

ARTICLE

Which phonological and lexical factors best influence whether a word is produced and pronounced well?

Margaret KEHOE*, Aya ABU LABAN and Romane LESPINASSE

Université de Genève, Switzerland

*Corresponding author. Margaret Kehoe, Faculté de psychologie et des sciences de l'éducation, Université de Genève, 28, bd du Pont-d'Arve, 1205 Genève, Switzerland. Email: Margaret.Winkler-Kehoe@unige.ch

(Received 30 June 2022; revised 16 November 2022; accepted 06 January 2023)

Abstract

This study examines lexical and phonological factors that influence word production and pronunciation. Specifically, we investigate whether phonological production (measured by percent consonants correct) contributes to word production and pronunciation over and above the properties of the target words (e.g., word frequency, neighborhood density, and phonetic complexity). Forty French-speaking monolingual and bilingual children, aged 1;11 to 3;1, participated in a spontaneous language sample and were administered a naming and a nonword repetition task. Their parents filled out the MacArthur Communicative Developmental Inventory (MCDI) and rated their children's pronunciation on an experimental version of the MCDI. Statistical models indicated that word frequency and the phonetic complexity of the target words influenced whether a word was produced. These factors along with neighborhood density and the children's production capacities influenced whether a word was pronounced poorly or well. Findings indicate that parents can provide reliable information on the word pronunciation of their children.

Keywords: vocabulary development; phonological production; bilingualism

Introduction

In recent years, investigators have examined what phonological and/or lexical factors best explain vocabulary development in children, aged two to three years (Hansen, 2017; Jones & Brandt, 2019a, 2019b; Kehoe, Patrucco-Nanchen, Friend, & Zesiger, 2018; Kehoe, Patrucco-Nanchen, Friend, & Zesiger, 2020; Maekawa & Storkel, 2006; Stokes, 2014; Stokes, Bleses, Basbøll, & Lambertsen, 2012a; Stokes, Kern, & dos Santos, 2012b; Stokes, de Bree, Kerkhoff, Momenian, & Zamuner, 2019; Storkel, 2004, 2009). Studies show that the psycholinguistic characteristics of the target words children produce, such as their neighbourhood density (ND), word frequency (WF), and

© The Author(s), 2023. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<http://creativecommons.org/licenses/by/4.0>), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

phonetic complexity¹, account for a significant proportion of the variance in vocabulary size (Kehoe et al., 2018, 2020; Stokes, 2010, 2014; Stokes et al., 2012a, 2012b, 2019). Studies also show that children's phonological production skills, as determined by phonetic inventory size or percent consonants correct (PCC), explain unique variance in vocabulary size; however, the contribution of production variables tends to be small in comparison with the psycholinguistic properties of the target words. The few studies that have explored the role of phonological production in lexical development have mainly employed measures extracted from spontaneous language samples (Kehoe et al., 2018, 2020). However, a spontaneous language sample may underestimate children's production abilities since they may exhibit lexical selection strategies, choosing the words they can produce (Ferguson & Farwell, 1975).

The current study explores the role of phonological production in vocabulary development. Specifically, we examine whether other methods of measuring phonological production (e.g., single-word naming or non-word repetition task) are more sensitive indicators of the influence of phonological production on vocabulary development than one based on a spontaneous language sample. One reason why previous studies have yielded a modest role of phonological production is that vocabulary knowledge is determined by a binary outcome: a word is produced or not. Here, we also explore whether a more fine-grained measure of vocabulary development which takes into account "pronunciation goodness" leads to a stronger role of phonological production in vocabulary development (Jones & Brandt, 2019b). Parents will be asked to judge the phonological production skills of their children on an experimental version of the MacArthur Communicative Developmental Inventory (MCDI; Fenson, Dale, Reznick, Thal, Bates, Hartung, Pethick, & Reilly, 1993); they will report whether their child pronounces the word poorly, adequately, or well (Jones & Brandt, 2019b). Thus, we will examine what factors influence whether a word is indicated as produced or not (binary outcome), and what factors influence whether it is pronounced poorly, adequately, well, or not at all (fine-grained outcome). This study extends previous work on phonological-lexical relations in young children by including bilingual along with monolingual children. In the following sections of the *Introduction*, we outline those studies that have examined the influence of phonological and lexical variables on vocabulary development in children. We complete the *Introduction* with a statement of the research questions.

Influence of lexical and phonological variables on vocabulary development

Target word variables: ND, WF, and phonetic complexity

In this study, we distinguish between target word variables, which refer to the lexical and phonological properties of the words in children's lexicons, and phonological production variables, which are based on phonological behavioural measures. The two target word variables most frequently studied, when examining factors that influence vocabulary size, are ND and WF. ND indicates the degree of phonological similarity between a given word and a set of other words. A phonological neighbour is a word that differs from another word by substitution, deletion, or addition of a sound in any word position (Luce & Pisoni, 1998). Stokes and colleagues in a series of studies found that ND accounted for a

¹The term phonemic complexity may be more appropriate than phonetic complexity since the dimensions used to determine phonetic complexity are arguably more phonological than phonetic. However, we remain with the original term for convenience.

high proportion of variance in the vocabulary size of children acquiring English, French, and Danish (Stokes, 2014; Stokes et al., 2012a, 2012b). In all of their studies, they coded the ND of one-syllable words appearing in two-year-old children's lexicons, as based on the MCDI. The general finding was that as vocabulary size increases, children choose words with fewer phonological neighbours.

Stokes and colleagues found that WF accounted for only a small proportion of variance in vocabulary size when compared with ND; however, Kern and Dos Santos (2017) reanalyzed Stokes et al.'s (2012b) French data separating out one-syllable words according to grammatical class: nouns versus predicates (verbs and adjectives). They reported that WF accounted for a greater proportion of variance for one-syllable nouns than ND did and neither ND nor WF accounted for variance in vocabulary size for predicates. This finding is consistent with Goodman, Dale, and Li's (2008) observation that within a grammatical class, high frequency is associated with earlier acquisition.

Kehoe et al. (2020) also found that WF accounted for greater variance in vocabulary size than ND did in a study of 40 French-speaking children followed longitudinally from 1;10 through to 4;0. They coded the ND and WF of both one- and two-syllable nouns in the children's lexicons based on the French version of the MCDI. The WF of one-syllable words accounted for the greatest degree of variance at 1;10, 2;5, and 3;0, and the WF of two-syllable words accounted for the greatest variance at 4;0. ND accounted for small degrees of variance at 1;10 and 2;5 only. Similarly, Hansen (2017) reported that WF played a stronger role than ND once grammatical class was controlled in a study based on the Norwegian MCDI. In a more recent study, Stokes et al. (2019) found that both WF and ND contribute to whether a known word appears in Dutch-speaking children's spoken lexicons.

Several authors have coded the phonetic complexity of the target word and examined its influence on vocabulary development. Phonetic complexity has been operationalized in different ways, either by counting the number of phonemes in the target word – that is, its word length – or by coding the featural and structural properties of the target word (e.g., presence of dorsals, fricatives, codas, clusters, etc.) using the Index of Phonetic Complexity (IPC, Jakielski, 2000), or even by coding the phonological mean length of utterance (Gendler-Shalev, Ben-David, & Novogrodsky, 2021). Studies indicate that word length (in phonemes) is a significant predictor of a word's age of acquisition (Jones & Brandt, 2019b; Maekawa & Storkel, 2006; Storkel, 2004). Children learn short before long words. Similarly, the IPC² has been found to significantly predict vocabulary development. Kehoe et al. (2018) reported that the IPC accounted for 11% unique variance in the vocabulary size of children aged 2;5. In a later longitudinal study of vocabulary development, Kehoe et al. (2020) employed both word length and the IPC as measures of phonetic complexity. They reported that word length and the IPC were highly correlated (correlations ranging from 0.8 to 0.9), and, thus, could be assumed to be accounting for the same underlying construct. Nevertheless, word length accounted for unique variance in the younger children whereas the IPC accounted for unique variance in the older children, suggesting subtle differences in how these two variables influence lexicon size across age.

Phonological production variables

In another field of literature, investigators have explored the links between vocabulary size and phonological production, a general finding being that children with small vocabularies

²We will use the abbreviation "IPC" to refer to phonetic complexity based on the "IPC".

have poorer phonological production abilities than children with large vocabularies (Fletcher, Chan, Wong, Stokes, Tardif, & Leung, 2004; Kehoe, Chaplin, Mudry, & Friend, 2015; Paul & Jennings, 1992; Petinou & Okalidou, 2006). Rescorla and Ratner (1996), for example, found that children with small vocabularies had smaller consonantal and vocalic inventories and employed a more restricted set of syllable shapes than their typically developing peers. At the other end of the spectrum, Smith, McGregor, and Demille (2006) found that two-year-old children with large vocabularies were superior to their age-matched peers in terms of the number of singleton consonants correct and the percentage of final consonants correct. They evidenced fewer phonological processes such as cluster reduction and final consonant deletion.

Phonological memory, the capacity to recall sequences of meaningless sounds such as nonwords, is also highly correlated with vocabulary ability. In both learning a new word and repeating a non-word, the learner is required to form an acoustic representation of the underlying speech units. The representation has to be robust enough to support the articulation of a sequence of sounds and syllables. According to Hoff, Core, and Bridges (2008), phonological memory is the link between phonological knowledge (interpreted here as phonological representation) and word learning, and the association between phonological representation and memory provides an explanatory account for many of the observed relations between lexical and phonological development. That non-word repetition (NWR) is related to speech production experience has also been demonstrated. Keren-Portnoy, Vihman, DePaolis, Whitaker, and Williams (2010) found that phonological strings that have been previously articulated are represented more robustly in memory than strings that have been heard but not articulated.

Numerous studies show that NWR scores are highly correlated with vocabulary measures (Gathercole, 2006). In the case of young children, Hoff et al. (2008) found that NWR accuracy was significantly correlated with vocabulary size in a small group of two-year-olds. Stokes and Klee (2009) found that results on a NWR task were the strongest predictor of vocabulary scores among a variety of other demographic and behavioural variables in a large group of two-year-olds, accounting for 36% of the variance in vocabulary scores.

In sum, studies examining the relation between phonological production, phonological memory, and vocabulary have observed strong correlations between all variables; however, few studies have examined the amount of variance in vocabulary size accounted for by phonological production versus other phonological and lexical variables with the exception of two studies by Kehoe and colleagues. Kehoe et al. (2018) found that three factors accounted for variance in vocabulary size in 40 French-speaking children, aged 2;5. ND accounted for the bulk of variance (57%) followed by the IPC (11%), and phonological production (8%), as measured by the size of the syllable-initial phonetic inventory extracted from spontaneous language samples. In a later study, Kehoe et al. (2020) examined the phonological and lexical factors influencing vocabulary size in the same set of 40 children followed longitudinally through to 4;0 years. In contrast to the earlier study, they included WF, alongside ND and IPC. They observed that WF accounted for the bulk of variance at all age ranges. Phonological production accounted for unique variance at 2;5 and 3;0 years, although its contribution was very small (approximately 1-2%).

While it is not surprising that the phonological and lexical properties of the words themselves influence their acquisition, one may wonder whether other measures of phonological production could lead to a stronger role of production in accounting for vocabulary development. In both studies by Kehoe and colleagues, the two measures of

phonological production (phonetic inventory size and PCC) were extracted from spontaneous language samples. The use of a different sampling condition could lead to other results (DuBois & Bernthal, 1978; Morrison & Shriberg, 1992; Yeh & Liu, 2021). Morrison and Shriberg (1992) observed that the contribution of cognitive-linguistic and pragmatic processes is different in a conversational speech sample versus a single-word naming task. On the one hand, the act of formulating sentences from thought in a conversational situation results in a more cognitively demanding task; on the other hand, the liberty of choosing words and sentence structures within one's production capacities leads to a less demanding task. As first proposed by Ferguson and Farwell (1975), children may select only those words that fall within their production capacities. This may mean that a conversational sample does not sample all productions of sounds across all word positions; rather, only the sounds present in the words children choose to produce. In contrast, the purpose of a single-word naming task is to elicit a wide range of sounds across a variety of word positions. Nevertheless, several authors have reported that a connected speech sample may reveal a greater numbers of errors than a single-word naming task (DuBois & Bernthal, 1978; Glaspey, Wilson, Reeder, Tseng, & MacLeod, 2022; Morrison & Shriberg, 1992). This is because phonological processes which are influenced by context may emerge in connected speech but not in isolated word production.

