JOURNAL ARTICLE

Variation in children's vowel production: Effects of language exposure and lexical frequency

Helena Levy[1] and Adriana Hanulíková[1,2,3]

[1] GRK "Frequency effects in language," University of Freiburg, DE

[2] German Department, University of Freiburg, DE

[3] Freiburg Institute for Advanced Studies (FRIAS), Freiburg, DE

Corresponding author: Helena Levy (helena.levy@frequenz.uni-freiburg.de)

According to usage-based models of phonology, the more frequently a word is used and perceived in accented pronunciation variants, the more exemplars of accented tokens are stored and then used for subsequent productions of this word. This may lead to greater production variability in speakers with more variable input than in speakers with less variable input (cf. Pierrehumbert, 2001). This contrasts with abstractionist theories and with proposals according to which children unconsciously filter out accent features. This study assesses the effects of variable input and lexical frequency on speech production by children (mean age 9;10) growing up with one or more languages and with exposure to regional varieties and foreign accents. In a picture-naming task, 60 children were tested on their production of eight German vowels. Children who experience more input variability produced more variable vowels in terms of greater Euclidean distances. Vowels in frequent words were produced with more variability than in infrequent words. Vowel position (F1) differed depending on language background (monolingual versus bilingual) and amount of input in regional varieties. The results imply that greater input variation can account for variable vowel production, in line with usage-based theories.

Keywords: Vowel production; school-aged children; bilingualism; regional varieties; foreign accents; lexical frequency; input variation

1. Introduction

The majority of research concerning the early acquisition of two languages has concluded that native-like pronunciation in both languages can be attained. In contrast, late bilinguals' speech production is often marked by a foreign accent in the second language. However, in recent years, studies increasingly have reported more variability in speech production of children exposed to more than one language from early on, as compared to children growing up with only one language (Darcy & Krüger, 2012; Gildersleeve-Neumann, Kester, Davis, & Peña, 2008; Khattab, 2007; McCarthy, Mahon, Rosen, & Evans, 2014; Marecka, Wrembel, Otwinowska-Kasztelanic, & Zembrzuski, 2015). Most of these studies examined vowel production, showing that children with different language backgrounds (henceforth bilinguals) produce vowels with greater variability than children who are exposed to mainly one language (henceforth monolinguals). Darcy and Krüger (2012), Bosch and Ramon-Casas (2011), and Khattab (2007) suggest that similarly to bilingual children, children exposed to foreign accents and regional varieties might also show a mutual influence of vowel categories from various varieties, yet they have not tested this hypothesis.

The aim of this paper is to examine the conditions under which variability in production occurs. First, we examine whether acoustic characteristics of German vowels produced by children aged eight to eleven years are modulated by the child's exposure to more than one language. Second, we examine whether input variability due to regional and foreign accents leads to greater production variability in monolingual and bilingual children. Since vowel articulation tends to be affected by word frequency, our third aim is to investigate whether vowels in high-frequency words are produced with greater variability than the same vowels in low-frequency words.

1.1. Effects of language background on vowel production

Prior research on variability in vowel production in monolingual and bilingual children is limited and inconclusive. Darcy and Krüger (2012) examined the influence of bilingual input on speech production in early Turkish-German bilingual and German monolingual children of primary school age. The bilingual children showed greater variability in the localization of the vowels (first two formants), which could be a consequence of greater input variability stemming from different languages and possibly foreign accents. Bosch and Ramon-Casas (2011) tested distinctiveness in production of Catalan front vowels between two different groups of predominantly Catalan-speaking bilingual adults using formant analyses. One group consisted of adults who were raised in monolingual Catalan homes and were first exposed to Spanish and Catalan from the age of four years. Participants in the second group were raised in bilingual Spanish-Catalan homes, or had been exposed to both languages before the age of three years. In contrast to the first group, speakers of the second group more often used the wrong vowel category in Catalan target words. The authors therefore concluded that early exposure to variation can lead to less stable lexical representations, thus affecting vowel production. Baker, Trofimovich, Flege, Mack, and Halter (2008) and Baker and Trofimovich (2005) also report that bilingual children perform differently from native speakers in vowel production tasks, even if exposed to two languages intensively and from early on. Using a picture-naming task, Baker and Trofimovich (2005) showed that the interaction between two vowel systems in early Korean-English bilingual children results in acoustic differences when compared to monolinguals' vowels. However, Tsukada et al. (2005) found no differences in the production of Korean-English nine-year-old bilingual children who had been living in the United States for three-five years when compared to age-matched monolingual American English speaking children (see also Aoyama, Flege, Guion, Akahane-Yamada, & Yamada, 2004, and Oh et al., 2011, for Japanese-English bilingual children). Studies in which differences between early bilinguals and monolinguals were examined (Baker & Trofimovich, 2005; Fowler, Sramko, Ostry, Rowland, & Hallé, 2008) did not directly assess vowel variability, though. Baker and Trofimovich (2005) used acoustic measures (differences in vowel position of English and Korean vowels), finding a bidirectional L1–L2 influence in early bilinguals. Similarly, Fowler et al. (2008) showed differing VOT durations for plosives as a result of the mutual influence of two languages in French-English adult bilinguals, which could be interpreted in terms of increased variability in bilingual speakers, even though the authors did not address the issue of variability.

Inconclusive results can also be seen in accent rating studies. Flege, Yeni-Komshian, and Liu (1999) report no differences between early Italian/English adult bilinguals and native English speakers in an accent intelligibility rating of Italian learners' vowels. In contrast, Khattab (2006, 2009) showed that despite having developed two separate phonetic systems by the age of five years, early Arabic-English bilinguals (English started at six months) lacked a full mastery of certain phonetic aspects by the age of ten years. In addition, bilingual children have been shown to display considerable variability in vowel-length

contrasts at the age of two-three years (Kehoe, 2002, for Spanish-German bilinguals) and at the age of nine-twelve years (Whitworth, 2000, for German-English bilinguals). In contrast to these results, De Houwer (2009, p. 179) claims that bilingual children over the age of three years usually sound like their monolingual peers. However, De Houwer admits that differences between monolinguals and bilinguals are possible due to differences in the input. Bilinguals often have only one or two speakers as representative(s) for one of their languages; they model their own speech after them and the outcome is probably different from speech in a monolingual setting with exposure to many model speakers. Taken together, it is unclear whether input in two languages leads to more production variability in children than input in only one language.

1.2. Effects of input variability due to regional and foreign accents

Similar to input in two languages, input in foreign accents and regional varieties might also lead to a mutual influence of vowel categories from various varieties. One of the few studies that addresses this issue was conducted by Khattab (2007), who measured children's input based on parental vowel production. She demonstrated a great range of variability in the parents' vowels, putatively contributing to the variability in production observed in their children. Khattab's study of English-Arabic bilingual children between the ages of five and ten years showed that children are able to switch between a more English-like and a more Arabic-like pronunciation as a result of sociolinguistic competencies. Depending on the situation (e.g., when code-switching into English from Arabic), their vowels were rated as more foreign-accented, suggesting that phonological overlap might be under active control of speakers as young as five years.

Research examining speech production of children who grow up with different regional varieties and/or foreign accents in their language community is still largely lacking. In Germany, children are often exposed to regional varieties as a result of living in areas where these varieties are spoken alongside Standard German.[1] In addition, they are also often exposed to non-native varieties and foreign accents as a result of migration and globalization. We are thus confronted with the question of how input variation may influence the pronunciation of monolingual and bilingual children who grow up with more than one input variety. Do children who receive more input variation due to regional varieties or foreign accents show greater variation in their production of German vowels than children who are exposed to less variable input?

1.3. Lexical frequency

In addition to factors of language exposure, such as the influence of language background and regional varieties or foreign accents, production variability may also vary as a function of word frequency. Evidence stemming from studies conducted on adults has demonstrated that frequent words are produced with more variability than infrequent words. For example, reduction is more prevalent in frequent words than in infrequent words (Jurafsky, Bell, Gregory, & Raymond, 2001; Pluymaekers, Ernestus, & Baayen, 2005). Gahl (2008) showed that more frequent members of homophonic word pairs are produced with shorter durations than their infrequent counterparts (see also Kang, Yoon, & Han, 2015). Moreover, Tomaschek, Tucker, Wieling, and Baayen (2014) showed that word frequency affects vowel articulation (movement patterns of the tongue body in German [a:]). Vowels in monosyllabic words were more centralized and thus produced with less effort in frequent words, whereas the opposite pattern was observed for disyllabic words. Bell, Brenier, Gregory, Girand, and Jurafsky (2009, p. 106) explain that frequency

---

[1] Standard German refers to a variety defined by a consensus among model speakers and writers, language experts, language codifiers, and language-norm authorities (Ammon, 2015, p. 57).

effects on articulation depend on how easily a lexeme can be accessed. Frequent words can be accessed faster because the articulatory plan is created more swiftly, which leads to shorter acoustic durations. Furthermore, several studies have investigated whether high-frequency words have more pronunciation variants (e.g., more reduction) than low-frequency words (Keating, 1998; Schertz & Ernestus, 2014). Schertz and Ernestus (2014) showed that vowel duration in the English definite article 'the' is shorter if it is followed by more frequent words compared to less frequent words, and that consonantal variation (such as substitutions, e.g., [ɾ], [z] for /ð/) is more likely in the context of frequent words than in the context of infrequent words. Taken together, the above studies suggest that word frequency affects vowel variability and that vowels in frequent words have shorter durations than in infrequent words. Here we examine whether school-aged children, similar to adults, show more production variability in frequent as compared to infrequent words.

