words, phonological representations develop first in high-density neighborhoods (Storkel, 2002; Sosa & Stoel-Gammon, 2012). Furthermore, we assume that more specified phonological representations are associated with lower intraword variability. In contrast, the expected effect that highly frequent words are associated with low variability was not confirmed in this study. In the correlational analyses, WF was not significantly related to intraword variability and in the regression model; while controlling for PND, complexity, and syllable number, high WF was modestly associated with high intraword variability, a trend opposite to the one reported by Sosa and Stoel-Gammon (2012). Several factors may explain the reduced role of WF in the current study. We employed frequency values extracted from an adult database, which may not reflect the frequency of words to children. It is likely that the words pertaining to animals (grenouille "frog," crocodile "crocodile," zèbre "zebra," éléphant "elephant"), of which there were several in our variability test, may be more frequently heard by children than by adults. In addition, although we selected words to have a range of WF values, there were only two that were of very high frequency (e.g., chien "dog," porte "door") based on adult norms. The use of adult norms and the reduced range of WF in the current data set may underlie the lack of a WF effect.

Limitations

This study has several limitations, particularly relating to the low numbers and heterogeneous nature of the participants. Although we controlled for uneven numbers of males and females as well as monolinguals and bilinguals across the different age groups in the statistical model, it would be preferable to have balanced numbers of males and females as well as monolinguals and bilinguals in a future study. We also determined whether children were typically developing on the basis of parent report; however, direct assessment of speech, language, and hearing would provide for more rigorous selection criteria. Another factor that should be considered in future studies is counterbalancing the order of test administration to prevent any uncontrolled factors influencing intraword variability. In addition, studies suggest differences in consonant accuracy between spontaneous and imitated productions (McLeod & Masso, 2019) so it cannot be excluded that there may be differences in intraword variability between spontaneous and imitated productions, a finding that should be confirmed in future studies. Finally, information on the PND and WF of the target words in the variability test was based on values taken from adult corpora. It would be important to determine whether the same results hold when values are taken from child corpora.

Source of Variability

One major message stemming from studies by Macrae, Sosa, and colleagues on intraword variability in typically developing children is that the source of intraword variability is not just at the level of speech motor planning (Macrae, 2013; Macrae & Sosa, 2015; Sosa, 2015). The presence of word variability in young typically developing children and its relation to factors such as vocabulary knowledge and the lexical and phonological characteristics of the target word suggest that intraword variability reflects difficulty at higher levels of speech processing. Our findings are consistent with this general message since we documented variability in typically developing children and observed it to be related to expressive vocabulary (although it did not emerge as a significant factor in our statistical model) and to the lexical and phonological characteristics of the target word. Nevertheless, our findings do not contradict the association of intraword variability with speech motor planning since the results of an OMA significantly predicted intraword variability. In reality, several factors, which reflect levels higher up (e.g., phonological representations, phonological planning) and lower down (speech motor planning) in the speech processing system (see model by Stackhouse & Wells, 1997), appear to influence intraword variability rate.

We advocate caution in using intraword variability as a differential diagnostic marker of CAS. Macrae and Sosa (2015) have already pointed out that small vocabularies, often observed in children with CAS, should be ruled out as a source of intraword variability. Our own findings indicate that low consonant accuracy as measured by PCC, a characteristic likely to be present in children with CAS, may also be associated with high intraword variability. In summary, children with CAS may present with high levels of intraword variability; however, the source of the variability may not necessarily be a speech motor deficit.

Finally, we attribute differences in intraword variability rate in typically developing English and French children to the phonological characteristics of the target language; however, it is unclear whether such target language effects will be observed in disordered populations. Martikainen et al. (2020) found comparable variability rates in Finnish children with speech sound disorders relative to English children with speech sound disorders (Macrae et al., 2014), whereas variability rates were not comparable in typically developing Finnish relative to English children.[4] Sosa and Bunta (2019) observed a significant main effect of language on intraword variability in bilingual Spanish–English children with normal hearing and with cochlear implants. Nevertheless, the influence of language was more pronounced in the normal hearing than in the implant group. It may be the case that target language effects are reduced in disordered populations, a finding that should be confirmed in future studies.