Non-word repetition, while also measuring phonological production ability, taps other phonological skills – most specifically, phonological memory. Hoff et al. (2008) administered both a real-word and non-word repetition task to two-year-old children. They found a robust relation between vocabulary and NWR scores, even after partialling out variance due to the repetition of real words which presumably reflects the articulatory demands of the repetition task. They argued that this finding was evidence that NWR was tapping into “something more” than articulation accuracy – namely, the phonological memory component. It could be assumed that a measure that reflects both phonological production and memory may be more sensitive to lexical development than one that reflects production alone. Please note that we do not exclude other interpretations of phonological memory including G. Jones and Macken's (2018) view that it reflects domain-general associative learning processes operating on children's long-term linguistic experience or Vihman's (2022) proposal that it is the product of dynamic sensorimotor (perceptual and production) processes, including children's own vocal practice, that mediate the mapping of new forms onto phonological representations. Neither of these approaches considers that a dedicated short-term memory component is necessary to explain children's results on NWR tasks. While acknowledging these different views, we do not enter into the specifics of the origins of phonological memory in the current study.

Alternate ways of investigating the role of phonological production in vocabulary development should also be considered. Jones and Brandt (2019b), in discussing the limitations of the MCDI as a vocabulary measure, note that it does not provide information on children's pronunciation. Parents will still indicate that their child knows the word ‘giraffe’ regardless of whether he/she says ‘raffe’ or ‘giraffe’. Rather than using a binary approach “produces” or “does not produce”, they recommend more fine-grained response options such as “does not produce”, “produces poorly”, “produces adequately”, and “produces well”. The use of such options may provide more information on how phonological and lexical factors influence vocabulary development. As far as we are aware, no study to date has asked parents to judge the pronunciation of the words known by their children on the MCDI, something we intend to do in the current study.

Monolingual versus bilingual children

Most of the studies that have examined the influence of phonological and lexical factors on vocabulary development have tested monolingual children only (see however Dos Santos & Kern, 2015). Languages vary in their syllable and word structure which means that target word properties such as neighbourhood density and phonetic complexity may exert different influences on vocabulary development across languages. For example, children speaking languages with many long words (i.e., three-syllables or more) tend to have less dense neighbourhoods than children speaking languages with many monosyllables such as English (Stoel-Gammon, 2011). We focus on French vocabulary acquisition in this study, but it cannot be excluded that due to cross-linguistic interaction, target word properties influence vocabulary development differently in monolingual versus bilingual children. We consider this possibility in the current study.

Statement of research questions

In sum, several studies indicate that the psycholinguistic characteristics of the words children produce influence their acquisition, and a small number of studies show that children's own phonological abilities, as determined from measures extracted from spontaneous language samples, explain unique variance in vocabulary size; however, there has been little research on whether other measures of phonological production are more sensitive to the influence of phonology on vocabulary acquisition. The current study examines whether different measures of phonological production account for vocabulary development over and above that which is accounted for by the psycholinguistic characteristics of the target word. This study extends previous research by asking parents to judge the pronunciation of their children, and by including both monolingual and bilingual children.

The study has three aims. The first aim is to examine which set of phonological and lexical factors best influence vocabulary development. Factors include target word (ND, WF, word length, and IPC), and phonological production variables (PCC), the latter based on three sampling techniques: a spontaneous language sample, a single-word naming task, and a non-word repetition task. We investigate whether phonological production contributes unique variance to accounting for vocabulary development and whether its contribution is dependent upon the sampling method. We depart from previous studies which have used hierarchical linear regression to examine the influence of lexical and phonological factors on vocabulary size. Instead, we use mixed-effect logistic regression to examine the influence of these factors on whether a word is produced or not.

The second aim is to examine whether parents are able to reliably judge the pronunciation of their children on an experimental version of the MCDI, which we refer to as MCDI-pronunciation. We use this inventory in two ways. First, we compare the results obtained with this more fine-grained measure versus those obtained with the standard form of the MCDI (which uses a binary outcome; see the first aim) so as to determine whether it is influenced by the same or by a different set of factors. Thus, we examine what factors influence whether a word is pronounced poorly, adequately, well, or not produced at all. We hypothesize that phonological production plays a stronger role in accounting for variance on the MCDI-pronunciation than on the standard version since the latter reports on production accuracy. Second, we use the inventory as a measure of phonological production by calculating the proportions of words indicated by the parents to be

well-pronounced. We compare this percentage score with the other measures of phonological production to ascertain whether it is also correlated with vocabulary development. The third aim is to investigate whether the same set of lexical and phonological factors that influence vocabulary development in monolingual children also influence bilingual children.

Method

Participants

40 children (21 monolinguals; 22 girls), aged 1;11 to 3;1 months (Mean age = 2;5 months), took part in the study. They were recruited via flyers deposited at preschools and kindergartens in the canton of Geneva, Switzerland. Fifty children were originally tested. Eight children (4 monolinguals and 4 bilinguals) were excluded because they did not produce enough speech during the testing session and two were excluded due to recording difficulties. Parents reported that their children did not have any history of otitis media or hearing difficulties and that their children were developing normally. The parents also signed an informed consent form as required by the ethics committee at the University of Geneva.

Bilingual status was determined by a parental questionnaire, in which the parents indicated whether their child spoke another language at least 30% of the time. Parents also indicated the language the child spoke at home and with whom, and at what age the child had acquired French. In addition, they judged the dominance of French versus the other language on a scale from 1 to 5. Of the 19 bilinguals, 15 were dominant in French, three were balanced, and one child was dominant in the other language. The bilinguals spoke the following languages: Spanish ($n = 6$), English (5), Italian (4), Portuguese (1), German (1), Catalan, (1), and Arabic (1). [Appendix A](#) provides information on the children's age, gender, bilingual status, languages, and dominance.

Procedure

Children took part in a single session of 30 to 60 minutes in which they were engaged in three tasks: a spontaneous language sample, a single word production task, and a NWR task. The sessions took place in a quiet room in the children's homes. At the same time, the parents completed several questionnaires:

1. A general questionnaire to provide information on bilingual status and speech and language development;
2. The MCDI in French;
3. The MCDI in the L1 (i.e., language spoken at home that was not French) if the child was bilingual;
4. The MCDI-pronunciation to provide information on the child's pronunciation in French.

Parental questionnaires

General parental questionnaire

The general parental questionnaire was developed for the purposes of the study and was loosely based on the PABIQ (Tuller, 2015). Parents provided information on the bilingual

status of the child if he/she was bilingual, whether the child attended preschool and how often, and his/her speech and language development. They also indicated the number of years of schooling they had received and their professions to provide information on socioeconomic status (SES).

MCDI in French

Parents completed the Mots et Phrases form of L'Inventaire Français du Développement Communicatif (IFDC) (Kern & Gayraud, 2010) (the French adaptation of the MCDI). This form is designed for children from 1;4 to 2;6. It consists of a list of 690 words organized into 22 semantic categories. Parents were asked to indicate whether their child produced the word. The IFDC is sensitive to vocabulary development over time and has strong short-term test-retest reliability ($r = .90$; Kern & Gayraud, 2010). Although the MCDI is typically used with children through to 2;6, we employed it with children through to 3;1 to avoid having to employ different vocabulary measures with younger versus older children. We didn't observe any ceiling effects on the IFDC with the older children.

MCDI in the L1

Parents of the bilingual children completed the MCDI in the L1. Spanish, British English, Italian, Portuguese, Catalan, German, and Arabic versions of the MCDI were used. Scores on this questionnaire allowed us to calculate the total vocabulary, which consisted of the scores on the French MCDI and the MCDI in the L1.

MCDI-pronunciation

The MCDI-pronunciation was based on a reduced set of words ($n = 279$) in the IFDC (see below). Parents were presented with a list of words organized in a similar fashion to the IFDC, and three columns corresponding to three levels of pronunciation.³ Parents were asked to indicate whether their child pronounced the word poorly ("produit mal"), adequately ("produit assez bien"), or well ("produit bien"). If the child did not produce the word at all, the parents were requested to leave the response blank. Examples were provided to the parents on the first page of the questionnaire to assist them in completing the questionnaire. Poorly pronounced words were characterized by syllable omission in the case of two-syllable words (e.g., cadeau → deau 'present') and segment omissions in the case of one-syllable words (e.g., bouche → bou 'mouth'; livre → li 'book'). Adequately pronounced words were characterized by substitutions (e.g., cadeau → tadeau 'present'; bouche → bousse 'mouth') and, in the case of words with clusters, also segment omissions (e.g., livre → liv 'book'). Well-pronounced words were characterized by target-like production (e.g., cadeau → cadeau; bouche → bouche; livre → livre). We asked parents to provide examples of their children's pronunciations when possible.

³Although it would have been preferable to incorporate the pronunciation component into the standard version of the MCDI, we did not do so for several reasons. First, the utilization of the standard version of the MCDI (or IFDC) allowed our results to be comparable to those of other studies that have used the same version. Second, there was insufficient space on the standard version to permit the inclusion of additional columns for rating pronunciation.

Phonological production tasks

Spontaneous language sample

Children participated in a free play session of approximately 20 minutes in which they interacted with the experimenter and, on occasion, their parents. The play items were the same for each child (playmobile items of animals, vehicles, & little people), thus ensuring a uniform set of vocabulary items per child. The children produced on average 68.75 utterances (range: 40-130).

Single-word naming task

Children took part in a single-word naming task of approximately 15 minutes. The task consisted of 35 words that targeted most of the consonants of French in word-initial position. Each phoneme was represented in both a one- and two-syllable word.⁴ We did not target consonants in all word positions to avoid having a test that was overly long for young children. The words were selected from the IFDC and were, thus, familiar to children of two to three years. The children were required to name the pictures following the question “Qu’est-ce que c’est?” ‘What is that?’ or “Comment ça s’appelle?” ‘What is that called?’ When the children were unable to name the word, they were given semantic or phonological prompts; or, if they still could not say the word, the word was elicited via delayed imitation (“C’est un chapeau. Est-ce que tu peux dire ça?” ‘That’s a hat. Can you say that?’). On average, 24 of the 35 words were spontaneously produced and the remaining words were produced by elicitation. [Appendix B](#) presents a list of the words in the single-word naming task.

Non-word repetition task

The NWR test was based on one previously developed by Parra, Hoff, and Core (2011) for young English- and Spanish-speaking children. In the case of the French adaptation, non-words were constructed from real words taken from the IFDC following two basic principles: (1) monosyllabic non-words were created by changing the first phoneme (e.g., dame /dam/ ‘lady’ → /bam/); (2) multi-syllabic nonwords were created by combining syllables of words from the IFDC which occurred in the same word position (e.g., maman /mamã/ ‘mummy’ and ballon /balõ/ ‘balloon/ (foot) ball’ → /malõ/). The test included two training trials with monosyllabic non-words and 12 test trials with one-, two-, and three-syllable non-words (four trials per group). To administer the task, the examiner showed a toy, said the non-word as if it was the toy’s name, and asked the child to repeat it back (e.g., “This guy is named Bam. Can you say Bam?”). If the child didn’t repeat the name, the examiner repeated the non-word up to three times. Only the first repetition produced by the child was scored, regardless of its accuracy. If a child failed to repeat the non-word for six consecutive trials, the test was ended. The accuracy of NWR was measured by calculating the total number of consonants presented that were repeated correctly by the child. All children succeeded in completing the task, although eight children did not produce all of the 12 items. [Appendix C](#) provides a list of the stimuli in the NWR task.

⁴The single-word production task did not target /p/ because it does not occur in word-initial position in French, nor /q/ because of its low frequency in words to young children. We only included one word with word-initial /z/ (e.g., zèbre ‘zebra’) because of its low frequency in French.