1.4. *Theoretical accounts*

From a theoretical perspective, variation in production due to different languages and accents in the input as well as lexical frequency can be accounted for within exemplar-based approaches. Usage-based models of phonology (Bybee, 2001; Bybee & Hopper, 2001) and exemplar-based models of phonology (Bybee, 2007; Goldinger, 1998; Kirchner & Moore, 2012) propose that the lexical representation of a word is updated every time that word is encountered. According to usage-based models, "linguistic units are gradient categories that have no fixed properties but rather are formed on the basis of experienced tokens," and experience "thus has an ongoing effect on mental representation" (Bybee & Beckner, 2010, p. 830). In line with this idea, exemplar-based models assume that an exemplar is created for every perceived variant of a word. Lexical entries thus encode variation in pronunciation. Such models assume a direct acoustic-lexical mapping: A perceived word is compared to its stored exemplars and in turn updates the exemplar storage. Consequently, frequently heard words have more exemplars than infrequent words (Schweitzer et al., 2015). Frequency effects on production are explained by biases towards frequently heard variants. They are a "product of having episodic representations," and variation is a "natural consequence" of the model (Drager & Kirtley, 2016, p. 3f.). In addition, word frequency may influence production because of automatized processes in articulation, which can be more or less routinized (Bürki, 2018; Bybee, 2001; Pierrehumbert, 2002).

There is relatively little work that can explain how exactly representations relate to the phonetic shape of words during production. It is unclear, for example, whether speakers 'choose' one of the stored exemplars for production or whether a generalization process of several exemplars in the vicinity of a target exemplar takes place. Possibly, similar pronunciation variants of words form a cluster or 'cloud' around a lexical entry, and the speaker accesses one of these exemplars during production (Foulkes & Docherty, 2006; Pierrehumbert, 2001, 2016). According to Pierrehumbert (2001), single exemplars are weighted depending on frequency and recency of the occurrence in perception, which can then drive the individual's production. Other models describe how speakers generalize over the cloud of exemplars in order to generate output that is based on input (Kirchner, Moore, & Chen, 2010). Overall, models that explain how input variation influences speech production are rare (see Bürki, 2018, for an overview of how variation in the speech signal can be used to explain the cognitive processes behind production). Moreover, it is unclear whether bilinguals' representations of sounds in one language are influenced by exemplars from the other language. Amengual (2012) extends exemplar models to include phonetic interference between related sounds of two languages in the bilingual mental lexicon. He tested Spanish-Catalan bilinguals' production (picture naming) and recognition (lexical

decision) of a Catalan-specific mid-vowel contrast (/o/-/ɔ/) in order to probe if cognates enhance cross-linguistic influence. He found effects of cross-linguistic influence both in production and perception, with cognate status influencing vowel height and fronting. He interprets these findings with respect to exemplar-based models, assuming that the exemplar clouds representing each of the bilingual's languages are mostly distinct but overlap in the case of cognates; such that exemplars for both languages exist in the same perceptual space. These bilinguals' productions then draw from both Catalan and Spanish exemplars (as an average of the overlapping region in the exemplar space) instead of restricting targets to language specific exemplars. This proposal, however, specifically applies to cognates (see also Brown, 2015; Pallier, Colomé, & Sebastián-Gallés, 2001).

Clopper (2014, p. 80) points out that "exposure to multiple different varieties leads to more variable representations. These variable representations are defined by distributions with greater variance or bandwidth than the distributions of less variable input." Building upon this idea, speakers who are exposed to regional or foreign accents as well as a standard variety are expected to have stored several accented and standard exemplars and might use different ones in production, depending on situational or communicative circumstances. In agreement with these theoretical accounts, Vihman (1993) describes the emergence of a production bias in children: Recurrent patterns of the child's input, including its own productions and frequently heard variants are enhanced and reinforced. Foulkes and Docherty (2006) expand this notion by predicting that children's production will diverge when entering school and become influenced more strongly by exemplars from other children. Presumably, this could extend the size of the cloud of exemplars and lead to more variability in children's productions.

According to both usage-based and exemplar-based theories, pronunciation variability should be greater in children who experience more input variability (i.e., different accents or varieties) as compared to children who are exposed to less variability (e.g., mainly one variety). Taking into account all children, whether monolingual or bilingual, those who have more exposure to regional varieties and/or foreign accents should thus show greater variability in vowel production than those with less variable input. It is unclear whether this also applies to input in more than one language; that is, whether bilingual children will exhibit more variable vowel productions than monolingual children. Usage-based and exemplar-based models also predict greater variation in the representations of frequent words compared to infrequent words (or even different phonological representations for different pronunciation variants of a single word; cf. Bürki, Ernestus, & Frauenfelder, 2010). We therefore expect frequent words to show higher vowel variability than low-frequency words. Furthermore, in agreement with Gahl (2008) and Schertz and Ernestus (2014), frequent words should be produced with shorter relative durations than infrequent words.

An alternative view to exemplar-based models is provided by abstractionist accounts, which assume an abstract lexical representation for each word (Norris, 1994; Norris & McQueen, 2008; Levelt et al., 1999). The mapping of acoustic input onto an abstract representation is mediated at an abstract pre-lexical level. Thus, there is an early normalization and abstraction away from speaker or situation specific characteristics. According to abstractionist models, there is only one lexical representation for each word and "pronunciation variants are derived from this single abstract representation by means of general processes, which apply to several words" (Ernestus, 2014, p. 28). While such approaches can account for productivity and generalization processes, it remains unclear how exactly they speak to production variability as a result of input variation and frequency effects (Guy, 2014). There are only very few abstractionist models that account for variation in the lexicon (e.g., Ranbom & Connine, 2007), and they do not

make predictions about consequences for *pronunciation* variability. Mitterer and Ernestus (2008) argue that the link between perception and production is abstract, and that only phonologically meaningful phonetic detail is accommodated in lexical representations. A possible prediction from an abstractionist viewpoint would thus be that neither variable input due to different varieties or languages, nor lexical frequency should increase vowel variability.

Indirect evidence in favor of abstract representations comes from a proposal that children who grow up with more than one language (and who receive input from native-accented and foreign-accented speakers) are equipped with an "innate accent filter," which enables them to unconsciously filter out foreign accent features in the input and thus produce natively accented speech (Chambers, 2002, p. 121). In principle, the idea is compatible with the critical period hypothesis (Lenneberg, 1967), according to which children can only achieve a full command of language if they are presented with adequate input early in life. Chambers refers to what he calls the "Ethan experience," which he feels generalizes to many other children of immigrant parents. He describes a boy of pre-school age born in Toronto, Canada, who, despite both of his parents having a "medium-to-strong" eastern European accent when speaking English, shows no foreign accent features in his own speech. Chambers attributes this to an unconscious filter that leads Ethan to perceive standardly accented speech when hearing his parents' foreign-accented pronunciation. When his mother produces a tap /r/, the child would "hear it as retroflex and pronounce it that way" (Chambers, 2002, p. 122), without even perceiving a difference between his mother's and his own pronunciation. Following Chambers' accent-filter theory, there should be no difference in vowel variability between children with more and less accented input experience.

1.5. The current study

We tested school children in southern Germany, in a rural region of Baden-Württemberg, where the local variety spoken by the majority of the population is Swabian (Ammon & Loewer, 1977). Baden-Württemberg is a multi-cultural area with immigrants from almost 200 foreign countries and a range of regional varieties that are in active use and may at times lead to comprehension problems (Weber & Häuser, 2008, p. 28). Swabian vowels differ from Standard German vowels in several aspects, including contrasts in length, the lowering of the high vowels /ɪ/ and /i/ and the interchangeability of /ɛː/ and /eː/ (see Table A1 in the Appendix for a comparison between Swabian and Standard German vowels). All the monolingual[2] and bilingual children we tested were exposed to the Swabian variety and Standard German (through parents and/or at school, through parents of friends, or in free-time activities).

We used a picture naming task (cf. Darcy & Krüger, 2012) and compared the distribution of the first two formant frequencies of each of the German vowels /iː/, /ɪ/, /eː/, /ɛː/, /ɛ/, /aː/, /a/, and /uː/ (embedded in high and low frequency words) across groups of monolingual and bilingual children with different amounts of experience with accented speech. We examined whether children who receive more variation in the input due to (1) two different languages and (2) regional varieties or foreign accents show greater variation in their production of German vowels than children who are exposed to less

---

[2] To put it more precisely, one could argue that these monolingual children were bidialectal, since they were exposed to Swabian as well as to Standard German (through parents, school, or media in some cases). In this paper, we refer to this group as monolinguals for three reasons: a) it is a heterogeneous group with passive as well as active knowledge of different varieties of German, b) because unlike the bilingual group, they did not grow up using two distinct languages, and c) because many varieties are roofed over by one standard variety (and this also holds for varieties of other languages mentioned for the bilingual children below, for example Arabic, which are roofed over by the respective standard varieties).

variable input. We further examined (3) whether frequent words show greater variability in vowel realization as compared to infrequent words. Addressing these questions will allow testing predictions of the different models discussed earlier.

Vowel variability has been operationalized differently in previous studies, ranging from systematic changes in F1 and F2 values (cf. Nicolaidis, 2003) and the (relative) duration of vowels (cf. Darcy & Krüger, 2012) to deviance from the mean metrics. Differences in F1 and F2 values between children with more and children with less variable input would indicate that children's vowels are either less prototypical (cf. Kuhl & Iverson, 1995) or that they even represent different vowel categories (possibly due to the influence from another language).