Conclusions

This study examined intraword variability in typically developing French-speaking monolingual and bilingual children. Our findings showed that variability was present to a lesser degree in French children than has been reported in English children, a finding we ascribe to linguistic differences between French and English. Our findings also revealed that several child- and word-related factors influenced intraword variability. Variability decreased as children got older and as their speech sound abilities improved. Variability decreased when children produced words with lower phonological complexity and more phonological neighbors. Variability was more often associated with bilinguals than monolinguals and with boys than girls. These findings suggest that the factors that condition intraword variability are multifaceted and cannot be localized to a single source.

Acknowledgments

This study is based on Emilie Cretton's master thesis in Logopédie defended at the "Université de Genève." We would like to thank Ilona Bray, Célia Hedbaut, and Oxana Kurzenova for their help in the reliability aspect of the study and Eric Menetre for his advice regarding statistics. In addition, we would like to thank the personnel and teachers at "EVE Espèces de vie enfantine du secteur université" and in crèches in the canton of Valais for their collaboration in the recruitment of children.

References

American Speech-Language-Hearing Association. (2007). *Childhood apraxia of speech*. https://www.asha.org/practice-portal/clinical-topics/childhood-apraxia-of-speech/

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Beckman, M. E., Yoneyama, K., & Edwards, J. (2003). Language-specific and language-universal aspects of lingual obstruent productions in Japanese-acquiring children. *Journal of the Phonetic Society of Japan, 7,* 18–28.

Bernthal, J. E. (2009). *Articulation and phonological disorders. Speech sound disorders in children*. Pearson.

Betz, S. K., & Stoel-Gammon, C. (2005). Measuring articulatory error consistency in children with developmental apraxia of speech. *Clinical Linguistics & Phonetics, 19*(1), 53–66. https://doi.org/10.1080/02699200512331325791

Brosseau-Lapré, F., & Rvachew, S. (2014). Cross-linguistic comparison of speech errors produced by English- and French-speaking preschool-age children with developmental phonological disorders. *International Journal of Speech-Language Pathology, 16*(2), 98–108. https://doi.org/10.3109/17549507.2013.794863

[4]Martikainen et al. (2020) note, however, that they transcribed their data at the phonetic level (including distortions), whereas Macrae et al. (2014) transcribed their data at the broad phonemic level, so there may be greater differences in the two groups of disordered children (Finnish vs. English) if transcription methods are taken into consideration.

Brysbaert, M., Mandera, P., & Keuleers, E. (2018). The word frequency effect in word processing: An updated review. *Current Directions in Psychological Science, 27*(1), 45–50. https://doi.org/10.1177/0963721417727521

Byers-Heinlein, K., & Fennell, C. T. (2014). Perceptual narrowing in the context of increased variation: Insights from bilingual infants. *Developmental Psychobiology, 56*(2), 274–291. https://doi.org/10.1002/dev.21167

Coquet, F., Ferrand, P., & Roustit, J. (2009). *EVALO 2–6*. Ortho Edition.

Core, C., & Scarpelli, C. (2015). Phonological development in young bilinguals: Clinical implications. *Seminars in Speech and Language, 36*(2), 100–108. https://doi.org/10.1055/s-0035-1549105

de Castro, M. M., & Wertzner, H. F. (2011). Speech inconsistency index in Brazilian Portuguese-speaking children. *Folia Phoniatrica et Logopaedica, 63*(5), 237–241. https://doi.org/10.1159/000323183

Delattre, P., & Olsen, C. (1969). Syllabic features and phonic impression in English, German, French and Spanish. *Lingua, 22,* 160–175. https://doi.org/10.1016/0024-3841(69)90051-5

Dell, F. (1984). L'accentuation dans les phrases en français [Accentuation in French sentences]. In F. Dell, D. Hirst, & J.-R. Vergnaud (Eds.), *Forme Sonore du Langage* (pp. 65–122). Hermann.