Phonological measure from MCDI-pronunciation

We also derived a measure of phonological production from the MCDI-pronunciation. We calculated the percentage of words that were indicated as poorly, adequately, and well-pronounced from the total number of words indicated on the questionnaire for each child. We used the category of percent well-pronounced (% well-pronounced) as the principal measure rather than percent poorly- or adequately-pronounced, since it was the category most frequently reported by the parents.

Data transcription

Children's speech was recorded with a portable digital tape recorder (MARANTZ, Professional PMD620). Using Phon, a software program designed for the analysis of phonological data (Rose & MacWhinney, 2014), each child's WAV format file was segmented, and stimulus words (or phrases) were identified and transcribed. Two French-speaking graduate students who had experience in phonetic transcription performed the analyses. They transcribed the data using broad phonemic transcription. The transcribed data were transferred to Excel. Calculations of percent consonants correct (PCC) were computed automatically for each child in Phon.

Inter-tester reliability

Four children (two monolingual and two bilingual) were transcribed independently by two transcribers using the blind transcription function in Phon. Point-to-point agreement for consonant transcription was 86% (range: 81.10 – 91.83%) for the spontaneous language sample; 85% (81.50 – 89.70%) for the single-word naming task, and 88% (84.72 – 91.7%) for the NWR task, indicating good reliability.

Data-coding of the target-word variables

The coding of target words was based on a reduced set of words of the IFDC, the same set of words used in the MCDI-pronunciation. It included one- and two-syllable nouns taken from 12 categories of the IFDC considered representative of core vocabulary (Stokes, 2010; Stokes et al., 2012b). It excluded onomatopoeia, games and routines, and context-based items (e.g., people and function words). We focused on one- and two-syllable nouns since previous studies indicated that the psycholinguistic properties of this reduced set of words were highly predictive of total vocabulary size (Kehoe et al., 2018, 2020). Furthermore, Kern and Dos Santos (2017) did not find that ND and WF based on predicates influenced lexicon size at least at 2;0 to 2;6. The 279 words of the reduced version of the IFDC were coded for phonetic complexity, ND, and WF.

Phonetic complexity

We coded each word in terms of word length in phonemes and the IPC. Concerning the latter, a point was assigned to each word if it contained a dorsal consonant (e.g., camion [kamjɔ̃] 'truck'), a fricative or liquid (e.g., avion [avjɔ̃] 'plane'; balle [bal] 'ball'), a final consonant (e.g., balle [bal] 'ball'), three or more syllables (e.g., animal [animal] 'animal'), two or more consonants with different places of articulation (e.g., balle [bal] 'ball', which

has labial and coronal places of articulation), a tautosyllabic cluster, a cluster that occurs within a syllable (e.g., crayon [kʁejɔ̃] ‘pencil’), or a heterosyllabic cluster, a cluster that is split across two syllables (e.g., tracteur [tʁaktœːʁ] ‘tractor’). The IPC also assigns points to rhotic vowels, but since rhotic vowels do not occur in French, this category was excluded.

Neighbourhood density

All items in the reduced set of the IFDC were coded for ND using the values generated by the Lexique3 database, a corpus of adult language (New, Brysbaert, Veronis, & Pallier, 2007). The most frequent phonological forms were chosen when two noun choices were provided (e.g., figure/visage ‘face’).

Word frequency

All items in the reduced set of the IFDC were coded for token frequency using the Lexique3 database (New et al., 2007).⁵ The frequency data are based on film subtitles from French films and French translations of English films and television series. In the case of words that can have multiple inflections, the frequency value corresponded to the sum of all possible inflections that had the same pronunciation as the target form (the most frequent phonological form). The WF values were log-transformed due to the skewed nature of the raw frequency values (Brysbaert, Madera, & Keuleers, 2018).

We acknowledge that it would have been preferable to use ND and WF values based on child corpora. We follow the arguments of Stokes and colleagues (Stokes, 2014; Stokes et al., 2012b) in noting that (a) children are exposed not only to child- but also to adult-directed speech, (b) ND and WF information from child corpora is not readily available in French, and (c) authors have observed high correlations between WF values generated in child- and adult-directed corpora (Gierut & Dale, 2007; Stokes, 2010).

Statistical analyses

Data were analyzed using mixed effect 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, 2020) and the lme4 package (Bates, Maechler, Bolker, & Walker, 2015) for mixed effects models.

The analyses included the control variables: age (in months), gender, bilingual status (monolingual, bilingual), and maternal education (in years). The target word predictor variables were ND, WF, word length (in phonemes), and IPC calculated separately for each word. Following Jones and Brandt (2019b), we also examined whether the influence of ND, WF, and phonetic complexity was dependent on the child’s overall vocabulary level, so we included the variable, vocabulary size, based on the standard version of the French-MCDI.

The phonological production variables were PCC for the three sampling conditions: spontaneous language sample (PCC-Spon), single-word naming task (PCC-Name), and NWR task (PCC-NWR), as well as the variable, %well-pronounced, extracted from the

⁵We were not able to obtain frequency values for 5 words of the reduced version of the IFDC in Lexique 3: kiwi, sandwich, sweat, cornflakes, and WC.

MCDI-pronunciation. We also measured phonemic inventory size for the three sampling measures. We found that correlations between inventory and vocabulary size were much lower than those between PCC and vocabulary size. Furthermore, inventory size did not emerge as significant in any of the statistical models. For this reason and to simplify the data presentation, we focus the analyses on PCC. We also did not include the variable % well-pronounced in our statistical models since it was a measure derived from the dependent variable and thus could have led to biased results. We included it, however, in our descriptive analyses. To assess multicollinearity, we ran simple binomial regression models in which the dependent variable was predicted by combinations of target word or phonological production variables. We then calculated the variance inflation factor (VIF) using the Car package in R. The VIFs for target word variables were low and did not exceed 1.6, whereas the VIFs for phonological production variables were high, in the vicinity of 6-7, thus, justifying our decision to add these variables separately into the models.

We fitted three different statistical models. In the first model, we examined the factors that influence whether a word was indicated as produced or not on the MCDI. This model is referred to as Model 1: Production and is based on the reduced set of words from the standard French MCDI ($n=279$). The dependent variable was produced (i.e., 1) or not produced (i.e., 0). This model addressed the first aim of the study. The second and third models were based on the MCDI-pronunciation and addressed the second aim of the study which was to examine whether parents could reliably judge the pronunciation of their children. One limitation of this experimental questionnaire was that parents were not always consistent in what they indicated on the standard versus experimental version of the MCDI. Sometimes, they indicated that children did not produce the word on the MCDI but indicated that the children pronounced it (poorly, adequately, or well) on the MCDI-pronunciation. At other times, they indicated that the children produced it on the MCDI but did not pronounce it (poorly, adequately, or well) on the MCDI-pronunciation. To offset these limitations, we coded the data in two ways. First, we took into consideration those words indicated as not produced on the MCDI and coded them as “0” on the MCDI-pronunciation. Those words indicated as produced on the MCDI and filled in on the MCDI pronunciation were coded as on the MCDI-pronunciation; those words indicated as produced on the MCDI but not filled in on the MCDI-pronunciation were not coded but left as missing data. Thus, the second model examines what factors influence whether a word is pronounced poorly, adequately, well, or not produced at all. This model is referred to as Model 2: Production and pronunciation. The dependent variable was a proportion score: actual production or pronunciation score/total score whereby “0” indicated not produced, “1” indicated “poorly-”, “2” indicated “adequately-”, and “3” indicated “well-pronounced” and the total score was “3”. Second, we considered only the items filled out on the MCDI-pronunciation regardless of whether the word was indicated as produced or not on the standard version of the MCDI. This model is referred to as Model 3: Pronunciation. All items that were not filled were left blank and counted as missing data. Thus, the third model focuses on what factors influence whether a word is pronounced poorly, adequately, or well. In this model, the dependent variable was once again a proportion score: actual pronunciation score/total pronunciation score, whereby “1 to 3” indicates the different levels of pronunciation and “3” is the total pronunciation score.

To establish the most parsimonious model, we added variables in several steps, removing variables that were not significant after each step. We first added the control variables. We then added the target word variables, WF and ND, and both measures of

phonetic complexity. Following this, we examined whether there were any significant interactions between the target word variables – that is, interactions between WF and ND, or between WF and IPC, etc. We also examined whether the influence of ND, WF, and phonetic complexity was dependent on the child's overall vocabulary level. Thus, we included the interaction of total vocabulary size based on the standard version of the French MCDI and ND, WF, or phonetic complexity variables. Finally, we added the production variables (PCC) for the different sampling conditions one at a time, as well as their interaction with target word variables, and included in the model the variable/variables that were the most significant. The final most optimal model was the one that had the lowest Akaike Information Criterion (AIC) and highest log-likelihood ratio. The random part of the model included random intercepts for participants and items (i.e., words). The model was fitted using maximum likelihood estimation.⁶

Results

Descriptive statistics

First, we present descriptive statistics on the tests conducted in the study. [Table 1](#) indicates the means, standard deviations (SD), and ranges of scores for the vocabulary and phonological measures for the entire sample of children, and [Table 2](#) indicates the means and standard deviations of scores according to monolingual and bilingual groups. Children knew on average 410 French words and the bilingual children knew on average another 250 words in their L1 (see [Table 2](#)). Their PCCs ranged from 70 to 80%, with slightly higher scores in the spontaneous language sample than in the single-word naming and NWR tasks. Parents judged approximately 70 to 80% of children's words as well-pronounced. There was a slight tendency for bilinguals to have higher scores than monolinguals on the lexical and phonological measures; however, a series of independent

Table 1. Descriptive statistics for the lexical and phonological measures

Measure	Mean	Sd.	Range
MCDI-Fr ^a	409.98	152.77	41–603
PCC-Spon	79.76	12.56	50.77–97.48
PCC-Name	71.13	16.24	25–92.31
PCC-NWR	73.70	18.66	33.33–96.87
%Well-Pronounced	72.60	21.71	18.63–98.71

^aMCDI-Fr = MacArthur Communicative Developmental Inventory in French (i.e., IFDC); PCC-Spon = Percent consonants correct based on spontaneous language sample; PCC-Name = Percent consonants correct based on single-word naming task; PCC-NWR = Percent consonants correct based on nonword repetition task; % Well-Pronounced = Percent of words indicated as well-pronounced by parents on the MCDI-Pronunciation.

⁶Based on procedures by Stokes et al. (2019) we also examined the by-participant random slopes for WF, ND, length, and ICP to determine whether there were individual differences in the effect of these variables on the probability that a word would be produced. Like Stokes et al. (2019), we found that the by-participant random slope for WF significantly contributed to the model but the by-participant random slope for the other variables (ND, length, IPC) did not. However, we found that adding the interaction between WF and vocabulary size resulted in a similar model fit and allowed us to avoid model convergence difficulties. Thus, we present the findings of models with a simpler set of random factors.

Table 2. Descriptive statistics (means and standard deviations) for the lexical and phonological measures separated according to monolingual and bilingual children

Measure	Monolinguals (n = 21)		Bilinguals (n = 19)	
	Mean	Sd.	Mean	Sd.
MCDI-Fr ^a	404.19	147.69	412.26	162.17
MCDI-Total	-		650.63	261.13
PCC-Spon	79.44	12.72	80.11	12.72
PCC-Name	67.46	17.96	75.19	13.41
PCC-NWR	70.48	21.24	77.26	15.09
%Well-Pronounced	69.95	23.62	78.85	17.97

^aMCDI-Fr = MacArthur Communicative Developmental Inventory in French (i.e., IFDC); MCDI-Total = Total vocabulary score based on MCDI-FR and MCDI of the L1 (i.e., language spoken at home that is not French); PCC-Spon = Percent consonants correct based on spontaneous language sample; PCC-Name = Percent consonants correct based on single-word naming task; PCC-NWR = Percent consonants correct based on nonword repetition task; % Well-Pronounced = Percent of words indicated as well-pronounced by parents on the MCDI-Pronunciation.

t-tests indicated no significant differences between monolinguals and bilinguals on any of the measures except for total vocabulary, whereby bilinguals knew more words than monolinguals (MCDI-Fr: $t(38)=.17$, $p=.87$; MCDI-total: $t(38)=3.7$, $p<.001$; PCC-Spon: $t(38)=.16$, $p=.87$; PCC-Name: $t(38)=1.53$, $p=.07$; PCC-NWR: $t(38)=1.15$, $p=.26$; %well-pronounced: $t(38)=1.78$, $p=.08$).