In this study, variability was measured using Euclidean distances (cf. Chiswick & Miller, 2005; Leinonen, 2011; Pickl et al., 2014). In order to examine pronunciation variability, we measured acoustic characteristics of eight vowels by means of F1 and F2 formants. We then determined the Euclidean distance between each vowel token produced by a monolingual child to the mean F1 and F2 formant values of this vowel produced by all monolingual children, and for each vowel token produced by a bilingual child to the mean F1 and F2 values of this vowel by all bilingual children respectively. Variability is defined as a larger distance of vowels to the mean of the vowel category. Our measure of variability thus does not refer to variability within each child or vowel. We used Euclidean distances in the confirmatory analyses, which refer to our hypotheses that more input variability due to different languages and due to regional varieties or foreign accents may lead to greater pronunciation variability. In order to increase comparability with other studies, we also performed exploratory analyses on F1 and F2 formant values and relative vowel duration, for which we did not formulate explicit hypotheses.
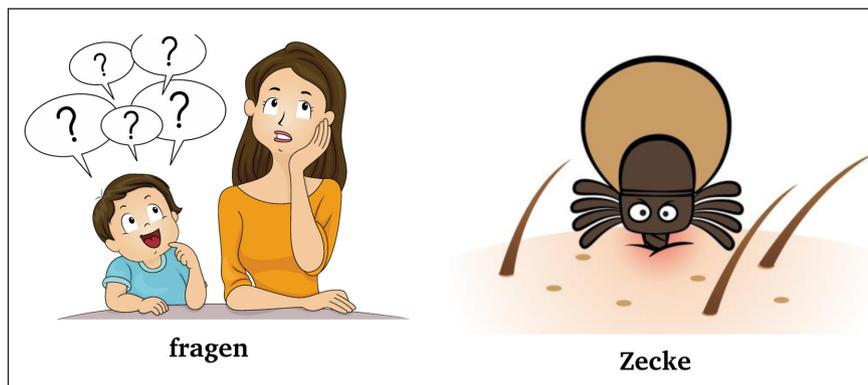
2. Methods

2.1. Participants

Twenty-seven bilingual (17 female, 10;0 years old, *SD* 0.7) and thirty-three monolingual (19 female, 9;9 years old, *SD* 0.85) children from the same primary school in southern Germany (Alb-Donau district) participated in the experiment. Monolingual children were defined as those children who grew up understanding and speaking only German (up to the age of six years, at which point some of them might have been enrolled in foreign language courses at school). Bilingual children were defined here as those who grew up understanding and speaking one or more language(s) in addition to German and started learning German before or at the age of three years. Five additional bilingual children were tested but not included in the final analysis because they were born outside of Germany and had moved to Germany between the ages of three and nine. All children were exposed to Swabian, the local variety spoken by the majority of the population. Most of the children had lived exclusively in this region (n = 57) or had spent their entire school time in this region. Most children had substantial experience with Standard German (if not from listening to their parents then from listening to teachers and TV programs). The bilingual children had various language backgrounds (Russian: n = 9, Turkish: n = 5, Albanian: n = 3, Serbian: n = 2, other languages: Portuguese, Spanish, Greek, Ewe, Arabic, Urdu, Croatian, Italian).

2.2. Stimuli and procedure

We used a picture-naming task without an auditory model (Baker et al., 2008; Darcy & Krüger, 2012) in order to elicit productions of eight German vowels (/iː/, /ɪ/, /eː/, /ɛː/, /ɛ/, /aː/, /a/, and /uː/) embedded in minimal or near minimal pairs (see Table A2 in the Appendix) consisting of one frequent and one infrequent word (e.g., *fragen* –

Figure 1: Example pictures with the labels *fragen* ('to ask') and *Zecke* ('tick').

Kragen, 'to ask' – 'collar').[3] Word frequency was taken from the childlex corpus (Schroeder et al., 2015), which lists 10 million words from over 500 children's books specifically for six to twelve-year-olds. Normalized lemma frequency was above 100 per million for high frequency words and below 30 for low frequency words (lexical frequency was operationalized as a continuous variable in all analyses). Two word stimuli were taken from Darcy and Krüger (2012; *Hand* 'hand,' *Biest* 'beast'). Twenty-eight words were taken from the vocabulary subpart of the 'cito-language test' (cito language test, 2015) and two additional words were taken from the childlex corpus. Every child produced sixteen minimal pairs (four words per vowel, with two frequent and two infrequent words) and every word was produced twice, adding up to a total of 64 words per child.

Stimuli were presented to the children in the form of cards in the format 99 mm × 99 mm (see examples in **Figure 1**). Pictures were colored line drawings of familiar words (taken from Shutterstock, www.shutterstock.com). There were two identical cards for each word, one labeled with the corresponding word and one unlabeled.

Subjects were tested one by one in a quiet classroom in their school. There were two experimenters present, one interacting with the child and the other one for documentation purposes. Children were shown the picture cards one by one. In the first round, we used labeled pictures in order to ensure familiarity with words that children were expected to produce. In the second round, the unlabeled pictures were used in order to elicit spontaneous productions. Children were recorded using both an Edirol R-09 audio recorder and the software Audacity (Audacity Team, 2014) on a Macbook Pro (44.1 kHz sampling frequency). The picture-naming task lasted approximately 15 minutes. Children also performed a battery of other tasks (hearing screening, vocabulary test, working memory, perception task after the production experiment), which were part of a different experiment and are not presented here. Parents completed an informed consent form several days before testing. Each child received a €5 voucher for participation in the experiment. The study was approved by the ethical committee of the University of Freiburg (application no. 73/16).

2.3. Experience with accents and languages

We operationalized Experience with Regional Varieties and Experience with Foreign Accents as continuous variables (cf. Porretta et al., 2016), quantifying each participant's weekly exposure to regional varieties and foreign accented speech in percent. Children's experience with accents and languages was assessed via a parental questionnaire prior to

---

[3] Nineteen stimulus pairs were prepared; however, three pairs had mistakenly been assigned to different frequency categories (high or low) even though both members corresponded to a single frequency category. Hence, we excluded these three pairs from further analyses.

Table 1: Amount (%) of weekly exposure to standard German, regional varieties, foreign accented German, and to other languages for monolingual and bilingual children according to the parental questionnaire.

	Monolinguals (n = 33)			Bilinguals (n = 27)		
	median	min	max	median	min	max
Weekly exposure to standard German	30.0	0	88.5	29.0	6.4	74.1
Weekly exposure to regional varieties	62.0	9.9	93.4	10.5	0	46.6
Weekly exposure to foreign accents	4.1	0	32.5	15.4	0	74.4
Weekly exposure to other languages	0	0	10.4	26.6	5.7	71

testing. Parental questionnaires were used in similar studies (Bent & Atagi, 2017; Darcy & Krüger, 2012; Van Heugten & Johnson, 2017). Validity of such questionnaires is often debated; we therefore assessed the reliability by using several questions on language and accent exposure. We obtained significant correlations between hours per week of input in a regional variety and percentage of input in a regional variety (r = .79, $p < .001$) as well as between hours per week of other-language input and percentage of other-language input (r = .72, $p < .001$). In order to find out how much experience children had with their various languages, varieties, and accents we calculated the number of hours per week that each child spends with a) Standard German, b) languages other than German, c) regional varieties of German, and d) foreign accented German (see De Houwer, 2017, on absolute input frequencies of multilingual children). We asked parents to indicate how many hours per week their children spend with each parent, with other adults, with relatives, with friends, at school, at free time activities and with media or on the phone, as well as which language or accent they are exposed to within these specific time periods. We also used a teacher questionnaire where teachers indicated their own variety used in interaction with the children. Only one teacher indicated the use of Swabian. For children in this class, school-time was then calculated as time spent with a regional variety (at this age, children usually spend most of their school hours with one teacher). We calculated a percentage value of the entire amount of waking hours spent with each language, regional variety, or foreign accent for each child as indicated by their parents. For each subject (monolinguals and bilinguals), we thus had one value for experience with regional varieties and one value for experience with foreign accents. For bilinguals, we also had one value for experience with their other language (see **Table 1**). Monolingual children had no experience with other languages with the exception of one child who, according to the questionnaire, did not speak or understand other languages but heard other languages spoken among others at the after-school daycare center. All bilingual children were reported to understand at least one language other than German, but the amount of input in both languages varied. There was one child with only 5.7% other language-input, but parents reported she understood Arabic. Even though all children were most likely exposed to both Standard German and to the Swabian variety to some extent, the parents of one child (who did not attend the class with the Swabian speaking teacher) reported he heard only Standard German, and the parents of another child (who attended the class with the Swabian speaking teacher) reported he had no exposure to Standard German.

2.4. Coding

Responses were transcribed and then annotated and analyzed acoustically using the software Praat (Boersma & Weenik, 2012, version 6.0). As in Darcy and Krüger (2012), we measured word durations and formant frequencies (F1 and F2) at the temporal mid-point of the vowel (which reduces the possible impact of coarticulation, see e.g., Bosch & Ramon-
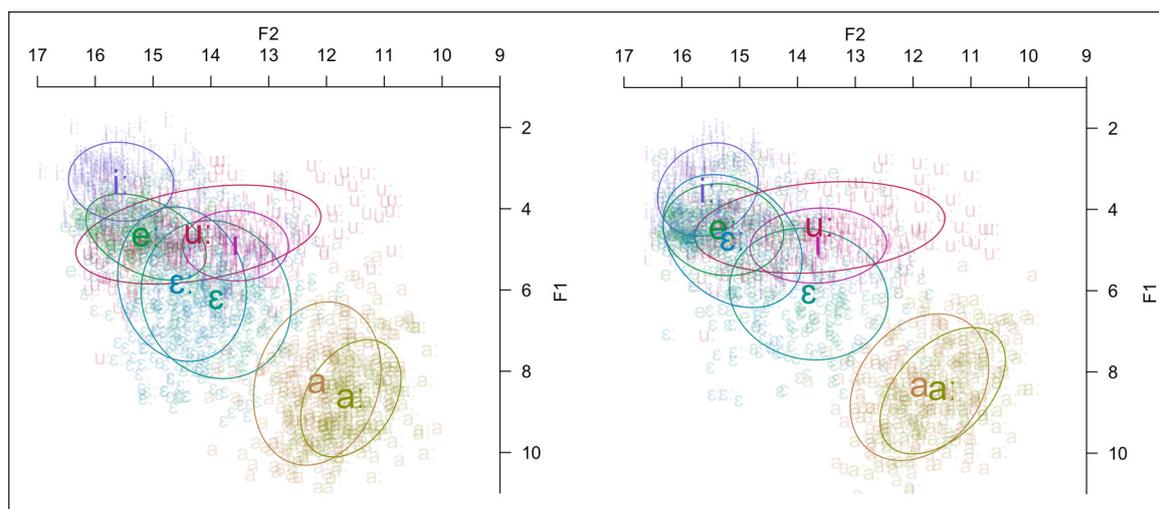
Casas, 2011). Using a custom Praat script (Lennes, 2017), the maximum formant value was set to 6000 Hz (Styler, 2011) and the number of formants was set to five. Prior to the analysis, all F1 and F2 values were normalized using the Bark difference metric method (Munson & Solomon, 2004; Zwicker & Ternhardt, 1980), using the *normalizeVowels* function from the package *phonR* (McCloy, 2016), which employs the Traunmüller (1990) formula. F1 and F2 thus correspond to Bark-transformed F1 and F2 values in Hertz. F1 and F2 frequencies above 2.5 standard deviations from each vowel's mean were measured again manually (9.8% of the tokens), adjusting the LPC settings with maximum formant values between 5000 Hz and 8000 Hz and five, six, or seven formants (see Derdemezis et al., 2016). Hand-correction was necessary for example in cases where Praat identified the F0 or F2 formant as F1, merged two formants that were very close together, or identified spurious formants. When recording quality was bad, tokens were excluded from the analysis, as were extreme outliers (more than ± 2.5 standard deviations from the mean for each vowel). Out of 3840 possible vowels (60 children × 8 target vowels × 4 stimulus words × 2 trials), 620 tokens (16.1%) had to be discarded, leaving 3220 vowel tokens for the acoustical analysis (all of the children retained more than 50% of the tokens). We determined relative durations from the ratio of absolute durations of each vowel token to the mean duration of all vowels in the group of monolingual children and bilingual children respectively.[4]