Dodd, B. (1995). *Differential diagnosis and treatment of children with speech disorder* (*1st ed.*). Wiley.

Dodd, B. (2005). *Differential diagnosis and treatment of children with speech disorder* (*2nd ed.*). Wiley.

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* (*2nd ed.*, pp. 44–70). Wiley.

Fabiano-Smith, L., & Goldstein, B. A. (2010). Phonological acquisition in bilingual Spanish–English speaking children. *Journal of Speech, Language, and Hearing Research, 53*(1), 160–178. https://doi.org/10.1044/1092-4388(2009/07-0064)

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*(2), 706–720. https://doi.org/10.1044/2017_AJSLP-16-0130

Fenson, L. (1993). *MacArthur Communicative Development Inventories: User's guide and technical manual*. Singular.

Ferguson, C. A., & Farwell, C. B. (1975). Words and sounds in early language acquisition. *Language, 51*(2), 419–439. https://doi.org/10.2307/412864

Gildersleeve-Neumann, C. E., Kester, E. S., Davis, B. L., & Peña, E. D. (2008). English speech sound development in preschool-aged children from bilingual English–Spanish environments. *Language, Speech, and Hearing Services in Schools, 39*(3), 314–328. https://doi.org/10.1044/0161-1461(2008/030)

Goldman, R., & Fristoe, M. (2000). *Goldman-Fristoe Test of Articulation–Second Edition*. APA PsycTests.

Goldstein, B. A., & Bunta, F. (2012). Positive and negative transfer in the phonological systems of bilingual speakers. *International Journal of Bilingualism, 16*(4), 388–401. https://doi.org/10.1177/1367006911425817

Goldstein, B. A., Fabiano, L., & Washington, P. S. (2005). Phonological skills in predominantly English-speaking, predominantly Spanish-speaking, and Spanish–English bilingual children. *Language, Speech, and Hearing Services in Schools, 36*(3), 201–218. https://doi.org/10.1044/0161-1461(2005/021)

Grech, H., & Dodd, B. (2008). Phonological acquisition in Malta: A bilingual language learning context. *International Journal of Bilingualism, 12*(3), 155–171. https://doi.org/10.1177/1367006908098564

Hambly, H., Wren, Y., McLeod, S., & Roulstone, S. (2013). The influence of bilingualism on speech production: A systematic review. *International Journal of Language & Communication Disorders, 48*(1), 1–24. https://doi.org/10.1111/j.1460-6984.2012.00178.x

Hoff, E., Core, C., Place, S., Rumiche, R., Señor, M., & Parra, M. (2012). Dual language exposure and early bilingual development. *Journal of Child Language, 39*(1), 1–27. https://doi.org/10.1017/S0305000910000759

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*(4), 467–486. https://doi.org/10.1080/13682820600988967

Jakielski, K. J. (2000). *Quantifying phonetic complexity in words: An experimental index*. Annual Child Phonology Conference, Cedar Falls, IA, United States.

Jones, S. D. (2020). Accuracy and variability in early spontaneous word production: The effects of age, frequency and neighbourhood density. *First Language, 40*(2), 128–150. https://doi.org/10.1177/0142723719894768

Keffala, B., Barlow, J. A., & Rose, S. (2018). Interaction in Spanish–English bilinguals' acquisition of syllable structure. *International Journal of Bilingualism, 22*(1), 16–37. https://doi.org/10.1177/1367006916644687

Kehoe, M., & Havy, M. (2019). Bilingual phonological acquisition: The influence of language-internal, language-external, and lexical factors. *Journal of Child Language, 46*(2), 292–333. https://doi.org/10.1017/S0305000918000478

Kern, S., & Gayraud, F. (2010). *The French Inventory of Communicative Development*. Editions La Cigale.