Table 3 presents the correlation coefficients between age (in months), lexical, and phonological measures. There were no significant correlations between age and vocabulary size but there were between age and all phonological production measures. There were significant correlations between the French MCDI and all phonological production measures. In addition, there were significant correlations between total vocabulary and all phonological production measures with the exception of PCC-Spon. In general, correlations

Table 3. Correlation matrix between lexical and phonological variables

Variables	Age	MCDI-Fr ^a	MCDI-Total	PCC-Spon	PCC-Name	PCC-NWR	%Well Pron
Age	-	.21	.19	.39 ^{ab}	.47**	.44**	.39*
MCDI-Fr	-	-	.74***	.51**	.62***	.62***	.63***
MCDI-Total	-	-	-	.30	.53***	.55***	.62***
PCC-Spon	-	-	-	-	.81***	.82***	.67***
PCC-Name	-	-	-	-	-	.91***	.79***
PCC-NWR	-	-	-	-	-	-	.78***
%Well-Pron	-	-	-	-	-	-	-

^aMCDI-Fr = MacArthur Communicative Developmental Inventory in French (i.e., IFDC); MCDI-Total = Total vocabulary score based on MCDI-FR and MCDI of the L1 (i.e., language spoken at home that is not French); PCC-Spon = Percent consonants correct based on spontaneous language sample; PCC-Name = Percent consonants correct based on single-word naming task; PCC-NWR = Percent consonants correct based on nonword repetition task; % Well-Pronounced = Percent of words indicated as well-pronounced by parents on the MCDI-Pronunciation.

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

Table 4. Correlation coefficients between psycholinguistic variables

Variables	ND ^a	WF	length	IPC
ND	–	.26*** ^b	–.77***	–.45***
WF	–	–	–.31***	–.21**
length	–	–	–	.62***
IPC	–	–	–	–

^aND = Neighbourhood density; WF = word frequency; Length = length of a word in phonemes; IPC = phonetic complexity based on the Index of Phonetic Complexity (IPC).

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

were higher between the French MCDI and phonological production than between total MCDI and phonological production, with the exception of the parent-based measure in which the correlations were similar (.62/.63). Correlations among the phonological production scores were all very high (correlations ranging from .67-.91). Important for our purposes and relevant to the second aim of the study was the finding that the parent-based phonology score (% well-pronounced) correlated highly with both the French and total MCDI (.62/.63) and correlated highly with all the other phonological production measures (.67-.79).

Table 4 presents the correlation coefficients between the psycholinguistic variables ND, WF, length (in phonemes), and IPC. These correlations were based on the values for each individual word of the reduced version of the IFDC. All correlations among the variables were significant. Moderately high correlations were obtained between ND and word length, and IPC and word length.

Statistical models

We now present the findings of the statistical models. Table 5 presents the best fitting model for Model 1: Production – that is, the model that best accounts for the factors that influence whether a word was indicated as produced or not on the MCDI. This model addresses the first aim of the study. This model was based on 10960 coded words from all 40 children. In terms of control variables, only gender emerged as significant: girls were more likely to produce words than boys ($\chi^2(1)=4.14$, $p=.04$). In terms of target word variables, WF was the only variable that had a significant main effect. Words of high frequency were produced more often than words of low frequency ($\chi^2(1)=71.45$, $p<.001$). Frequency interacted with two other variables: IPC ($\chi^2(1)=8.65$, $p=.003$) and vocabulary size ($\chi^2(1)=83.62$, $p<.001$). First, the effect of frequency was modulated by phonetic complexity. As seen in Figure 1, there was a general tendency for words of higher phonetic complexity to be lower in word frequency; however, the difference between the word frequency of produced and non-produced words widened with increasing phonetic complexity suggesting some trade-off effects between phonetic complexity and word frequency: children could produce more complex phonetic forms when they were of higher frequency. Second, the effect of frequency was modulated by vocabulary size. As seen in Figure 2, as vocabulary size increased, the frequency of produced words decreased, and the difference between the frequency of produced and non-produced words also increased, suggesting stronger effects of word frequency with increasing vocabulary size. Similarly, phonetic complexity was modulated by vocabulary size ($\chi^2(1)=24.20$, $p<.001$). As vocabulary size increased, the phonetic complexity of produced words increased. Finally, in terms of phonological

Table 5. Model 1: Production. Best fitting model to explain the factors that influence whether a word is produced or not on the French MCDI

Fixed Effects	Estimate	Std. Error	z value	p value
(Intercept)	-0.41573	0.32921	-1.263	0.21
gender	-0.53200	0.26045	-2.043	0.04
WF ^a	2.11631	0.23947	8.837	< .001
WF:IPC	-0.12298	0.04145	-2.967	0.003
WF:MCDI	0.62400	0.06449	9.675	< .001
IPC:MCDI	0.09765	0.01910	5.114	<.001
PCC-NWR	0.32963	0.14054	2.345	0.02
Random Effects				
Group Name	Variance		Std.Dev.	
Word (Intercept)	3.5794		1.8919	
Participant (Intercept)	0.5709		0.7556	

^aWF = word frequency; WF:IPC = interaction between word frequency and phonetic complexity based on the Index of Phonetic Complexity (IPC); WF:MCDI = interaction between word frequency and vocabulary size based on the French MacArthur Communicative Developmental Inventory (MCDI); IPC:MCDI = interaction between phonetic complexity based on the IPC and vocabulary size; PCC-NWR: Percent consonants correct (PCC) based on the nonword repetition (NWR) test.

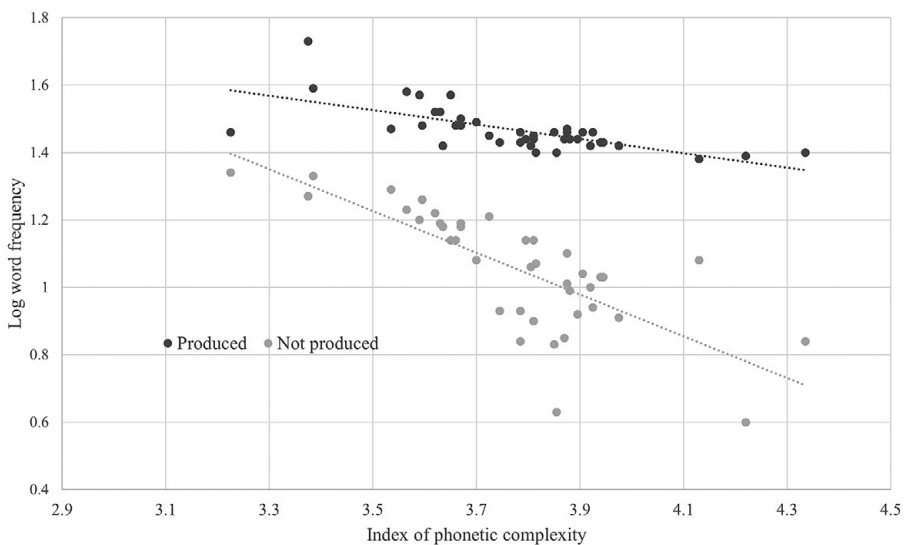


Figure 1. This graph shows the relation between phonetic complexity (IPC) and word frequency (WF). It plots the mean WF for produced and non-produced words across the mean IPC for the 40 children.

production variables, we found that PCC-NWR contributed to model fit ($\chi^2(1)=5.74$, $p<.05$). Children with superior phonological memory capacities produced more words. Consonant precision based on the spontaneous language sample or the single-word naming task did not emerge as significant in the model.

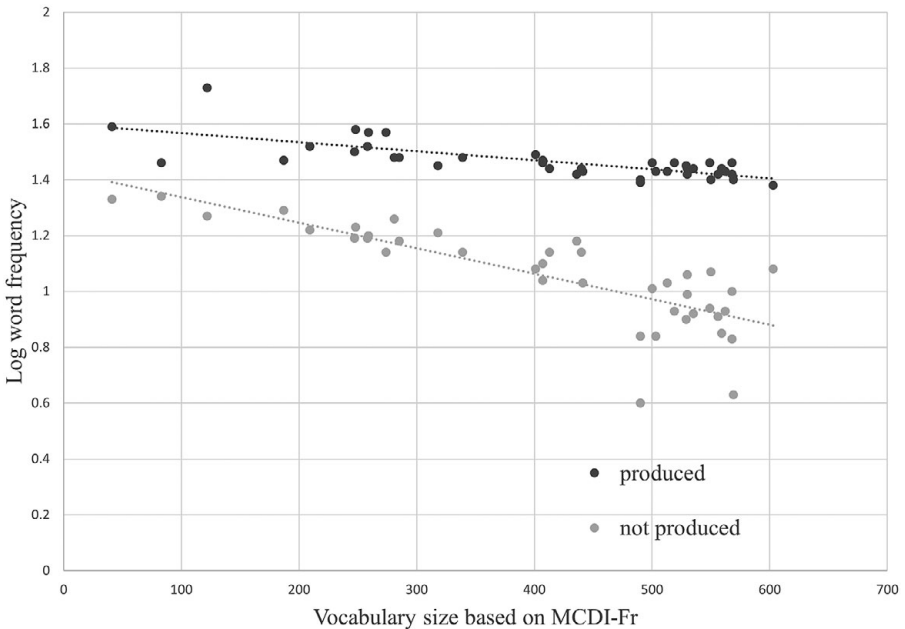


Figure 2. This graph shows the relation between word frequency (WF) and vocabulary size. It plots the mean WF for produced and non-produced words across vocabulary size for the 40 children based on the French MCDI.

The remaining two models address the second aim of the study concerning whether parents could provide information on their children's word pronunciation. To remind the reader, we were required to set up two models to offset the fact that parents were inconsistent when completing the vocabulary questionnaires. Table 6 presents the best fitting model for Model 2: Production and pronunciation – that is, the model that best accounts for the factors that influence whether a word was indicated as poorly, adequately, well-pronounced (on the MCDI-pronunciation), or not produced at all (as determined by the standard version of the MCDI). This model was based on 10468 coded words from all 40 children. There were strong similarities between Models 1 and 2 in terms of the factors that emerged as significant. As in Model 1, WF emerged as a main effect ($\chi^2(1)=68.26, p<.001$). As shown in Figure 3, words that were pronounced well had higher mean frequency than words that were pronounced adequately which in turn had higher mean frequency than words that were not produced at all. WF interacted with phonetic complexity ($\chi^2(1)=33.27, p<.001$), and vocabulary size ($\chi^2(1)=22.23, p<.001$) as was also observed in Model 1. In addition, there were differences between Models 1 and 2. First, the control variable, gender, did not explain any additional variance. Second, ND and the phonological production variables played an important part in the model. ND did not emerge as a significant main effect but it interacted with vocabulary size ($\chi^2(1)=14.90, p<.001$). As vocabulary size increased, children produced and pronounced well words of lower ND. In terms of phonological production, all three variables (PCC-Spon, PCC-Name, PCC-NWR) when entered on their own significantly improved model fit to data. The phonological production variable that resulted in the best AIC score was PCC-Name. It emerged as a significant main effect ($\chi^2(1)=9.79, p=.002$) and also significantly interacted with two target-word variables: WF ($\chi^2(1)=40.00, p<.001$) and ND ($\chi^2(1)=26.21, p<.001$). Children with superior PCCs based on the single-word naming task produced and pronounced well more words than children with inferior PCCs. Furthermore, children with

Table 6. Model 2: Production and Pronunciation. Best fitting model to explain the factors that influence whether a word is pronounced poorly, adequately, well, or not at all produced on the French MCDI

Fixed Effects	Estimate	Std. Error	z value	p value
(Intercept)	-1.10721	0.29550	-3.747	< .001
WF ^a	1.35905	0.15521	8.756	< .001
WF:IPC	-0.40656	0.06865	-5.922	< .001
WF:MCDI	0.20508	0.04301	4.768	< .001
ND:MCDI	0.10204	0.02641	3.864	< .001
PCC-Name	0.62983	0.18989	3.317	< .001
PCC-Name:WF	0.23837	0.03747	6.362	< .001
PCC-Name:ND	-0.11701	0.02275	-5.143	< .001
Random Effects				
Group Name	Variance		Std.Dev.	
Words (Intercept)	2.491		1.578	
Participant (Intercept)	1.359		1.166	

^aWF = word frequency; WF:IPC = interaction between word frequency and phonetic complexity based on the Index of Phonetic Complexity (IPC); WF:MCDI = interaction between word frequency and vocabulary size based on the French MacArthur Communicative Developmental Inventory (MCDI); ND:MCDI = interaction between neighbourhood density (ND) and vocabulary size; PCC-Name = Percent consonants correct (PCC) based on the single-word naming task; PCC-Name:WF = interaction between percent consonants correct (PCC) based on the single word naming task and word frequency; PCC-Name:ND = interaction between percent consonants correct (PCC) based on the single word naming task and neighbourhood density.