3. Results

3.1. Descriptive results

A descriptive summary of acoustic characteristics (F1 and F2 formants) of monolingual and bilingual children's vowels is provided in **Figure 2**.[5] Each F1/F2 value (in Bark) is marked by a single IPA vowel symbol. As can be seen in **Figure 2**, the vowel spaces of monolingual and bilingual children are largely similar with respect to vowel variability.

For several of the vowels (/aː/, /ɪ/, and /uː/), however, F2 values in bilinguals appear more widely spread. Within both the monolingual and the bilingual group, high overlap is visible in the vowels /eː/, /ɛː/, and /ɛ/, as well as /aː/ and /a/, suggesting a shared space possibly due to overlapping categories for these vowels (Ammon & Loewer, 1977). This is not surprising, taking into account that all children live in the Swabian dialect area

Figure 2: F1 and F2 formants in Bark for all vowels in monolinguals (left, n = 33) and bilinguals (right, n = 27). Ellipses show ±1 standard deviation.

---

[4] A second way to determine relative duration would be the ratio of absolute duration to a mean for each vowel within each subject (Wells, 1962; Porzuczek, 2012). This procedure led to the same results.

[5] The graph was created with the *phonR* package in R (McCloy, 2016).

Table 2: Average Bark difference scores (F1 and F2) for monolinguals and bilinguals.

	Monolinguals				Bilinguals			
	F1	SD	F2	SD	F1	SD	F2	SD
eː	4.697	0.695	15.116	0.686	4.496	0.739	15.274	0.685
ɛ	6.250	1.283	13.905	0.855	6.067	1.071	13.813	0.905
ɛː	5.852	1.242	14.495	0.731	4.803	1.090	15.108	0.804
aː	8.664	0.948	11.581	0.570	8.456	1.020	11.488	0.709
a	8.307	1.321	12.161	0.726	8.380	1.185	11.891	0.788
ɪ	4.909	0.570	13.572	0.601	4.889	0.607	13.639	0.775
iː	3.335	0.639	15.552	0.598	3.509	0.756	15.556	0.566
uː	4.634	0.799	14.217	1.389	4.443	0.738	13.616	1.417

where these vowels are often interchangeable. The figure also shows that /uː/ displays the greatest amount of dispersion in F2 values. The Bark values (means and standard deviations) for each vowel in each group are summarized in **Table 2**. F1 frequency values correlate negatively with the position of the highest part of the tongue (open vowels have higher F1 values than closed vowels), while second-formant values correlate with the position of the tongue (front/back), with front vowels showing higher F2 values than back vowels.

3.2. Confirmatory statistics

Linear mixed-effects regression models were run using the function *lmer* from the R (R Core Team, 2016) packages *lme4* (Bates, Maechler, Bolker, & Walker, 2014) and *lmerTest* (Kuznetsova et al., 2016). All continuous predictors were z-standardized and group-mean centered in the individual groups (monolinguals and bilinguals) before running the models. The Euclidean distance was the dependent variable in all models. We used random intercepts for subject and item and random slopes for language background (mono-/bilingual) by vowel (nested within item). Thus, differences between monolingual and bilingual children were taken into account by fitting this random slope to each factor entering the model. Random slopes ensure that the results generalize to other items and participants (Barr, Levy, Scheepers, & Tily, 2013). Matuschek, Kliegl, Vasishth, Baayen, and Bates (2017) argue that model comparisons justify the selection of the random effects structure, aiming at the most parsimonious model (see also Stroup, 2012). We therefore used the function *anova* to compare the models against one another. We report only the results of the best model. Initially, we fitted the maximal model with the maximal random structure and first eliminated random factors and then fixed factors, always using model comparisons (following Zuur, Ieno, Walker, Saveliev, & Smith, 2009). We removed all non-significant predictors one-by-one until the model contained only predictors that significantly contributed to the model fit (backward fitting procedure; see also Rathcke & Smith, 2015). We used sum coding, which means that the estimates in the following tables are in contrast to the grand mean and not to a reference condition. Within-subject variability is high in children under the age of twelve years (Lee, Potamianos, & Narayanan, 1999), which is why we initially performed analyses over all vowels.

The maximal model included a three-way interaction between the predictor variables Language Background (mono-/bilingual), amount of Experience with Regional Varieties, and amount of Experience with Foreign Accents. The model also included the predictor Lexical Frequency, and the variables Age and Sex as previous production studies with children of a similar age-group showed higher formant frequencies for females and for younger children (Vorperian & Ken, 2007; Huber, Stathopoulos, Curione, Ash, & Johnson,

1999). The fixed factors Language Background (mono-/bilingual), Age, and Sex were not significant in the first full model. The three-way interaction was also not significant and the model was better without the three-way interaction. We thus removed the three-way interaction. In a second model with only two-way interactions, only the interaction between Experience with Regional Varieties and Experience with Foreign Accents was significant. In a further model with only this interaction and the fixed factors Lexical Frequency and Language Background (mono-/bilingual), Language Background was not a significant predictor (β = –0.04, t = –1.222, p = 0.226), nor did it contribute to the model fit and was thus removed (as were the predictors Age and Sex). This suggests that the effect of variable input (input in two varieties) on vowel variability cannot be exclusively explained by the monolingual/bilingual status.
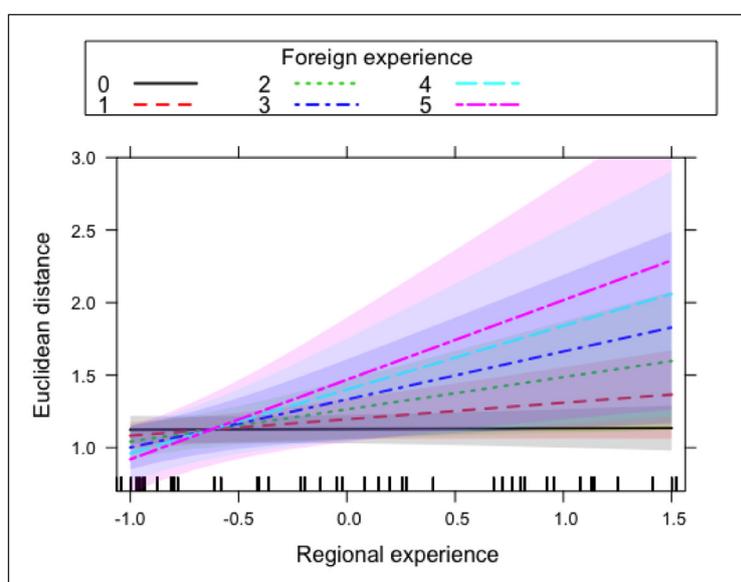
The best fitting model therefore included the interaction between Experience with Regional Varieties and Experience with Foreign Accents as well as the predictor Lexical Frequency. **Table 3** shows the summary of the regression model.

There was no significant effect of the single variables Experience with Regional Varieties and Experience with Foreign Accents; only the interaction between these two was significant (see **Figure 3**). Children with more experience with *both* regional and foreign

Table 3: Summary of the mixed-effects regression model with Euclidean distance as outcome variable.

	β	SE	t	p
(Intercept)	1.14	0.051	22.364	<0.001
Experience with Regional Varieties	–0.014	0.034	–0.404	0.688
Experience with Foreign Accents	0.079	0.041	1.917	0.060
Lexical Frequency	0.098	0.034	2.908	<0.01
Experience with Regional Varieties * Experience with Foreign Accents	0.112	0.039	2.862	<0.01

*Formula: Euclidean distance ~ Experience with Regional Varieties * Experience with Foreign Accents + Lexical Frequency + (1|subject) + (1+Language Background|vowel:item)*

Figure 3: Interaction effect between Experience with Regional Varieties and Experience with Foreign Accents, plotted as percentage (z-standardized) of input in regional varieties and foreign accents per week. The values for Experience with Foreign Accents are plotted in six equally spaced levels (level 0 corresponds to 0% foreign accent experience, level 5 to 74%).

accents showed greater Euclidean distances than children with less experience with both accent types. Thus, only a combination of experience with regional and foreign accents led to greater variability in vowel production. This result is in line with our hypothesis that greater input variability due to different accents may lead to greater pronunciation variability. Furthermore, vowels in lexically frequent words were produced with greater Euclidean distances than in infrequent words. This confirms our hypothesis that greater lexical frequency leads to greater variability in the production of vowels.

3.3. Exploratory statistics

3.3.1. Language background and accent experience: 'F1/F2'-analysis

In order to increase comparability with other studies, we also analyzed whether more input variability leads to different vowel positions (F1/F2) by running models with the outcome variables F1 and F2. As in the confirmatory analysis, we derived the models via stepwise model comparisons and removed all non-significant predictors one-by-one (Rathcke & Smith, 2015). We used random intercepts for subject and item, and random slopes for Language Background (mono-/bilingual) by Vowel (nested within *item*). Model results are displayed in the Appendix.