Lleó, C. (2006). The acquisition of prosodic word structures in Spanish by monolingual and Spanish–German bilingual children. *Language and Speech, 49*(2), 205–229. https://doi.org/10.1177/00238309060490020401

Lleó, C., & Demuth, K. (1999). Prosodic constraints on the emergence of grammatical morphemes: Cross-linguistic evidence from Germanic and Romance languages. In A. Greenhill, H. Littlefield, & C. Tano (Eds.), *Proceedings of the 23rd annual Boston University Conference on Language Development* (pp. 407–418). Cascadilla Press.

Luce, P. A., & Pisoni, D. B. (1998). Recognizing spoken words: The neighborhood activation model. *Ear and Hearing, 19*(1), 1–36. https://doi.org/10.1097/00003446-199802000-00001

MacLeod, A. N., Sutton, A., Trudeau, N., & Thordardottir, E. (2011). The acquisition of consonants in Québécois French: A cross-sectional study of pre-school aged children. *International Journal of Speech-Language Pathology, 13*(2), 93–109. https://doi.org/10.3109/17549507.2011.487543

Macrae, T. (2013). Lexical and child-related factors in word variability and accuracy in infants. *Clinical Linguistics & Phonetics, 27*(6–7), 497–507. https://doi.org/10.3109/02699206.2012.752867

Macrae, T., & Sosa, A. (2015). Predictors of token-to-token inconsistency in preschool children with typical speech-language development. *Clinical Linguistics & Phonetics, 29*(12), 922–937. https://doi.org/10.3109/02699206.2015.1063085

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*(1), 27–35. https://doi.org/10.1044/1058-0360(2013/12-0037)

Martikainen, A.-L., Savinainen-Makkonen, T., & Kunnari, S. (2019). Intra-word consistency and accuracy in Finnish children aged 3–6 years. *Clinical Linguistics & Phonetics, 33*(9), 815–830. https://doi.org/10.1080/02699206.2019.1576770

Martikainen, A.-L., Savinainen-Makkonen, T., Laukkanen-Nevala, P., & Kunnari, S. (2020). Intra-word accuracy and consistency in Finnish-speaking children with speech sound disorder compared to their typically developing peers. *Clinical Linguistics & Phonetics, 34*(8), 718–733. https://doi.org/10.1080/02699206.2019.1696610

Martinez Perez, T., Masson, P., & James, E. G. (2015). Construction et normalization d'épreuves aidant au diagnostic de la dyspraxie verbale en Belgique Francophone [Construction and standardization of tests to aid in the diagnosis of childhood dyspraxia of speech in French-speaking Belgium]. *Speech Therapy Rehabilitation, 263,* 149–164.

Masson, P. (2017). *Etude de la spécificité d'épreuves langagières chez des enfants présentant une dyspraxie verbale versus un trouble phonologique* [Study of the specificity of language tests in children with verbal dyspraxia versus phonological disorder] [Unpublished master's thesis, University of Liège].

McLeod, S. (2009). Speech sound acquisition. In J. Bernthal, N. Bankson, & P. Flipsen (Eds.), *Articulation and phonological disorders: Speech sound disorders in children* (*6th ed.*, pp. 63–120). Pearson.

McLeod, S., & Hewett, S. R. (2008). Variability in the production of words containing consonant clusters by typical 2- and 3-year-old children. *Folia Phoniatrica et Logopaedica, 60*(4), 163–172. https://doi.org/10.1159/000127835

McLeod, S., & Masso, S. (2019). Screening children's speech: The impact of imitated elicitation and word position. *Language, Speech, and Hearing Services in Schools, 50*(1), 71–82. https://doi.org/10.1044/2018_LSHSS-17-0141

Menn, L., & Matthei, E. (1992). The "two-lexicon" approach of child phonology: Looking back, looking ahead. In C. Ferguson, L. Menn, & C. Stoel-Gammon (Eds.), *Phonological development: Models, research, implications* (pp. 211–248). York Press.