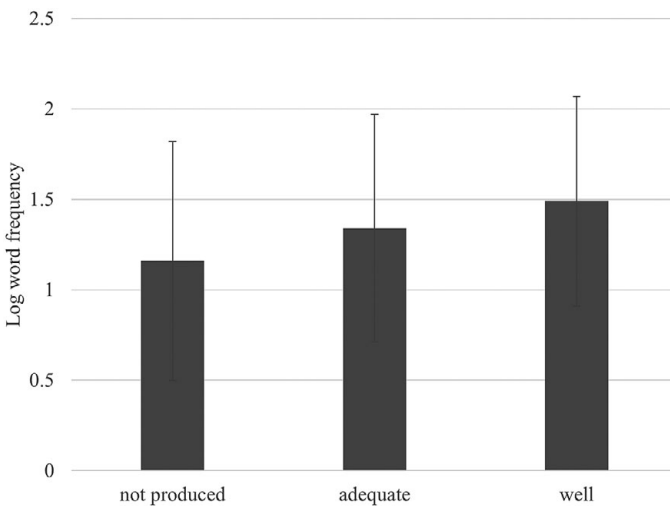


Figure 3. This graph shows the mean log word frequency (and standard deviation) for words that were not produced, pronounced adequately and well for the 40 children. Results for poorly pronounced words are not shown because they were not frequently reported by parents.

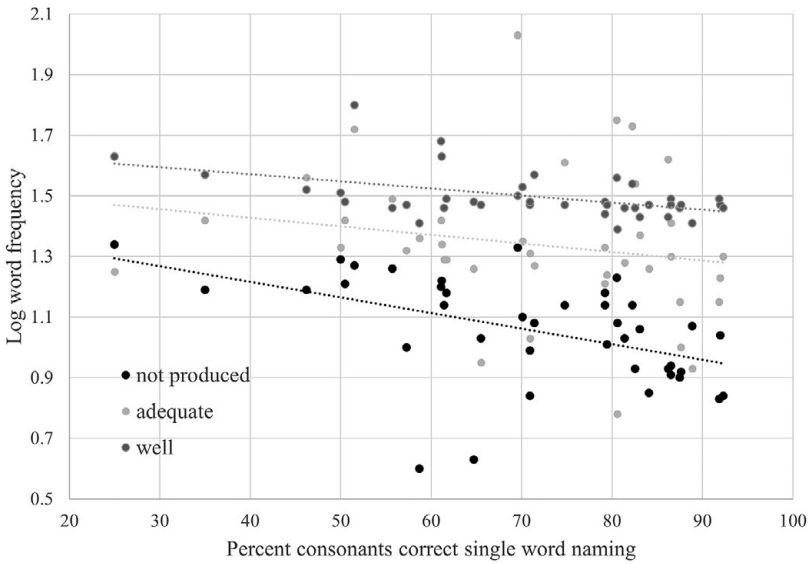


Figure 4. This graph shows the relation between log word frequency (WF) and percent consonants correct on the single-word naming task (PCC-Name). The mean log WF for non-produced words as well as for words that were pronounced adequately and well are plotted across the mean PCC-Name for the 40 children.

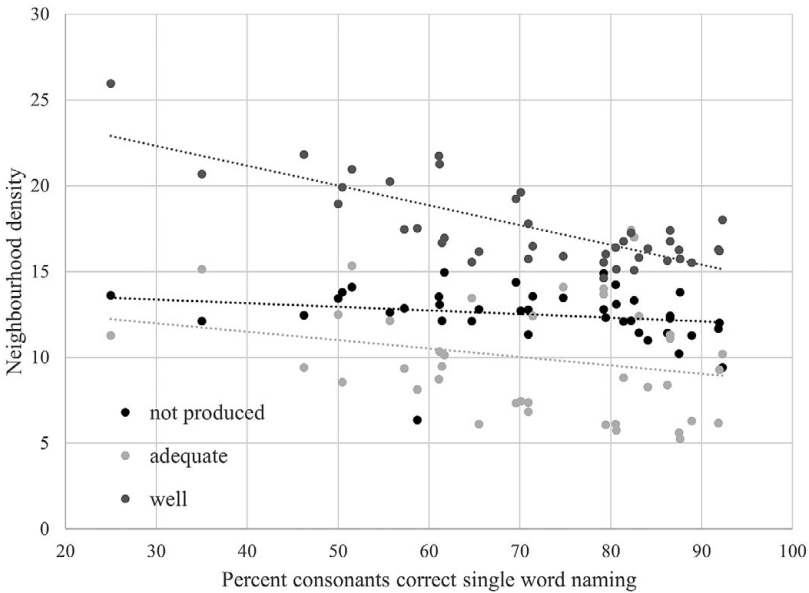


Figure 5. This graph shows the relation between neighbourhood density (ND) and percent consonants correct on the single-word naming task (PCC-Name). The mean ND for non-produced words as well as for words that were pronounced adequately and well are plotted across mean PCC-Name for the 40 children.

superior PCCs based on the single-word naming task produced and pronounced well words with lower frequency and lower NDs than children with inferior PCCs. These patterns are shown in Figures 4 to 5. We do not display findings for the poorly produced words because

they were not frequently indicated by the parents and were not present across all children. Also, note that non-produced words had lower frequency than adequately and well-pronounced words; however, the ND of non-produced words fell between those of adequate and well-pronounced words. This may suggest that WF determines whether words are produced or not but ND plays a greater role in determining whether words are pronounced well or not.

Table 7 presents the best fitting model for Model 3: Pronunciation – that is, the factors that influence whether a word was indicated as poorly, adequately, or well-pronounced on the MCDI-pronunciation. There were 8071 coded items across the 40 children. No control variables emerged as significant in the model. Several target word variables emerged as significant: WF ($\chi^2(1)=10.12$, $p=.002$), ND ($\chi^2(1)=60.74$, $p<.001$), and IPC ($\chi^2(1)=106.76$, $p<.001$). As seen in Figure 6, well-pronounced words were characterized by higher WF, higher ND, and lower phonetic complexity than adequately pronounced words. We do not show results for poorly produced words as they were not frequently reported by parents. In addition, and as observed in the other models, WF interacted with vocabulary size. As vocabulary size increased, children pronounced well words of lower frequency. In terms of production variables, all variables contributed significantly to model fit when entered on their own. Of the three variables, the variable PCC-Name resulted in the lowest AIC score ($\chi^2(2)=25.61$, $p<.001$). Thus, children with the highest phonological ability based on the single-word naming task had the best pronunciation scores. PCC-Name also interacted with the IPC ($\chi^2(2)=14.70$, $p<.001$). As phonological abilities increased, children pronounced well words of higher phonetic complexity (see Figure 7).

Table 7. Model 3: Pronunciation. Best fitting model to explain the factors that influence whether a word is pronounced poorly, adequately, or well on the MCDI-pronunciation

Fixed Effects	Estimate	Std. Error	z value	p value
(Intercept)	2.50158	0.16241	15.403	< .001
WF ^a	0.20830	0.06445	3.232	0.001
ND	0.38623	0.04755	8.123	< .001
IPC	-0.52272	0.04646	-11.251	< .001
WF:MCDI	0.19606	0.03420	5.733	< .001
PCC-Name	0.79926	0.13495	5.923	< .001
PCC-Name:IPC	-0.10160	0.02627	-3.867	< .001
Random Effects				
Group Name	Variance		Std.Dev.	
Word (Intercept)	0.2341		0.4838	
Participant (Intercept)	0.6576		0.8109	

^aWF = word frequency; ND = neighbourhood density; IPC = phonetic complexity based on the Index of Phonetic Complexity (IPC); WF:MCDI = interaction between word frequency and vocabulary size based on the French MacArthur Communicative Developmental Inventory (MCDI); PCC-Name: Percent consonants correct (PCC) based on the single word naming test; PCC-Name:IPC = interaction between percent consonants correct (PCC) based on the single word naming test and phonetic complexity based on the IPC.

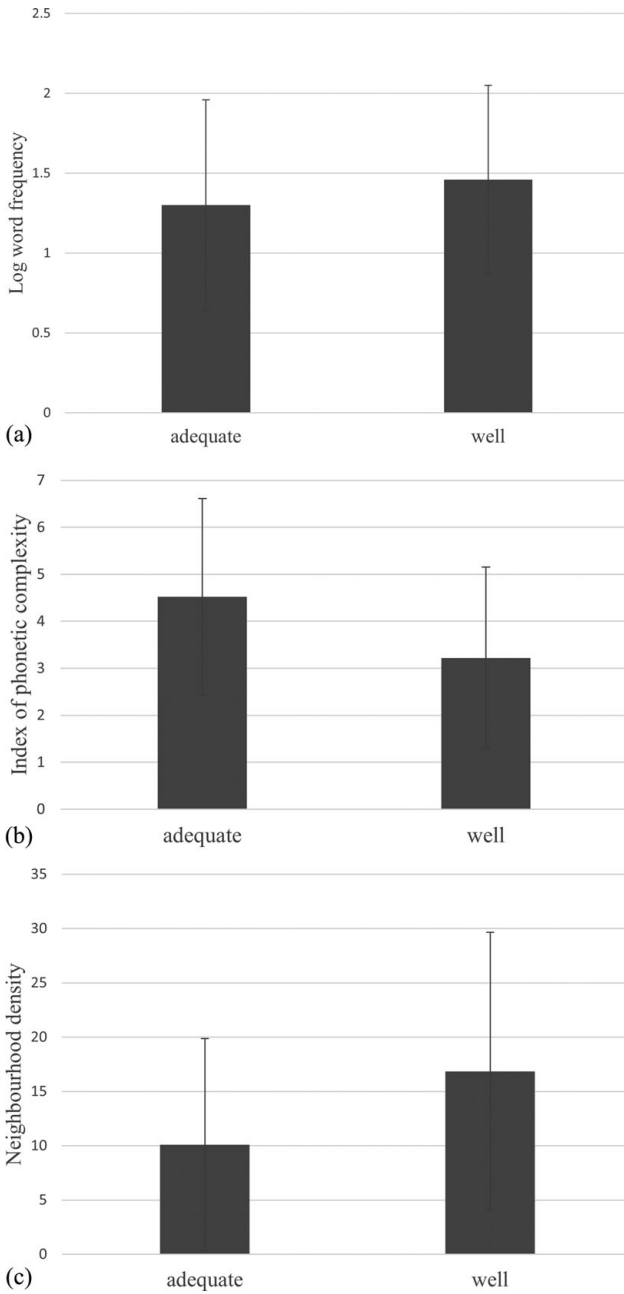


Figure 6. Means and standard deviation of word frequency (6a), phonetic complexity (6b), and neighbourhood density (6c) for adequately and well-pronounced words across the 40 children.