The maximal F1 model over all children contained a three-way interaction between the predictor variables Language Background (mono-/bilingual), amount of Experience with Regional Varieties, and amount of Experience with Foreign Accents, as well as the predictor Lexical Frequency, and the variables Age and Sex. The three-way interaction and the predictor Lexical Frequency did not yield significant results and did not contribute to model fit and were thus removed. The best fitting F1 model over all children (see Table A3) included an interaction between Experience with Foreign Accents and Language Background, as well as the predictors Experience with Regional Varieties, Age, and Sex. As can be seen in Table A3, children with more experience with regional varieties produced lower F1 values than children with less experience with regional varieties. This suggests that they produced more closed and more fronted vowels, possibly due to Swabian influence or influence from their other languages. Children with more foreign accent experience also produced vowels with lower F1 values than children with less foreign accent experience.

As the interaction between Language Background (mono-/bilingual) and Experience with Foreign Accents was significant in the F1 model, we ran separate models for monolingual and bilingual children. We report only the results for the best fitting models for monolingual and bilingual children. The best F1-model for monolinguals contained only the fixed factors Experience with Regional Varieties and Age (Table A4). F1-analyses for the monolingual group showed an effect of experience with regional varieties (lower F1 values, more closed and more fronted vowels, possibly due to Swabian influence). There are several reasons why the interaction between Language Background and Experience with Foreign Accents was significant in the model over all children but the factor Experience with Foreign Accents was not a significant predictor in the separate models. Less data can lead to different outcomes, and non-linear distribution of data points can make patterns visible only when all data is considered. In addition, monolinguals generally had little experience with foreign accents. For bilinguals, the best fitting model contained only the fixed factors Age and Sex. The F1 model for bilinguals did not yield significant effects of Experience with Regional Varieties or Foreign Accents (see Table A5).

The maximal model with F2 as an outcome variable contained the same interactions and predictors as the maximal F1 model. The three-way interaction (Language Background, amount of Experience with Regional Varieties, and amount of Experience with Foreign Accents) and the predictors Experience with Foreign Accents and Lexical Frequency were removed because they did not yield significant results and did not contribute to model

fit. The best model predicting F2 values over all children thus contained an interaction between Language Background and Experience with Regional Varieties, and the factors Age and Sex (see Table A6). The interaction between Language Background and Experience with Regional Varieties was significant, and thus we also ran separate models for monolinguals and bilinguals. The best F2 model for monolinguals contained only the fixed factor Age, and random intercepts for subject and for Vowel nested within item. No significant effects of Experience with Regional Varieties or Foreign Accents were found (see Table A7). For bilinguals (Table A8), the model contained the fixed factors Experience with Regional Varieties, Sex, and Age and had the same random effects structure as the F2 model for the monolinguals. More experience with regional varieties led to greater F2 values in bilingual children, which indicates greater variability, possibly due to the influence of overlapping vowel categories for vowels of the other language and due to greater accent experience. As these children are exposed to an additional language over and above Standard German and a German regional variety, they are exposed therefore to a large amount of input variability.

According to all F1/F2 models, female children produced vowels with higher F1 and F2 values as compared to male children, and older children generally produced vowels with significantly lower F1 and F2 values. This is not surprising, given that vocal cords tend to be longer in males than in females, correspondingly affecting frequencies. The significant effect of the factors Age and Sex on F1 and F2 values (despite Bark transformation) suggest differences in tongue position, which has an impact on the resonating cavity. These results are consistent with findings from other studies that have shown higher formant frequencies for females and for younger children (Fant, 1966; Traunmüller, 1984). Considering that children in this study were between 8;2 and 11;9, developmental differences are to be expected (cf. Vorperian & Kent, 2007; Huber et al., 1999). The results for Sex and Age in the separate models for monolinguals and bilinguals (Tables A4, A5, A7 and A8) were consistent with the results for the models over all children (lower F1 and F2 values in older children; females produced vowels with higher F1 and F2 values). As there was a fairly even distribution of male and female subjects across both groups of monolingual and bilingual children in our study, it is unlikely that the effects of language background and accent experience were influenced by these factors.

3.3.2. Lexical frequency: 'Relative duration'-analysis

The mean vowel duration for vowels in frequent words was descriptively shorter (161.44 ms, *SD* 65.33) than in infrequent words (163.73 ms, *SD* 66.89). **Table 4** shows that roughly half of the eight vowels across all children were descriptively slightly shorter in frequent words than in infrequent words.

The maximal model with relative duration as an outcome variable contained the same interactions and predictors as the maximal F1 and F2 models and, additionally, the predictor Vowel in an interaction with Lexical Frequency. We included Vowel as a fixed factor because we wanted to examine the effect of word frequency on the individual vowels in order to increase comparability with other studies. Since we applied sum coding, there is not one single vowel mapped onto the intercept. The intercept is the grand mean for all vowels. For every listed vowel in the model then, we can see how the specific vowel differs from the grand mean.[6] We were interested in the interaction between Lexical Frequency and Vowel because we wanted to know whether frequency

---

[6] In the models, due to the degrees of freedom, there is always one level of the factor not explicitly listed. This is the base group. The vowel /ɛː/ served as the base group, because its mean differed the least from the grand mean.

Table 4: Average durations (in ms) for vowels in frequent and infrequent words across all children.

Vowel	Frequent words	SD	Infrequent words	SD
eː	154.08	50.20	165.29	50.88
ɛ	116.56	50.20	114.00	33.70
ɛː	180.88	53.13	197.36	56.76
aː	244.64	63.57	237.98	58.36
a	112.26	30.30	121.58	36.85
ɪ	116.57	43.74	109.86	32.85
iː	183.63	52.22	184.26	61.50
uː	184.90	55.59	193.50	70.29

affects only some vowels. The three-way interaction (Language Background, amount of Experience with Regional Varieties, and amount of Experience with Foreign Accents) as well as the predictors Language Background, amount of Experience with Regional Varieties, and Experience with Foreign Accents were removed because they did not yield significant results and did not contribute to model fit. Further reduced models without the three-way-interaction and without the two-way-interaction term (Vowel and Lexical Frequency) did not yield any significant results, nor did any of the fixed factors apart from the variable Sex. The only reliable effect in the regression model (see Table A9 in the Appendix for the last relevant model including the predictor Vowel) was that female subjects produced vowels with longer relative vowel durations.

4. Discussion

We examined how exposure to different languages and to regional varieties and foreign accents affects vowel production in school-aged children and whether vowels vary as a function of lexical frequency. We measured vowel formants of monolingual and bilingual children and used regression models to predict variability (expressed in Euclidean distance) depending on language background, experience with regional varieties and foreign accents, as well as lexical frequency. We will discuss each of the outcomes in turn.

4.1. Language background

We predicted that children who receive input in two languages would show greater Euclidean distances and thus more production variability. Contrary to our prediction, we did not find differences between monolingual and bilingual children in vowel variability. This null result is in line with Tsukada et al. (2005) and Oh et al. (2011), who found no differences between monolingual and early bilingual children in vowel production. Several previous studies, however, do find differences between monolingual and bilingual children. Darcy and Krüger (2012) found slightly greater variability in the localization of bilingual children's vowels as compared to monolinguals (although only for the vowels /a/, /aː/, and /eː/). In contrast to our study, they examined bilinguals with a homogenous language background (Turkish-German) and differences in vowel position were predicted based on the mutual influence of the two vowel systems. Similarly, Bosch and Ramon-Casas (2011) and Baker and Trofimovich (2005) found differences in vowel production between monolinguals and early bilinguals. However, they measured differences in the position of vowels (F1/F2) and did not use distance metrics that yield information on vowel variability concerning the dispersion of vowels, such as Euclidean distances.

In order to increase comparability with these studies, we performed exploratory analysis on F1 and F2 formant values and found differences in the F1 formant values between monolingual and bilingual children. Bilinguals showed lower F1 values (and there was a tendency for lower F2 values) in their vowel productions compared to monolinguals. This suggests that there are differences in vowel positions between monolingual and bilingual children, despite the early acquisition of German (all bilinguals were born in Germany). A possible explanation for this result is that the vowel systems of the bilinguals' other language influence German vowel categories, leading to less precise realizations of the German vowels. There is ample evidence that two languages in contact during acquisition influence how listeners perceive, discriminate, and categorize speech sounds (Strange, 1995; Cutler, 2012), and that perception affects subsequent productions (Flege, 2007; Paradis, 2001). Models of phonetic perception such as Flege's Speech Learning Model (SLM; Flege, 1995) and Best's Perceptual Assimilation Model (PAM; Best & Tyler, 2007) predict discriminability of phoneme categories by reference to the relationship between the phoneme repertoires of both languages in contact. These models mainly relate to speech perception; however, Flege (2007) specifically links the relationship of the two phoneme repertoires to production and suggests that bilinguals' phonetic subsystems have mutual effects on each other, assuming what he calls "a common phonological space." While this account does not directly consider exposure effects from listening to accented speech, it implies that the phonetic similarity between sounds in a bilingual's two languages is important for perception and production from an early age on. Transferred to simultaneous or early bilingual children, Flege's account would predict an influence by the other language on the production of German vowels, which is compatible with our result for F1 formant differences between monolingual and bilingual children.

Taken together, we did not find an effect of language background on vowel variability but we did find differences in vowel position. It is debatable whether such differences also reflect vowel variability (cf. Nicolaidis, 2003). The question arises whether different vowel positions in production also lead to stronger foreign accent features in the speech of bilinguals. We did not use accent ratings in our study but future research could examine how much of this variation is audible in children's speech. Studies that have employed accent ratings suggest that early bilinguals do not show accent features in their productions (Baker, Trofimovich, Mack, & Flege, 2002; Flege et al., 1999; Piske, Flege, MacKay, & Meador, 2002). Piske et al. (2002) found that early English-dominant bilinguals did not have an accent in either of their languages (English and Italian). This is in line with Chambers' observation that no foreign-accent features were audible in the speech of an English-dominant bilingual child, despite his parents' heavily accented English. This would apply particularly in the situation when the language of a community was acquired from very early on, as was the case in our bilingual children. While accent rating studies usually do not employ acoustic measurements, they consistently show a lack of a foreign accent in early bilinguals' speech. Thus, future studies could combine both methods to determine which acoustic parameters of vowels are more likely linked to perceptions of a foreign accent.