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 J. L. Metsala & L. C. Ehri (Eds.), *Word recognition in beginning literacy* (pp. 89–120). Erlbaum.

Murray, E., McCabe, P., Heard, R., & Ballard, K. J. (2015). Differential diagnosis of children with suspected childhood apraxia of speech. *Journal of Speech, Language, and Hearing Research, 58*(1), 43–60. https://doi.org/10.1044/2014_JSLHR-S-12-0358

New, B., Brysbaert, M., Veronis, J., & Pallier, C. (2007). The use of film subtitles to estimate word frequencies. *Applied Psycholinguistics, 28*(4), 661–677. https://doi.org/10.1017/S014271640707035X

Nicolaidis, K., Edwards, J., Beckman, M., & Tserdanelis, G. (2003). *Acquisition of lingual obstruents in Greek*. Proceedings of the 6th International Conference of Greek Linguistics, Rethymno, Crete, Greece.

R Development Core Team. (2015). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing.

Ramus, F., Nespor, M., & Mehler, J. (1999). Correlates of linguistic rhythm in the speech signal. *Cognition, 73*(3), 265–292. https://doi.org/10.1016/S0010-0277(99)00058-X

Roark, B., & Demuth, K. (2000). Prosodic constraints and the learner's environment: A corpus study. In S. C. Howell, S. A. Fish, & T. Keith-Lucas (Eds.), *Proceedings of the 24th annual Boston University Conference on Language Development* (pp. 597–608). Cascadilla Press.

Rose, Y., & MacWhinney, B. (2014). The PhonBank initiative. In J. Durand, J. U. Gut, & G. Kristoffersen (Eds.), *The Oxford handbook of corpus phonology* (pp. 380–401). Oxford University Press.

Rose, Y., MacWhinney, B., Byrne, R., Hedlund, G., Maddocks, K., O'Brien, P., & Wareham, T. (2006). Introducing Phon: A software solution for the study of phonological acquisition. In D. Bamman, T. Magnitskaia, & C. Zaller (Eds.), *Proceedings of the 30th Boston University Conference on Language Development* (pp. 489–500). NIH Public Access.

Rvachew, S., Leroux, E., & Brosseau-Lapré, F. (2014). Production of word-initial consonant sequences by Francophone preschoolers with a developmental phonological disorder. *Canadian Journal of Speech-Language Pathology and Audiology, 37*(4), 252–267.

Savinainen-Makkonen, T. (2000). Learning long words: A typological perspective. *Language and Speech, 43*(2), 205–225. https://doi.org/10.1177/00238309000430020401

Schwartz, R. G., & Goffman, L. (1995). Metrical patterns of words and production accuracy. *Journal of Speech and Hearing Research, 38*(4), 876–888. https://doi.org/10.1044/jshr.3804.876

Sosa, A. V. (2015). Intraword variability in typical speech development. *American Journal of Speech-Language Pathology, 24*(1), 24–35. https://doi.org/10.1044/2014_AJSLP-13-0148

Sosa, A.V., & Bunta, F. (2019). Speech production accuracy and variability in monolingual and bilingual children with cochlear implants: A comparison to their peers with normal hearing. *Journal of Speech, Language, and Hearing Research, 62*(8), 2601–2616. https://doi.org/10.1044/2019_JSLHR-S-18-0263

Sosa, A. V., & Stoel-Gammon, C. (2012). Lexical and phonological effects in early word production. *Journal of Speech, Language, and Hearing Research, 55*(2), 596–608. https://doi.org/10.1044/1092-4388(2011/10-0113)

Stackhouse, J., & Wells, B. (1997). *Children's speech and literacy impairment*. Whurr Publishers.