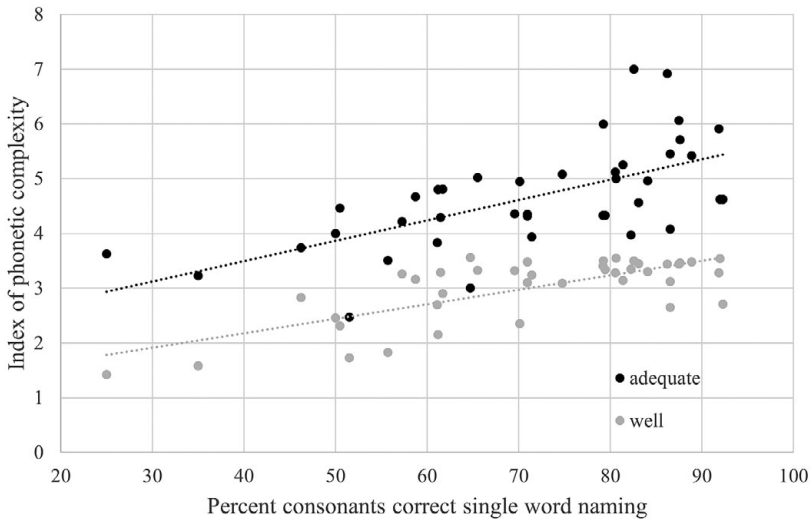


Figure 7. This graph shows the relation between phonetic complexity (ICP) and percent consonants correct on the single-word naming task (PCC-name). The mean IPCs for words that were pronounced adequately and well are plotted across mean PCC-Name for the 40 children.

Table 8 summarizes the findings of the three different models organized according to control variables, target word variables and their interactions, and phonological production variables and their interactions. Of the control variables, gender influenced Model 1. Girls had higher scores in word production than boys. Bilingual status did not emerge as a significant main effect nor as a significant interaction effect in any of the models. Other control variables, age, and maternal education also did not emerge as significant. Of the four target variables, WF played a major role and was a main effect in all statistical models. ND and IPC emerged as main effects only in the models based on pronunciation. Significant interactions included those involving WF. It was modulated either by phonetic complexity or vocabulary size. IPC and ND were also modulated by vocabulary size in Models 1 and 2 respectively. Word length in phonemes did not contribute to any of the models either as a main or interaction effect. In terms of phonological production, we observed that PCC-NWR contributed significantly to Model 1 whereas PCC-Name contributed most to Models 2 and 3. In Models 2 and 3, phonological production also interacted with target word variables.

We used the MuMIn function in R to determine the unique variance contributed by phonological production. The function provides a R^2_{GLMM} based on a revised statistic by Nakagawa, Johnson, and Schielzeth (2017). For each model, we calculated the R^2_{GLMM} for the full model and for the model without phonological production variables. In the first model (Production), PCC-NWR contributed 2% unique variance out of a total of 34.17% variance contributed by all fixed effects (70.89% for fixed and random effects). In the second model (Production and pronunciation), PCC-Name and its interactions contributed 10% unique variance out of a total of 26.49% variance contributed by all fixed effects (66.13% for fixed and random effects). In the third model (Pronunciation), PCC-Name and its interactions contributed 13% unique variance out of a total of 25.70% contributed by all fixed effects (41.54% for fixed and random effects). Thus, we observed an increasing role of phonological production across the three models.

Table 8. Summary of the three different statistical models

Model 1	Model 2	Model 3
Production	Production & Pronunciation	Pronunciation
Control variables		
gender		
Target word variables		
WF ^a	WF	WF
		IPC
		ND
Target word variables + interactions		
WF:IPC	WF:IPC	
WF:MCDI	WF:MCDI	WF:MCDI
IPC:MCDI	ND:MCDI	
Production variables		
PCC-NWR	PCC-Name	PCC-Name
Production variables + interactions		
	PCC-Name:WF	PCC-Name:IPC
	PCC-Name:ND	

^aWF = word frequency; ND = neighbourhood density; IPC = phonetic complexity based on the Index of Phonetic Complexity (IPC); WF:IPC = interaction between word frequency and phonetic complexity; WF:MCDI = interaction between word frequency and vocabulary size based on the French MacArthur Communicative Developmental Inventory (MCDI); IPC: MCDI = interaction between phonetic complexity and vocabulary size; ND:MCDI = interaction between neighbourhood density and vocabulary size; PCC-NWR = Percent consonants correct (PCC) based on the nonword repetition (NWR) test; PCC-Name = Percent consonants correct (PCC) based on the single word naming test; PCC-Name:WF = interaction between percent consonants correct (PCC) based on the single word naming test and word frequency; PCC-Name:ND = interaction between percent consonants correct (PCC) based on the single word naming test and neighbourhood density; PCC-Name:IPC = interaction between percent consonants correct (PCC) based on the single word naming test and phonetic complexity.

Discussion

This study examined lexical and phonological factors that influence vocabulary development in children aged two to three years. First, we were interested in whether phonological production contributes unique variance to vocabulary development over and above psycholinguistic variables such as ND, WF, and phonetic complexity, and whether its contribution was dependent upon the sampling method. Second, we sought to determine whether parents' judgments of children's production accuracy would yield a stronger role of phonology in vocabulary development and whether this measure correlated with other phonological production measures. Third, we investigated whether the same set of lexical and phonological factors influences monolingual and bilingual children alike.

Concerning the first aim, we found that phonological production does indeed contribute unique variance to accounting for vocabulary development although its contribution was modest. Phonological production based on the NWR task appeared to be the most sensitive measure for capturing the role of production in vocabulary development. Concerning the second aim, parents were found to reliably judge the pronunciation of

their children. Statistical models based on judgment scores revealed an important role of phonological production, and phonological measures extracted from the pronunciation questionnaire were highly correlated with vocabulary size and other phonological measures. Finally, concerning the third aim, no effect of bilingualism emerged in our statistical models suggesting similar results for monolingual and bilingual children. In the following paragraphs, we summarize the results in more detail and discuss their relevance to the role of phonological production in lexical development.

Lexical and phonological factors that influence word production

Our findings on what factors influence whether a word is indicated as produced or not on the MCDI are consistent with previous findings on factors that influence vocabulary size (Kehoe et al., 2020). We found that word frequency was a key player. It influenced word production either as a simple effect or in interaction with vocabulary size. Children were more likely to produce high than low-frequency words; however, more specifically, it was children with small vocabulary sizes who were more likely to produce high than low-frequency words. The importance of word frequency in accounting for vocabulary development has been observed in children acquiring English (Jones & Brandt, 2019a, 2019b), Norwegian (Hansen, 2017), French (Kehoe et al., 2020; Kern & Dos Santos, 2017), and Dutch (Stokes et al., 2019; Verhagen, Stiphout, & Blom, 2022). We also found that the phonetic complexity of the target word influenced word production. Still, phonetic complexity did not have a simple effect on word production. Its influence was only when it interacted with word frequency and vocabulary size. Word length (in phonemes) did not emerge as significant in any of the statistical models. Kehoe et al. (2020) found that both word length and the IPC influenced vocabulary size, whereas Jones and Brandt (2019b) found word length to be more significant than ND and WF in determining whether a word on the MCDI was produced or not by children, aged 1;6. The influence of word length appeared to be eclipsed in this study by that of the IPC which suggests that other aspects of phonetic complexity apart from word length (i.e., phonetic features) influence whether a word is produced, at least in children aged two to three years. Similarly, ND did not influence whether a word was produced or not, although a recent study by Stokes et al. (2019) found it to play an important role in whether a known word appeared in children's spoken lexicons. Jones and Brandt (2019a) also found ND to play a role in word production but not word comprehension. In our study, we assume that the variable, IPC, accounted for similar aspects of word production as did ND, and the former appeared to be the more sensitive measure.

As for the role of phonological production in word production, we observed that it does play a role but a small one. It accounted for two percent additional variance, similar to what has been reported by Kehoe et al. (2020) using linear regression to explain vocabulary size. Furthermore, only one phonological measure emerged as significant in statistical models: the measure based on the NWR task. Thus, a measure that reflects phonological memory capacities appears to be the one that best captures phonological influences on word production. Overall, the findings suggest that the main lexical and phonological factors that influence whether a word is produced or not are the properties of the words themselves.

Lexical and Phonological Factors that Influence Word Production and Pronunciation: Parents' Judgments of their Children's Pronunciations

Inspired by comments from Jones and Brandt (2019b) that the binary outcome option (produced or not produced) on the MCDI does not inform us on pronunciation goodness, we developed an experimental version of the MCDI which offered parents more nuanced response options for their children's pronunciations. We ran statistical models similar to the one described above but with the graded rather than the binary responses as the dependent variable. We hypothesized that if parents were able to reliably judge their children's pronunciations, phonological production would play a stronger role in explaining vocabulary development in the model with the graded than with the standard response options. Unfortunately, parents, although found to be reliable judges of children's pronunciation (see below) were not found to be always consistent when completing multiple vocabulary inventories. We, thus, adapted the parent questionnaires in two ways: We combined responses from the standard and pronunciations forms of the MCDI to obtain information on production and pronunciation and used information from the pronunciation MCDI alone to obtain information on pronunciation. In the latter case, even if parents completed pronunciation information on words that they previously had indicated were not produced by their children, it is likely they were projecting their knowledge of their children's pronunciation onto these words. Overall, we found that the factors that influenced production and pronunciation versus pronunciation alone were relatively similar.

As was the case in the model based on production alone (see first aim), word frequency and phonetic complexity influenced production and pronunciation. What was new, however, was the stronger role of ND and phonological production in explaining model fit. In model 2, ND interacted with both vocabulary size and phonological production ability, as did WF. As vocabulary size increased, children were able to produce and pronounce well words of lower WF and ND. Similarly, as phonological production capacity increased, children were able to produce and pronounce well words of lower WF and ND. Given that both effects emerged as significant in the model, we assume that vocabulary size and phonological abilities modulated separately ND and WF and contributed unique variance. In addition, ND emerged as a significant main effect in Model 3. The importance of ND in influencing word production has been related to its role in reducing auditory memory demands; however, based on the current findings, we observe that saying a word that has the same sound sequences as many other words aids word pronunciation, presumably because the child practices producing the same sound sequences. These findings are consistent with other reports indicating that words from dense neighborhoods are produced more accurately and with less variability than words from less dense neighborhoods (Sosa, 2008; Sosa & Stoel-Gammon, 2012; Stoel-Gammon, 2011). Overall, we observed that ND played a role in whether a word was pronounced well but not in whether it was produced. As can be seen in Figure 5, words that were not produced had higher NDs than words that were produced but not well pronounced suggesting a complex relationship between ND, word production, and pronunciation.

As for phonological production, it emerged as a significant main effect in Models 2 and 3. This is a logical finding given that the dependent variable in these models was based on the judgments of the parents concerning their children's word pronunciations. Nevertheless, it was not foreseeable that parents could make accurate judgments of their children's productions. The results suggest that they can judge the pronunciation of their children very well. In these models, all three phonological measures contributed to model fit; however, PCC-Name contributed the most variance in Models 2 and 3. It was superior

to the phonological measures based on the NWR test and the spontaneous language sample. Thus, we observe that when our model included only a binary option (word produced or not), the measure that tapped phonological memory was the best one related to vocabulary development; whereas when our models included fine-grained response options on word pronunciation, the measure based on the word-naming task accounted for the greatest variance in word production and pronunciation. PCC-Name may have had the “edge” over the other phonological measures for pronunciation since it was designed to elicit all (or most) of the phonemes of French and thus more comprehensively measured speech sound production than the spontaneous sample or the NWR test.

In addition to using the parent-reported judgments as the dependent variable in the statistical models, we extracted a phonological measure from the MCDI-pronunciation, % well-pronounced. We found that it correlated well with vocabulary size (.63) and with all the other phonological variables (.67 - .79). Indeed, when parents provided orthographic examples of their children’s errors, it was possible to identify the phonological processes in the children’s speech, which could then be validated by examining the children’s IPA transcriptions in the different sampling conditions. [Table 9](#) provides examples of productions extracted from the MCDI-pronunciation, spontaneous language sample, and single-word naming task for three children. SD had poor phonological production results (only 38% of productions were judged as well-pronounced). His parents indicated orthographic forms consistent with assimilation, syllable deletion, and palatal fronting. AB3 was judged to have good phonology (94% of productions were well-pronounced). His parents indicated forms consistent with palatal fronting only. SGP was also judged to have good phonology (91% of productions were well-pronounced). The orthographic examples provided by the parent were consistent with cluster reduction, word-initially, -medially, and -finally. In all cases, the parents’ orthographic examples correlated well with the phonetic examples in the spontaneous language sample and word naming task.