4.2. Effects of input variability (regional varieties and foreign accents)

Arguing from a usage-based perspective, we hypothesized that children with more experience with regional varieties and foreign accents would show greater variability in their vowel productions than children who hear mostly one variety or accent. Our results showed that children who had experience with *both* regional varieties and foreign accents showed greater variability in vowel production (expressed by greater Euclidean distances). This confirmed our hypothesis that more input variability due to regional

varieties and foreign accents leads to greater variability in the production of vowels. Increased experience with regional varieties or with foreign accents alone did not lead to greater vowel variability. It is reasonable to assume that input variability is accounted for best by exposure to different varieties and accents, and not by the amount of input in one variety. In a hypothetical setting, children who hear Swabian from one parent and foreign accented German from the other parent would be exposed to greater variability and thus show vowel productions with larger dispersions than children who hear Swabian or foreign accented German from both parents. The result that greater input variability leads to greater production variability is in line with several proposals.

Usage-based and exemplar-based models predict that the lexical representation of a word is updated every time the word is encountered (Bybee & Beckner, 2010). According to Clopper (2014), exposure to different varieties should thus lead to more variable representations with greater distributions (Clopper, 2014, p. 80); or possibly, even to different phonological representations for different pronunciation variants of a single word (cf. Bürki et al., 2010). This may lead to greater variability of production in speakers with more variable input than in speakers with less variable input (cf. Pierrehumbert, 2001). It is unclear, however, whether our subjects produced greater vowel variability because speakers 'chose' exemplars with different vowel realizations, or whether they generalized over the stored exemplars' vowel realizations, thus producing vowels with features that were merged from several exemplars in the vicinity of a target exemplar (Kirchner et al., 2010).

In contrast to the usage-based view, and in line with abstractionist theories, Chambers' (2002) accent-filter theory suggests that there should be no foreign accent features in the speech of bilingual children with more or less accented input, as accent features are filtered out during perception. While this theory originally addresses foreign accents only, the same reasoning could be applied to regional accents too. Our result of greater vowel variability in children with more variable input seems to contradict the accent-filter view, although we did not measure regional or foreign accent features, or the extent to which vowel variability contributes to being perceived as speaking with a regional or a foreign accent.

We also have to take into consideration that increased experience with one regional variety (as opposed to experience with more than one variety or accent) might not necessarily cause greater variability, especially as all children in our study lived in an area where one dialect is spoken by the majority of speakers. Children themselves probably produce a variety containing features from both Standard German and Swabian, which they use at school and among friends, and which could be viewed as an instance of dialect leveling (cf. Clopper, 2014). This is described in Francot, van den Heuij, Blom, Heeringa, and Cornips (2017, 94), who report an "extreme case of dialect leveling" in children of the same age group who develop an intermediate variety between Standard Dutch and the Limburgian dialect of their region. In our case, monolingual and bilingual children could be expected to produce such a leveled variety and, accordingly, to have formed rather stable lexical representations. Future studies could take into consideration which accents or varieties children hear from their peers and determine whether children's productions become more leveled (Francot et al., 2017) or more variable (Foulkes & Docherty, 2006) after entering school.

Our findings on vowel position (F1/F2) showed that increased experience with one regional variety or one foreign accent alone does not cause greater variability but can lead to different vowel positions as compared to speakers with less accent experience. Children with more experience with regional varieties (mostly Swabian) produced more closed and more fronted vowels. Children with more foreign accent experience produced

vowels with lower F1 values than children with less foreign accent experience. Separate analyses for monolingual and bilingual children suggested that more experience with regional varieties leads to different vowel positions in monolingual and bilingual children. Whereas monolinguals show more closed and more fronted vowels (F1 values), possibly due to influence from Swabian, bilinguals with a greater amount of input in a regional variety show higher F2 values. As our group of bilingual children consisted of children with various other language backgrounds it is unclear whether effects on vowel position stem from influence of the bilinguals' other language or from the influence of regional varieties and foreign accents, or both.

Related to this issue, the mutual influence of bilinguals' two different phonologies likely differs from the mutual influence of two related varieties, in particular when the respective phonological spaces differ substantially. The bilinguals in our study were exposed to both different varieties and different languages. An unresolved question remains: Which weighs heavier for production variability, being exposed to input in two different varieties or languages by merely hearing them, or actively speaking, as well as hearing, the different languages or varieties? Possibly, bilingual and bidialectal children understand but do not actively speak two languages or varieties. Overall, bidialectals are more frequently exposed to variation in related varieties than bilinguals, who split their exposure time between two different phonological systems. Since this study was not set up to compare fully functioning bidialectal children with bilingual children, future studies may address this issue by assessing active and passive knowledge of different languages and varieties as well as the respective phonological spaces.

Furthermore, we measured Euclidean distances between the target vowel in each token to the mean of this vowel category (for monolingual and bilingual children to the mean of their group, respectively). Measuring the distance of each produced vowel from a *speaker* mean would possibly show whether production variability is due to variability within the single speaker (as opposed to the group of speakers). This would, however, require more tokens per vowel from each speaker.

4.3. Lexical frequency

Evidence from studies conducted on adults suggested that vowels in frequent words would be produced with more variability and with shorter durations than in infrequent words (Jurafsky et al., 2001; Gahl, 2008). Our main finding was that vowels in frequent words were produced with greater variability (larger Euclidean distances from the mean of each vowel) than vowels in infrequent words. As implied by Pierrehumbert (2001), an increase in possible candidates can lead to greater variation in production. Words with higher lexical frequency (e.g., /leːʁɐ/ 'teacher') are perceived in more tokens in different variants, whereas low-frequency words (e.g., /leːdɐ/, 'leather') are perceived in only a limited amount of tokens. This effect was observed independently of children's language background.

Based on previous studies on the impact of lexical frequency on duration (Gahl, 2008), we expected shorter relative durations of vowels in more frequent words. Therefore, we also analyzed the relative duration of vowels as a function of lexical frequency and found that frequent words were not produced with shorter relative durations. Our results on vowel duration thus do not confirm the study by Gahl (2008), who showed that the more frequent members of word pairs are produced with shorter durations than their infrequent counterparts.

It is important to note that frequency values from a corpus (in our case a corpus of written language) can always only approximate reality. Words occur with different frequencies in

spontaneous speech, and every child has individual experiences with words. Furthermore, the consonantal context of the items we used might cause differences in vowel production, which might override effects of lexical frequency.

Taken together, our Euclidean distance measurements implied greater variability in frequent words than in infrequent words but frequent words were not produced with shorter relative durations than infrequent words. This suggests that greater lexical frequency accounts for more variability in vowel productions. In line with exemplar models, frequent words are more likely to have been perceived in different variants, and this may affect subsequent productions.

To conclude, the results of this study suggest that input variability leads to greater variability in the production of vowels, in line with usage-based phonology and exemplar theory. Children who had experience with *both* regional varieties and foreign accents showed greater variability in vowel production (measured by Euclidean distances). Exposure to a language other than German (bilinguals) did not lead to greater variability compared to monolinguals but we did observe different F1 formant values for monolinguals' and bilinguals' vowels, in line with several previous studies. These results are consistent with proposals according to which lexical representations in speakers who are exposed to more variable input exhibit a greater bandwidth than in speakers who experience less input variability (Clopper, 2014); the consequence being increased variability in pronunciation (Pierrehumbert, 2001; Darcy & Krüger, 2012; Khattab, 2007). Additionally, we replicate previous findings that frequent words are produced with more variability (greater Euclidean distances) than less frequent words by speakers with different accent and language backgrounds, as these words have been perceived in more tokens in different varieties. Overall, we have shown that input variability and lexical frequency can account for increased variability in vowel production. These results are difficult to explain without assuming the storage of individual word tokens, with rich acoustic detail, in a single lexicon used for comprehension and production.

Additional File

The additional file for this article can be found as follows:

- **Appendix.** This document includes tables that list the Swabian vowels, the picture-naming task stimuli, and summaries of the mixed-effects regression models. DOI: https://doi.org/10.5334/labphon.131.s1

Acknowledgements

Competing Interests

The authors have no competing interests to declare.

References

Amengual, M. (2012). Interlingual influence in bilingual speech: Cognate status effect in a continuum of bilingualism. *Bilingualism: Language and Cognition, 15*(3), 517–530. DOI: https://doi.org/10.1017/S1366728911000460

Ammon, U. (2015). On the social forces that determine what is standard in a language–with a look at the norms of non-standard language varieties. *Bulletin VALS-ASLA 2015,* 53–67.

Ammon, U., & Loewer, U. (1977). *Dialekt/Hochsprache–kontrastiv: Schwäbisch*. Düsseldorf: Schwann.

Aoyama, K., Flege, J. E., Guion, S. G., Akahane-Yamada, R., & Yamada, T. (2004). Perceived phonetic dissimilarity and L2 speech learning: The case of Japanese /r/ and English /l/ and/r/. *Journal of Phonetics*, *32*(2), 233–250. DOI: https://doi.org/10.1016/S0095-4470(03)00036-6

Audacity Team. (2014). Audacity(R): Free Audio Editor and Recorder [Computer program]. Version 2.0.2. Retrieved 7 April 2014 from http://audacity.sourceforge.net/.