Stoel-Gammon, C. (2010). The word complexity measure: Description and application to developmental phonology and disorders. *Clinical Linguistics & Phonetics, 24*(4–5), 271–282. https://doi.org/10.3109/02699200903581059

Storkel, H. L. (2002). Restructuring of similarity neighbourhoods in the developing mental lexicon. *Journal of Child Language, 29*(2), 251–274. https://doi.org/10.1017/S0305000902005032

Tuller, L. (2015). Clinical use of parental questionnaires in multilingual contexts. In S. Armon-Lotem, J. de Jong, & N. Meir (Eds.), *Assessing multilingual children: Disentangling bilingualism from language impairment* (pp. 301–330). Walter de Gruyter GmbH. https://doi.org/10.21832/9781783093137-013

Appendix A

Information on the Monolingual Children Including Gender, Age, and Number of Years of Supplementary Schooling by the Mother

Participant ID	Gender	Age	Years of supp. schooling
1	G	2;7	9
2	G	3;8	6
3	B	3;5	9
4	B	3;5	10
6	G	2;10	10
8	G	2;10	12
9	B	2;10	9
13	G	2;11	8
15	B	4;1	9
17	B	3;0	12
18	B	3;10	3
20	B	3;10	9
21	B	4;1	8
22	B	3;10	9
23	B	4;3	15
24	B	4;2	7.5
25	B	4;2	11
29	G	4;0	0
32	B	3;5	3
33	G	3;5	8
34	G	3;11	2.5

Note. G = girl; B = boy.

Appendix B

Information on the Bilingual Children Including Gender; Age; Languages Spoken by Mother, Father, and Prominent Caretaker; Dominant Language; and Number of Years of Supplementary Schooling by the Mother

Participant ID	Gender	Age	Language of mother	Language of father	Language of caretaker	Dominant language	Years of supp. schooling
5	B	3;3	French	Portuguese	—	French	12
7	B	2;11	Polish	Polish	French	Polish	8
10	B	3;9	Spanish	Spanish	French	Spanish	12
11	B	4;6	Italian	Italian	French	Balanced	13
12	G	4;1	Italian French	Italian French	Swiss German	Balanced	8
14	B	3;9	Portuguese	French	—	French	11
16	G	3;3	Arab	French	—	French	12
19	G	4;4	Italian	Portuguese	French	Balanced	5
27	B	2;8	French Italian	French	—	French	8
28	G	2;8	French	English	—	French	6
30	B	4;8	French	Albanese	—	French	3
31	B	4;3	Albanese French	Albanese French	—	French	8
35	G	3;8	French	French	Albanese	French	7
36	B	3;3	French Polish	French	—	French	10
37	B	3;9	French Italian	French Italian	—	French	9
38	B	3;9	German	French	—	French	15
39	B	4;6	Swiss German	French	—	French	6
40	G	4;1	Swiss German	French	—	French	8
41	G	2;6	French Albanese	Albanese	—	French	9

Note. G = girl; B = boy.

Appendix C

Description of the Characteristics of the 25 Stimulus Words of the Variability Test Including Word Frequency, Number of Phonological Neighbors, Number of Syllables, Syllable Structure, and IPC (phonetic complexity) Value