Findings on monolingual versus bilingual children

In contrast to previous studies that have tested monolingual children only, this study included monolingual and bilingual children. Initial descriptive analyses revealed no differences between monolingual and bilingual children on French vocabulary scores. The monolingual children in this study had on average 404 words in their spoken lexicons, which is consistent with Kern and Gayraud’s (2010) finding that monolingual French children, aged 2;6, know on average 417 words. The fact that bilingual children knew a similar number of words ($n=412$) is a surprising result given that it is widely observed that bilingual children have smaller vocabularies than monolinguals when only one language is considered (Hoff, Core, Place, Rumiche, Señor, & Parra, 2012). We attribute this finding to the fact that the bilingual children were dominant in French. Indeed, a requirement of the study was that children needed to take part in a spontaneous language sample and denomination task and this would have necessitated a reasonably-sized active vocabulary. In all our statistical models, bilingual status did not emerge as significant on its own or in combination with any variable. These findings suggest that the same lexical and phonological factors influence monolingual and bilingual children alike, with the caveat that this finding applies to bilingual children dominant in French.

Limitations and implications

The study focused on lexical and phonological factors that influence whether a word is produced and pronounced well. We do not exclude that other factors not measured in our

Table 9. Examples of three children’s phonological errors from the different sampling methods

Phonological	MCDI-Pron	Spontaneous	Word-naming
Process		language sample	Task
SD, aged 2;0			
Assimilation	vanille ^a manille	bonhomme /bɔ̃nɔm/ [mɔ̃mɔ̃t]	lunette /lynɛt/ [nenɛt] drapeau /dʁapo/ [papo]
Syllable deletion	canard nard genou nou compote pote	revoir /ʁɔ̃vwaʁ/ [vaʁ], [waʁ]	fromage /fʁomaʒ/ [mat] stylo /stilo/ [lo] bouteille /bute:j/ [but]
Palatal fronting	chaussure saussure chaussette saussette		chapeau /ʃapo/ [sapo]
AB3, aged 2;2			
Palatal fronting	chips ssipss chaussette ssausset	accrocher /akʁɔʃe/ [akʁɔse]	chips /ʃips/ [sips]
	singe sinz nuage nuaz	mange /mɑ̃ʒ/ [mɑ̃ʒ] Genève /ʒənɛv/ [zənɛv]	singe /sɛ̃ʒ/ [sɛ̃ʒ] nuage /nyʒɔʒ/ [nyʒɔʒ]
SGP, aged 2;6			
	train tain traineau teineau tartine tatine tigre tigel	prénom /pʁenɔ̃/ [pʁɛnā] gros /gʁo/ [gɔ] serpent /sɛʁpɑ̃/ [sɛpā] chèvre /ʃɛvʁ/ [ʃɛb]	fromage /fʁomaʒ/ [fomaz] grenouille /gʁənuj/ [gɔnuj] zebra /zɛbʁ/ [zɛp]

^avanille ‘vanilla’; bonhomme ‘fellow/man’; lunette ‘glasses’, drapeau ‘flag’, canard ‘duck’, revoir ‘to see again’, fromage ‘cheese’, genou ‘knee’, stylo ‘pen’, compote ‘compote’, bouteille ‘bottle’, chaussure ‘shoe’, chapeau ‘hat’, chaussette ‘sock’, chips ‘chips’, accrocher ‘to hang’, singe ‘monkey’, mange ‘eat’, nuage ‘cloud’, Genève ‘Geneva’, train ‘train’, prénom ‘first name’, traineau ‘sleigh’, gros ‘big/large’, grenouille ‘frog’, tartine ‘slice of bread and jam’, serpent ‘snake’, tigre ‘tiger’, chèvre ‘goat’, zèbre ‘zebra’

study such as “imagery” and “concreteness” for word production (Jones & Brandt, 2019b; Verhagen et al., 2022), and “speech motor ability” for word pronunciation may also account for variance. We limited ourselves to the properties of one- and two-syllable nouns; however, future studies should widen the analyses to determine whether a different set of factors accounts for production and pronunciation when other grammatical classes are implicated. We also coded WF and ND according to adult-based norms and, in further studies, it would be important to use child-based norms to provide a more valid account of the influence of psycholinguistic properties on word production. Our sample size of 40 children was small – a larger group which includes bilingual children with varying degrees of dominance should be tested. Another limitation was that the SES of the parents/children tested may have been higher (in Geneva, Switzerland) than in other geographical regions. It would be important to determine whether all groups of parents, from both high and low SES, could provide reliable information on this questionnaire. In addition, the children were between two and three years of age and it is uncertain whether this questionnaire would work with a younger age group. In sum, more research is recommended with younger children and ones from more diverse SES backgrounds. Finally and as noted above, parents were not always consistent in

responding across multiple questionnaires. Our study, thus, confirms that reliability can be an issue when using parental report data (Feldman, Campbell, Kurs-Lasky, Rockette, Dale, Colborn, & Paradise, 2005), but that converging assessments and tailored coding schemes can help mitigate this issue.

Our findings indicated that parents were able to reliably judge their children's word pronunciations and the pronunciation data provided important information on the role of phonology in vocabulary development. Nevertheless, we could ask whether the standard version of the MCDI should be extended to include a pronunciation component. Anecdotally, parents complained of the excessive time required to fill out the questionnaires and complained of filling out the same questionnaire twice. Thus, further replication of this approach should ensure that production and pronunciation components appear on the same form to reduce administration time. Short versions of the MCDI have been found to be both reliable and valid (Mayor & Mani, 2019); hence, it may be possible to envisage both production and pronunciation components included in a short version of the MCDI. Another possibility is to construct a shortened version of the MCDI with words specifically selected to target a wide range of sounds and syllable structures. The MCDI-pronunciation could thus be used not only for the early screening of vocabulary but also for phonological production. Given that speech sound disorders are often associated with long-term academic and professional difficulties (Lewis, Freebairn, Tag, Igo, Ciesla, Iyengar, Stein, & Taylor, 2019; McCormack, McLeod, McAllister, & Harrison, 2009), detection of them at an early age would be beneficial for early targeted intervention. This measure could also inform us about individual profiles of vocabulary and phonological learning (e.g., some children may acquire words easily but pronounce them poorly, whereas other children may know few words but pronounce them well). Such information may be useful when working with clinical populations (e.g., children with developmental delays or children on the autism spectrum) who are difficult to test when employing standard methods. Interestingly, parents judged the speech of children, aged two to three years, to be well-pronounced 70 to 80% of the time which was a similar value to the PCCs obtained from our formal testing. Future research should determine whether this similarity was coincidental or would be found with other groups of children of varying ages and production capacities. If indeed this similarity holds out across varying ages and production abilities, the parent-based phonological measure (% words well-pronounced) would be a useful screening measure.

One other focus of this study was comparing the relationship between phonology and vocabulary using different sampling methods. Overall, we documented similar PCCs regardless of sampling methods. This finding is in agreement with several studies that show similar phonological abilities when children are tested via a language sample versus a single-word naming task (Masterson, Bernhardt, & Hofheinz, 2005; Morrison & Shriberg, 1992) and a single-word naming task versus a NWR task (Meziane & MacLeod, 2021). Nevertheless, the measures based on the NWR and the single-word naming task appeared to correlate best with vocabulary size and were the ones that emerged as significant in our statistical models, indicating that these measures tapped phonological processing skills more relevant to vocabulary development than the language sample. It is important to note that the NWR task contained 12 items and the denomination test contained 35 items, and were thus shorter than the language sample which consisted on average of 67 utterances; yet, they were the most sensitive measures. Nevertheless, a language sample is a more natural elicitation method than a single-word naming or NWR task. It yields valuable lexical, morphosyntactic, and pragmatic information in addition to phonological. As many authors argue, the utilization of multiple sampling methods of

phonological skills is important in the clinical setting to yield different perspectives on children's developing competencies (Morrison & Shriberg, 1992).

Conclusion

In sum, this study examined lexical and phonological factors that influence whether a word is indicated as produced or not and whether it is indicated as pronounced poorly, adequately, or well. We tested a group of French monolingual and bilingual children aged two to three years. Several important findings emerged. First, the psycholinguistic properties of the words themselves such as their WF, and to a lesser extent their phonetic complexity influenced whether a word was indicated as produced. Phonological production, as determined by a NWR task, accounted for only a small amount of unique variance in word production. Second, we found that parents could reliably inform us about production accuracy when they were asked to complete an experimental form of the MCDI. Statistical models based on their graded response options revealed an important contribution of ND and phonological production. In these models, phonological precision as measured by a single-word naming task was more closely related to vocabulary development than phonological precision based on a language sample. We also observed that phonological measures extracted from the pronunciation questionnaire were highly correlated with vocabulary size and other phonological measures and that the phonological examples the parents provided of their children's speech resembled the ones we documented in our recordings. Third, we did not document any effect of bilingualism in our statistical models suggesting that, in the current sample, lexical and phonological effects on vocabulary development in monolingual and bilingual children were similar. In terms of implications, this study highlights the need for a child-based corpus of WF and ND in French and underscores the potential utility of a pronunciation version of the MCDI. Taken together these data provide interesting perspectives on how phonological production interacts with the psycholinguistic properties of words in influencing vocabulary development. Phonology counts when explaining vocabulary development but primarily when one uses a fine-grained measure that taps production accuracy.

Acknowledgments. An early version of this study was presented at the Building Linguistic Systems conference in honour of Marilyn Vihman at the University of York in June 2022. We would like to recognize the contribution Marilyn Vihman has made to research on the relationship between phonological production and lexical development. We also thank all the parents and children who contributed to this study.

Competing interests. The authors declare none.

References

- Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67, 1–48.
- Brysbart, M., Madera, P., & Keuleers, E. (2018). The word frequency effect in word processing: An updated review. *Current Directions in Psychological Science*, 27, 45–50.
- Dos Santos, C., & Kern, S. (2015). Comparison between French monolingual and French–Portuguese simultaneous bilingual children: vocabulary size, grammatical distribution, and phonetic complexity. *Paper presented at the International Symposium on Monolingual and Bilingual Speech*, Chania, Greece.
- DuBois, E., & Bernthal, J. (1978). A comparison of three methods for obtaining articulatory responses. *Journal of Speech and Hearing Disorders*, 43, 295–305.