Baker, W., & Trofimovich, P. (2005). Interaction of native- and second-language vowel system(s) in early and late bilinguals. *Language and Speech*, *48*(1), 1–27. DOI: https://doi.org/10.1177/00238309050480010101

Baker, W., Trofimovich, P., Flege, J. E., Mack, M., & Halter, R. (2008). Child–adult differences in second-language phonological learning: The role of cross-language similarity. *Language and Speech, 51*(4), 317–342. DOI: https://doi.org/10.1177/0023830908099068

Baker, W., Trofimovich, P., Mack, M., & Flege, J. E. (2002). The effect of perceived phonetic similarity on non-native sound learning by children and adults. In: B. Skarabela, S. Fish, & A. H.-J. Do (Eds.), *BUCLD 26: Proceedings of the 26th Annual Boston University conference on Language Development* (pp. 36–47). Somerville, MA: Cascadilla Press.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255–278. DOI: https://doi.org/10.1016/j.jml.2012.11.001

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2014). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1–48. DOI: https://doi.org/10.18637/jss.v067.i01

Bell, A., Brenier, J. M., Gregory, M., Girand, C., & Jurafsky, D. (2009). Predictability effects on durations of content and function words in conversational English. *Journal of Memory and Language, 60*(1), 92–111. DOI: https://doi.org/10.1016/j.jml.2008.06.003

Bent, T., & Atagi, E. (2017). Perception of nonnative-accented sentences by 5-to 8-year-olds and adults: The role of phonological processing skills. *Language and Speech*, *60*(1), 110–122. DOI: https://doi.org/10.1177/0023830916645374

Best, C. T., & Tyler, M. D. (2007). Nonnative and second-language speech perception: Commonalities and complementarities. In: M. J. Munro, & O.-S. Bohn (Eds.), *Language experience in second language speech learning: In honor of James Emil Flege* (pp. 13–34). Philadelphia: John Benjamins. DOI: https://doi.org/10.1075/lllt.17.07bes

Boersma, P., & Weenik, D. (2012). Praat: Doing phonetics by computer [Computer program]. Version 5.3.23. Retrieved 7 September 2012 from http://www.praat.org/.

Bosch, L., & Ramon-Casas, M. (2011). Variability in vowel production by bilingual speakers: Can input properties hinder the early stabilization of contrastive categories? *Journal of Phonetics, 39*(4), 514–526. DOI: https://doi.org/10.1016/j.wocn.2011.02.001

Brown, E. L. (2015). The role of discourse context frequency in phonological variation: A usage-based approach to bilingual speech production. *International Journal of Bilingualism, 19*(4), 387–406. DOI: https://doi.org/10.1177/1367006913516042

Bürki, A. (2018). Variation in the speech signal as a window into the cognitive architecture of language production. *Psychonomic Bulletin & Review*, *25*(6), 1973–2004. DOI: https://doi.org/10.3758/s13423-017-1423-4

Bürki, A., Ernestus, M., & Frauenfelder, U. H. (2010). Is there only one "fenêtre" in the production lexicon? On-line evidence on the nature of phonological representations of pronunciation variants for French schwa words. *Journal of Memory and Language*, *62*(4), 421–437. DOI: https://doi.org/10.1016/j.jml.2010.01.002

Bybee, J. (2001). *Phonology and language use.* Cambridge: Cambridge University Press. DOI: https://doi.org/10.1017/CBO9780511612886

Bybee, J. (2007). *Frequency of use and the organization of language.* Oxford: Oxford University Press. DOI: https://doi.org/10.1093/acprof: oso/9780195301571.001.0001

Bybee, J., & Beckner, C. (2010). Usage-based theory. In: B. Heine, & H. Narrog (Eds.), *The Oxford handbook of linguistic analysis* (1 ed.). Oxford: Oxford University Press. DOI: https://doi.org/10.1093/oxfordhb/9780199544004.013.0032

Bybee, J. L., & Hopper, P. J. (2001). *Frequency and the emergence of linguistic structure* (Vol. 45). Amsterdam: John Benjamins Publishing. DOI: https://doi.org/10.1075/tsl.45

Chambers, J. K. (2002). Dynamics of dialect convergence. *Journal of Sociolinguistics, 6*(1), 117–130. DOI: https://doi.org/10.1111/1467-9481.00180

Chiswick, B. R., & Miller, P. W. (2005). Linguistic distance: A quantitative measure of the distance between English and other languages. *Journal of Multilingual and Multicultural Development, 26*(1), 1–11. DOI: https://doi.org/10.1080/14790710508668395

CITO language test version 3. (2015). Digitale Sprachstandserhebung im Elementarbereich CITO Deutschland GmbH: http://www.de.cito.com/leistungen_und_produkte/cito_sprachtest

Clopper, C. G. (2014). Sound change in the individual: Effects of exposure on cross-dialect speech processing. *Laboratory Phonology, 5*(1), 69–90. DOI: https://doi.org/10.1515/lp-2014-0004

Cutler, A. (2012). *Native listening: Language experience and the recognition of spoken words.* Cambridge, MA: MIT Press. DOI: https://doi.org/10.7551/mitpress/9012.001.0001

Darcy, I., & Krüger, F. (2012). Vowel perception and production in Turkish children acquiring L2 German. *Journal of Phonetics, 40*, 568–581. DOI: https://doi.org/10.1016/j.wocn.2012.05.001

De Houwer, A. (2009). *Bilingual first language acquisition.* Clevedon, UK: Multilingual Matters. DOI: https://doi.org/10.21832/9781847691507

De Houwer, A. (2017). The role of language input environments for language outcomes and language acquisition in young bilingual children. In: D. Miller, F. Bayram, J. Rothman, & L. Serratrice (Eds.), *Bilingual cognition and language: The state of the science across its subfields* (pp. 127–154). Amsterdam: John Benjamins. DOI: https://doi.org/10.1075/sibil.54.07hou

Derdemezis, E., Vorperian, H. K., Kent, R. D., Fourakis, M., Reinicke, E. L., & Bolt, D. M. (2016). Optimizing vowel formant measurements in four acoustic analysis systems for diverse speaker groups. *American Journal of Speech-Language Pathology, 25*(3), 335–354. DOI: https://doi.org/10.1044/2015_AJSLP-15-0020

Drager, K., & Kirtley, M. J. (2016). Awareness, salience, and stereotypes in exemplar-based models of speech production and perception. In: A. Babel (Ed.), *Awareness and control in sociolinguistic research,* (pp. 1–24). Cambridge: Cambridge University Press. DOI: https://doi.org/10.1017/CBO9781139680448.003

Ernestus, M. (2014). Acoustic reduction and the roles of abstractions and exemplars in speech processing. *Lingua, 142*, 27–41. DOI: https://doi.org/10.1016/j.lingua.2012.12.006

Fant, G. (1966). A note on vocal tract size factors and non-uniform F-pattern scalings. *Speech Transmission Laboratory Quarterly Progress and Status Report, 1*, 22–30.

Flege, J. (1995). Second language speech learning: Theory, findings and problems. In: W. Strange (Ed.), *Speech perception and linguistic experience: Theoretical and methodological issues* (pp. 233–277). Baltimore: York Press.

Flege, J. (2007). Language contact in bilingualism: Phonetic system interactions. In: J. Cole, & J. Hualde (Eds.), *Laboratory phonology* (Vol. 9, pp. 353–382). Berlin: Mouton de Gruyter.

Flege, J. E., Yeni-Komshian, G. H., & Liu, S. (1999). Age constraints on second- language acquisition. *Journal of Memory and Language, 41*(1), 78–104. DOI: https://doi.org/10.1006/jmla.1999.2638

Foulkes, P., & Docherty, G. (2006). The social life of phonetics and phonology. *Journal of Phonetics, 34*(4), 409–438. DOI: https://doi.org/10.1016/j.wocn.2005.08.002

Fowler, C. A., Sramko, V., Ostry, D. J., Rowland, S. A., & Hallé, P. (2008). Cross language phonetic influences on the speech of French–English bilinguals. *Journal of Phonetics, 36*(4), 649–663. DOI: https://doi.org/10.1016/j.wocn.2008. 04.001

Francot, R. J., van den Heuij, K., Blom, E., Heeringa, W., & Cornips, L. (2017). Inter-individual variation among young children growing up in a bidialectal community. *Language Variation – European Perspectives, VI*, 85–98. DOI: https://doi.org/10.1075/silv.19.05fra

Gahl, S. (2008). Time and thyme are not homophones: The effect of lemma frequency on word durations in spontaneous speech. *Language, 84*, 474–496. DOI: https://doi.org/10.1353/lan.0.0035

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

Goldinger, S. D. (1998). Echoes of echoes? An episodic theory of lexical access. *Psychological Review, 105*(2), 251–279. DOI: https://doi.org/10.1037/0033-295X.105.2.251

Guy, G. R. (2014). Linking usage and grammar: Generative phonology, exemplar theory, and variable rules. *Lingua, 142*, 57–65. DOI: https://doi.org/10.1016/j.lingua.2012.07.007

Huber, J. E., Stathopoulos, E. T., Curione, G. M., Ash, T. A., & Johnson, K. (1999). Formants of children, women, and men: The effects of vocal intensity variation. *The Journal of the Acoustical Society of America, 106*(3), 1532–1542. DOI: https://doi.org/10.1121/1.427150

Jurafsky, D., Bell, A., Gregory, M., & Raymond, W. D. (2001). Probabilistic relations between words: Evidence from reduction in lexical production. *Typological Studies in Language*, 45, 229–254. DOI: https://doi.org/10.1075/tsl.45.13jur

Kang, Y., Yoon, T. J., & Han, S. (2015). Frequency effects on the vowel length contrast merger in Seoul Korean. *Laboratory Phonology, 6*(3–4), 469–503. DOI: https://doi.org/10.1515/lp-2015-0014

Keating, P. (1998). Word-level phonetic variation in large speech corpora. *ZAS Papers in Linguistics, 11*, 35–50. DOI: https://doi.org/10.1142/9781848160712_0003

Kehoe, M. (2002). Developing vowel systems as a window to bilingual phonology. *International Journal of Bilingualism, 6*(3), 315–334. DOI: https://doi.org/10.1177/13670069020060030601

Khattab, G. (2006). Phonological acquisition in Arabic–English bilingual children. *Phonological development and disorders: A cross-linguistic perspective*, (pp. 383–412). Bristol, Blue Ridge Summit: Multilingual Matters. DOI: https://doi.org/10.21832/9781853598906-017

Khattab, G. (2007). Variation in vowel production by English-Arabic bilinguals. In: J. Cole, & J. Hualde (Eds.), *Laboratory phonology* (Vol. 9, 383–410). Berlin: Mouton de Gruyter.