Stimulus words	Word frequency	No. of phon. neighbors	No. of syllables	Syllable structure[a]	IPC
porte	288.39	23	1	CVCC	5
chien	223.53	12	1	CGV	2
robe	72.72	20	1	CVC	4
arbre	49.29	5	1	VCCC	8
fleur	25.2	18	1	CCVC	7
clé	14.61	19	1	CCV	4
zèbre	2.65	1	1	CVCC	6
poisson	53.61	12	2	CGV.CV	3
lunettes	31.61	4	2	CV.CVC	2
fromage	25.68	3	2	CCV.CVC	7
étoile	21.65	5	2	V.CGVC	4
grenouille	5.74	2	2	CCV.CVC	6
tracteur	2.87	2	2	CCVC.CVC	9
carotte	2.45	17	2	CV.CVC	5
pantalon	31.49	0	3	CV.CV.CV	3
éléphant	10.71	3	3	V.CV.CV	4
papillon	8.12	2	3	CV.CV.CV	2
crocodile	6.14	0	3	CCV.CV.CVC	9
balançoire	1.93	0	3	CV.CV.CGVC	8
toboggan	0.59	0	3	CV.CV.CV	3
arrosoir	0.37	0	3	V.CV.CGVC	9
escalier	20.91	1	3	VCCVCGV	7
ordinateur	30.2	1	4	VC.CV.CV.CVC	8
hélicoptère	10.98	0	4	V.CV.CVC.CVC	8
aspirateur	4.33	1	4	VC.CV.CV.CVC	9

Note. IPC = Index of Phonetic Complexity.

[a]C = consonant; V = vowel; G = glide.

Appendix D

Words in the Variability Test Including Phonetic Form and English Translation

arbre /aʁbʁ/ "tree," arrosoir /aʁozwaʁ/ "watering can," aspirateur /aspiʁatœʁ/ "vacuum cleaner," balançoire /balɑ̃swaʁ/ "see-saw," carotte /kaʁɔt/ "carrot," chien /ʃjɛ̃/ "dog," clé /kle/ "key," crocodile /kʁɔkodil/ "crocodile," éléphant /elefɑ̃/ "elephant," escalier /ɛskalje/ "stairs," étoile /etwal/ "star," fleur /flœʁ/ "flower," fromage /fʁomaʒ/ "cheese," grenouille /gʁənuj/ "frog," hélicoptère /elikoptɛʁ/ "helicopter," lunettes /lynɛt/ "glasses," ordinateur /ɔʁdinatœʁ/ "computer," pantalon /pɑ̃talɔ̃/ "pants/trousers," papillon /papijɔ̃/ "butterfly," poisson /pwasɔ̃/ "fish," porte /pɔʁt/ "door," robe /ʁɔb/ "dress," toboggan /tobogɑ̃/ "slide," tracteur /tʁaktœʁ/ "tractor," zèbre /zɛbʁ/ "zebra."

Appendix E

Statistical Model

Model = glmer (Var ~ age + gender + maternaled + bilingual + vocab + OMA + PCC + complexity + syllno + ND + logfreq + (1|word) + (1|ID)

Random effects.

Group's name	Variance	*SD*
ID (Intercept)	0.02726	0.1651
Word (Intercept)	0.08178	0.2860

Note. Number of observations: 968, groups: ID, 40; word, 25.

Fixed effects.

Variable	Estimate	*SE*	*z* value	Pr(>|z|)
(Intercept)	−0.93701	0.34719	−2.699	0.006958**
age	−0.42254	0.11183	−3.779	0.000158***
gender	0.43835	0.18788	2.333	0.019639*
maternaled	−0.32043	0.10847	−2.954	0.003135**
bilingual	−0.59828	0.20509	−2.917	0.003532**
vocab	0.06085	0.12047	0.505	0.613479
OMA	−0.21946	0.09385	−2.338	0.019371*
PCC	−0.40836	0.10161	−4.019	5.85e-05***
complexity	0.43106	0.10801	3.991	6.58e-05***
sylno	−0.04540	0.13147	−0.345	0.729878
PND	−0.52492	0.15623	−3.360	0.000779***
WF	0.25344	0.12078	2.098	0.035872*

Note. Significant codes: ***$p < .001$; **$p < .01$; *$p < .05$; $p < .1$. The variables are age, gender, maternal education, bilingual status, vocabulary score, OMA (oral motor assessment), PCC (percent consonants correct), complexity (phonological complexity), syllable number, PND (phonological neighborhood density), and WF (word frequency).