- Feldman, H., Campbell, T., Kurs-Lasky, M., Rockette, H., Dale, P., Colborn, D. K., & Paradise, J. (2005). Concurrent and predictive validity of parent reports of child language at ages 2 and 3 years. *Child Development*, *76*, 856–868.
- Fenson, L., Dale, P., Reznick, S., Thal, D., Bates, E., Hartung, J., Pethick, S., & Reilly, J. (1993). *MacArthur Communicative Development Inventories: User's guide and technical manual*. San Diego: CA Singular Publishing Group.
- Ferguson, C., & Farwell, C. B. (1975). Words and sounds in early language acquisition. *Language*, *51*, 419–439.
- Fletcher, P., Chan, C., Wong, P., Stokes, S., Tardif, T., & Leung, S. (2004). The interface between phonetic and lexical abilities in early Cantonese language development. *Clinical Linguistics & Phonetics*, *18*, 535–545.
- Gathercole, S. E. (2006). Non-word repetition and word learning: The nature of the relationship. *Applied Psycholinguistics*, *27*, 513–543.
- Gendler-Shalev, H., Ben-David, A., & Novogrodsky, R. (2021). The effect of phonological complexity on the order in which words are acquired in early childhood. *First Language*, *41*, 779–793.
- Gierut, J. A., & Dale, R. A. (2007). Comparability of lexical corpora: Word frequency in phonological generalization. *Clinical Linguistics & Phonetics*, *21*, 423–33.
- Glaspey, A., Wilson, J., Reeder, J., Tseng, W.-C., & MacLeod, A. (2022). Moving beyond single word acquisition of speech sounds to connected speech development with dynamic assessment. *Journal of Speech, Language, and Hearing Research*, *65*, 508–524.
- Goodman, J., Dale, P., & Li, P. (2008). Does frequency count? Parental input and the acquisition of vocabulary. *Journal of Child Language*, *35*, 515–531.
- Hansen, P. (2017). What makes a word easy to acquire? The effects of word class, frequency, imageability and phonological density on lexical development. *First Language*, *37*, 205–225.
- Hoff, E., Core, C., & Bridges, K. (2008). Non-word repetition assesses phonological memory and is related to vocabulary development. *Journal of Child Language*, *35*, 903–916.
- 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*, 1–27.
- Jakielski, K. J. (2000, June). *Quantifying phonetic complexity in words: An experimental index*. Paper presented at the Annual Child Phonology Conference, Cedar Falls, IA.
- Jones, G., & Macken, B. (2018). Long-term associative learning predicts verbal short-term memory performance. *Memory and Cognition*, *46*, 216–229.
- Jones, S., & Brandt, S. (2019a). Do children really acquire dense neighbourhoods? *Journal of Child Language*, *46*, 1260–1273.
- Jones, S., & Brandt, S. (2019b). Neighborhood density and word production in delayed and advanced learners. *Journal of Speech, Language, and Hearing Research*, *63*, 2847–2854.
- Kehoe, M., Chaplin, E., Mudry, P., & Friend, M. (2015). La relation entre le développement du lexique et de la phonologie chez les enfants francophones. *Rééducation Orthophonique*, *263*, 61–85.
- Kehoe, M., Patrucco-Nanchen, T., Friend, M., & Zesiger, P. (2018). The relation between phonological and lexical development in French-speaking children. *Clinical Linguistics & Phonetics*, *32*, 1103–1125.
- Kehoe, M., Patrucco-Nanchen, T., Friend, M., & Zesiger, P. (2020). The relationship between lexical and phonological development in French-speaking children: A longitudinal study. *Journal of Speech, Language & Hearing Research*, *63*, 1807–1821.
- Keren-Portnoy, T., Vihman, M., DePaolis, R., Whitaker, C., & Williams, N. (2010). The role of vocal practice in constructing phonological working memory. *Journal of Speech, Language and Hearing Research*, *53*, 1280–1293.
- Kern, S., & Gayraud, F. (2010). *L'inventaire français du développement communicatif*. Grenoble: Editions La Cigale.
- Kern, S., & Dos Santos, C. (2017). Invariance in variation. Frequency and neighbourhood density as predictors of vocabulary size. In M. Hickmann, E. Veneziano, & H. Jisa (Eds.), *Sources of variation in first language acquisition. Languages, contexts, and learners. Trends in Language Acquisition Research (TILAR)* (pp. 183–200). Amsterdam: John Benjamin.
- Lewis, B., Freebairn, L., Tag, J., Igo, R. Jr, Ciesla, A., Iyengar, S., Stein, C., & Taylor, H. G. (2019). Differential long-term outcomes for individuals with histories of preschool speech sound disorders. *American Journal of Speech-Language Pathology*, *28*, 1582–1596.

- Luce, P., & Pisoni, D. (1998). Recognizing spoken words: The neighborhood activation model. *Ear & Hearing*, *19*, 1–36.
- Maekawa, J., & Storkel, H. (2006). Individual differences in the influence of phonological characteristics on expressive vocabulary development by young children. *Journal of Child Language*, *33*, 439–459.
- Masterson, J., Bernhardt, B., & Hofheinz, M. (2005). A comparison of single words and conversational speech in phonological evaluation. *American Journal of Speech-Language Pathology*, *14*, 229–241.
- Mayor, J., & Mani, N. (2019). A short version of the MacArthur–Bates Communicative Development Inventories with high validity. *Behavior Research Methods*, *51*, 2248–2255.
- McCormack, J., McLeod, S., McAllister, L., & Harrison, L. (2009). A systematic review of the association between childhood speech impairment and participation across the lifespan. *International Journal of Speech-Language Pathology*, *11*, 155–170.
- Meziane, R., & MacLeod, R. (2021). Internal and external factors contributing to variability in consonant accuracy of Arabic–French simultaneous bilingual children. *Journal of Child Language*, First View.
- Morrison, J., & Shriberg, L. (1992). Articulation testing versus conversational speech sampling. *Journal of Speech and Hearing Research*, *35*, 259–273.
- Nakagawa, S., Johnson, P., & Schielzeth, H. (2017). The coefficient of determination R² and intra-class correlation coefficient from generalized linear mixed effects models revisited and expanded. *Journal of the Royal Society Interface*, *14*, 20170213.
- New, B., Brysbaert, M., Veronis, J., & Pallier, C. (2007). The use of film subtitles to estimate word frequencies. *Applied Psycholinguistics*, *28*, 661–677.
- Parra, M., Hoff, E., & Core, C. (2011). Relations among language exposure, phonological memory, and language development in Spanish-English bilingually developing two-year-olds. *Journal of Experimental Child Psychology*, *198*, 113–125.
- Paul, R., & Jennings, P. (1992). Phonological behavior in toddlers with slow expressive language development. *Journal of Speech, Language, and Hearing Research*, *35*, 99–107.
- Petinou, K., & Okalidou, A. (2006). Speech patterns in Cypriot-Greek late talkers. *Applied Psycholinguistics*, *27*, 335–353.
- R Development Core Team (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna.
- Rescorla, L., & Ratner, N. B. (1996). Phonetic profiles of toddlers with specific expressive language impairment (SLI-E). *Journal of Speech, Language, and Hearing Research*, *39*, 153–165.
- Rose, Y., & MacWhinney, B. (2014). The PhonBank initiative. In J. Durand, U. Gut, and G. Kristoffersen (Eds.), *The Oxford handbook of corpus phonology*. Oxford: Oxford University Press, pp. 380–401.
- Smith, B. L., McGregor, K. K., & Demille, D. (2006). Phonological development in lexically precocious 2-year-olds. *Applied Psycholinguistics*, *27*, 355–375.
- Sosa, A. (2008). Lexical effects in typical phonological acquisition. Unpublished doctoral dissertation, University of Washington, Seattle.
- Sosa, A., & Stoel-Gammon, C. (2012). Lexical and phonological effects in early word production. *Journal of Speech, Language, and Hearing Research*, *55*, 596–608.
- Stoel-Gammon, C. (2011). Relationships between lexical and phonological development in young children. *Journal of Child Language*, *38*, 1–34.
- Stokes, S. (2010). Neighborhood density and word frequency in toddlers. *Journal of Speech, Language, and Hearing Research*, *53*, 670–683.
- Stokes, S. (2014). The impact of phonological neighbourhood density on typical and atypical emerging lexicons. *Journal of Child Language*, *41*, 634–657.
- Stokes, S., Bleses, D., Basbøll, H., & Lambertsen, C. (2012a). Statistical learning in emerging lexicons: The case of Danish. *Journal of Speech, Language, and Hearing Research*, *55*, 1265–1273.
- Stokes, S., de Bree, E., Kerkhoff, A., Momenian, M., & Zamuner, T. (2019). Phonology, semantics, and the comprehension-expression gap in emerging lexicons. *Journal of Speech, Language, and Hearing Research*, *62*, 4509–4522.
- Stokes, S., Kern, S., & dos Santos, C. D. (2012b). Extended statistical learning as an account for slow vocabulary growth. *Journal of Child Language*, *39*, 105–129.
- Stokes, S., & Klee, T. (2009). Factors that influence vocabulary development in two-year-old children. *Child Psychology and Psychiatry*, *50*, 498–505.

- Storkel, H.** (2004). Do children acquire dense neighborhoods? An investigation of similarity neighborhoods in lexical acquisition. *Applied Psycholinguistics*, **25**, 201–221.
- Storkel, H.** (2009). Developmental differences in the effects of phonological, lexical and semantic variables on word learning by infants. *Journal of Child Language*, **36**, 291–321.
- Tuller, L.** (2015). Clinical use of parental questionnaires in multilingual contexts. In S. Armon-Lotem, J. De Jong, and N. Meir (Eds.), *Assessing multilingual children: disentangling bilingualism from language impairment*. Bristol: Multilingual Matters, pp. 301–330.
- Verhagen, J., Stiphout, M., & Blom, E.** (2022). Determinants of early lexical acquisition: Effects of word- and child-level factors on Dutch children's acquisition of words. *Journal of Child Language*, **49**, 1193–1213.
- Vihman, M.** (2022). The developmental origins of phonological memory. *Psychological Review*. Advance online publication. <https://doi.org/10.1037/rev0000354>
- Yeh, L.-L., & Liu, C.-C.** (2021). Comparing the informativeness of single-word samples and connected speech samples in assessing speech sound disorder. *Journal of Speech and Hearing Research*, **64**, 4071–4084.

Appendix A1. Descriptive information on the monolingual participants

Child	Age (months)	Gender	No. years obligatory schooling (mother)	French vocabulary (MCDI)
KS	24	M	8	247
IB	24	F	8	281
AA	31	F	10	556
OD	26	M	8	413
NB	33	F	8	549
LT	34	M	9	568
BdG	25	M	5	407
TB	32	M	7	122
IG	26	M	4	441
NM	30	F	5	209
CN	28	M	5	503
AM	32	F	8	562
LV	30	M	10	401
AD	24	M	10	513
VR	33	F	9	530
AL	23	M	8	83
KA	34	F	7	440
JS	30	M	7	436
IP	26	F	9	318
RD	36	M	3	568
LMD	28	F	4	339

Appendix A2. Descriptive information on the bilingual participants

Child	Age (months)	Gender	L1 ^a	Dominant language	Age of acquisition French ^b	No. years obligatory schooling (mother)	Total vocabulary
CB	32	F	Italian	French	–	9	983
AAB	37	M	Arabic	Arabic	No response	0	146
TO	33	F	Portuguese	French	–	10	527
LR	25	F	Spanish	French	–	11	674
AB	34	F	Italian	French	–	0	848
SG	24	F	Italian	French	–	12	264
JD1	32	M	English	French	11 months	6	595
AC	30	F	Catalan	French	–	5	723
JD2	25	F	German	French	–	10	521
FB	24	F	English	French	–	7	1109
SD	24	M	Spanish	French	–	9	358
RK	32	M	English	Equal	–	10	708
GM	29	F	English	Equal	–	8	557
MM	26	M	Italian	French	–	11	318
MW	28	F	Spanish	French	–	3	781
SGP	30	F	Spanish	French	–	8	671
AB2	25	F	English	French	–	9	963
AB3	26	M	Spanish	French	–	9	650
AS	37	F	Spanish	Equal	–	4,5	966

^aL1 = home language other than French.

^b= at birth.

Appendix B. Items in the single word naming task

Phoneme	Monosyllabic words	Bisyllabic words
/p/	pomme	poussette
/b/	bottes	bouteille
/t/	table	tortue
/d/	dent	drapeau
/k/	clé	carotte
/g/	gants	grenouille
/m/	main	maison
/n/	nez	nuage
/f/	fleur	fromage
/v/	vache	voiture
/s/	singe	stylo
/z/	zèbre	-
/ʃ/	chips	cheval
/ʒ/	jambe	girafe
/l/	lune	lunettes
/ʁ/	renne	raisin
/w/	oie	oiseau
/j/	yeux	yaourt/yoghourt

Appendix C. Items in the non-word repetition task

	Monosyllabiques	Bisyllabiques	Trisyllabiques
Essai	/gɔt/		
	/däb/		
Test	/dyl/	/malš/	/nupikœj/
	/bam/	/gabu/	/pileko/
	/rup/	/sutœʁ/	/tefilš/
	/sob/	/fɔnet/	/sikɔmal/

Cite this article: Kehoe M., Abu Laban A., & Lespinasse R. (2023). Which phonological and lexical factors best influence whether a word is produced and pronounced well? *Journal of Child Language* 1–35, <https://doi.org/10.1017/S0305000923000089>