Khattab, G. (2009). Phonetic accommodation in children's code-switching. In: B. E. Bullock, & A. J. E. Toribio (Eds.), *Cambridge handbooks in linguistics. The Cambridge handbook of linguistic code-switching* (pp. 142–160). New York: Cambridge University Press. DOI: https://doi.org/10.1017/CBO9780511576331.010

Kirchner, R., & Moore, R. K. (2012). Modeling exemplar-based phonologization. In: Cohn, Fougeron, Huffman, & Renwick (Eds.), *The Oxford handbook of laboratory phonology* (pp. 332–344). Oxford: Oxford University Press. DOI: https://doi.org/10.1093/oxfordhb/9780199575039.001.0001

Kirchner, R., Moore, R. K., & Chen, T. Y. (2010). Computing phonological generalization over real speech exemplars. *Journal of Phonetics, 38*(4), 540–547. DOI: https://doi.org/10.1016/j.wocn.2010.07.005

Kuhl, P. K., & Iverson, P. (1995). Linguistic experience and the "perceptual magnet effect." In: Strange, W. (Ed.), *Speech perception and linguistic experience* (pp. 433–459). Baltimore, MD: York Press.

Kuznetsova, A., Brockhoff, P. B., & Bojesen-Christensen, R. H. (2016). lmerTest: Tests in linear mixed effects models. R Package Version 2.0-30. Retrieved September 6, 2017, from http://CRAN.R-project.org/package = lmerTest

Lee, S., Potamianos, A., & Narayanan, S. (1999). Acoustics of children's speech: Developmental changes of temporal and spectral parameters. *The Journal of the Acoustical Society of America, 105*(3), 1455–1468. DOI: https://doi.org/10.1121/1.426686

Leinonen, T. (2011). Aggregate analysis of vowel pronunciation in Swedish dialects. *Citerat, 5*, 75–95. Retrieved from https://www.journals.uio.no/index.php/osla/article/view/101

Lenneberg, E. H. (1967). The biological foundations of language. *Hospital Practice, 2*(12), 59–67. DOI: https://doi.org/10.1080/21548331.1967.11707799

Lennes, M. (2017). Spect – speech corpus toolkit for praat (Version 1.0.0). Retrieved from https://github.com/lennes/spect/releases.

Levelt, W. J., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences, 22*(1), 1–38. DOI: https://doi.org/10.1017/S0140525X99001776

Marecka, M., Wrembel, M., Otwinowska-Kasztelanic, A., & Zembrzuski, D. (2015). Do early bilinguals speak differently than their monolingual peers? Predictors of phonological performance of Polish-English bilingual children. In: E. Babatsouli, & D. Ingram (Eds.), *Proceedings of the international symposium on monolingual and bilingual speech 2015* (pp. 207–213). DOI: https://doi.org/10.13140/RG.2.1.3134.9845

Matuschek, H., Kliegl, R., Vasishth, S., Baayen, H., & Bates, D. (2017). Balancing type I error and power in linear mixed models. *Journal of Memory and Language, 94*, 305–315. DOI: https://doi.org/10.1016/j.jml.2017.01.001

McCarthy, K. M., Mahon, M., Rosen, S., & Evans, B. G. (2014). Speech perception and production by sequential bilingual children: A longitudinal study of voice onset time acquisition. *Child Development, 85*(5), 1965–1980. DOI: https://doi.org/10.1111/cdev.12275

McCloy, D. R. (2016). phonR: tools for phoneticians and phonologists. R package version 1.0-7. Package description on http://drammock.github.io/phonR/.

Mitterer, H., & Ernestus, M. (2008). The link between speech perception and production is phonological and abstract: Evidence from the shadowing task. *Cognition, 109*(1), 168–173. DOI: https://doi.org/10.1016/j.cognition.2008.08.002

Munson, B., & Solomon, N. P. (2004). The effect of phonological neighborhood density on vowel articulation. *Journal of Speech, Language, and Hearing Research, 47*(5), 1048–1058. Retrieved from DOI: https://doi.org/10.1044/1092-4388(2004/078)

Nicolaidis, K. (2003, August). Acoustic variability of vowels in Greek spontaneous speech. In: *Proceedings of the 15th international congress of phonetic sciences* (pp. 3221–3224). Retrieved from https://pdfs.semanticscholar.org/82e1/5dc27b0539dbd1f3718ae3665 9ffb2133cc7.pdf

Norris, D., 1994. Shortlist: a connectionist model of continuous speech recognition. *Cognition, 52*(3), 189–234. DOI: https://doi.org/10.1016/0010-0277(94)90043-4

Norris, D., & McQueen, J. M. (2008). Shortlist B: A Bayesian model of continuous speech recognition. *Psychological Review, 115*(2), 357–395. DOI: https://doi.org/10.1037/0033-295X.115.2.357

Oh, G. E., Guion-Anderson, S., Aoyama, K., Flege, J. E., Akahane-Yamada, R., & Yamada, T. (2011). A one-year longitudinal study of English and Japanese vowel production by Japanese adults and children in an English-speaking setting. *Journal of Phonetics, 39*(2), 156–167. DOI: https://doi.org/10.1016/j.wocn.2011.01.002

Pallier, C., Colomé, A., & Sebastián-Gallés, N. (2001). The influence of native-language phonology on lexical access: Exemplar-based versus abstract lexical entries. *Psychological Science, 12*(6), 445–449. DOI: https://doi.org/10.1111/1467-9280.00383

Paradis, J. (2001). Do bilingual two-year-olds have separate phonological systems? *International Journal of Bilingualism, 5*(1), 19–38. DOI: https://doi.org/10.1177/13670 069010050010201

Pickl, S., Spettl, A., Pröll, S., Elspaß, S., König, W., & Schmidt, V. (2014). Linguistic distances in dialectometric intensity estimation. *Journal of Linguistic Geography, 2*(1), 25–40. DOI: https://doi.org/10.1017/jlg.2014.3

Pierrehumbert, J. B. (2001). Exemplar dynamics: Word frequency, lenition and contrast. In: Joan Bybee, & Paul Hopper (Eds.), *Frequency effects and the emergence of linguistic structure* (pp. 137–157). Amsterdam: John Benjamins. DOI: https://doi.org/10.1075/ tsl.45.08pie

Pierrehumbert, J. B. (2002). Word-specific phonetics. In: C. Gussenhoven, & N. Warner (Eds.), *Laboratory phonology 7: Phonology and phonetics* (pp. 101–139). Berlin: Mouton de Gruyter.

Pierrehumbert, J. B. (2016). Phonological representation: Beyond abstract versus episodic. *Annual Review of Linguistics, 2*, 33–52. DOI: https://doi.org/10.1146/annurev-linguistics-030514-125050

Piske, T., Flege, J. E., MacKay, I. R. A., & Meador, D. (2002). The production of English vowels by fluent early and late Italian–English bilinguals. *Phonetica, 59*, 49–71. DOI: https://doi.org/10.1159/000056205

Pluymaekers, M., Ernestus, M., & Baayen, R. H. (2005). Lexical frequency and acoustic reduction in spoken Dutch. *Journal of the Acoustical Society of America, 18*, 2561–2569. DOI: https://doi.org/10.1121/1.2011150

Porretta, V., Tucker, B., & Järvikivi, J. (2016). The influence of gradient foreign accentedness and listener experience on word recognition. *Journal of Linguistic Phonetics, 58*, 1–21. DOI: https://doi.org/10.1016/j.wocn.2016.05.006

Porzuczek, A. (2012). Measuring vowel duration variability in native English speakers and polish learners. *Research in Language, 10*(2), 201–214. DOI: https://doi.org/10.2478/ v10015-011-0034-9

R Core Team. (2016). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Ranbom, L. J., & Connine, C. M. (2007). Lexical representation of phonological variation in spoken word recognition. *Journal of Memory and Language, 57*(2), 273–298. DOI: https://doi.org/10.1016/j.jml.2007.04.001

Rathcke, T. V., & Smith, R. H. (2015). Speech timing and linguistic rhythm: On the acoustic bases of rhythm typologies. *Journal of the Acoustical Society of America, 137*(5), 2834-2845. DOI: https://doi.org/10.1121/1.4919322

Schertz, J., & Ernestus, M. (2014). Variability in the pronunciation of non-native English *the*: Effects of frequency and disfluencies. *Corpus Linguistics and Linguistic Theory, 10*(2), 329–345. DOI: https://doi.org/10.1515/cllt-2014-0024

Schroeder, S., Würzner, K.-M., Heister, J., Geyken, A., & Kliegl, R. (2015). ChildLex: A lexical database of German read by children. *Behavior Research Methods, 47*, 1085–1094. DOI: https://doi.org/10.3758/s13428-014-0528-1

Schweitzer, K., Walsh, M., Calhoun, S., Schütze, H., Möbius, B., Schweitzer, A., & Dogil, G. (2015). Exploring the relationship between intonation and the lexicon: Evidence for lexicalised storage of intonation. *Speech Communication, 66*, 65–81. DOI: https://doi.org/10.1016/j.specom.2014.09.006

Strange, W. (ed.) (1995). *Speech perception and linguistic experience: Issues in cross-language research*. Baltimore: York Press.

Stroup, W. W. (2012). *Generalized linear mixed models: Modern concepts, methods and applications*. Boca Raton: CRC Press.

Styler, W. (2011). Using Praat for Linguistic Research. Version 1.4.5. Retrieved from http://savethevowels.org/praat/.

Tomaschek, F., Tucker, B. V., Wieling, M., & Baayen, R. H. (2014). Vowel articulation affected by word frequency. *Proceedings of the 10th international Seminar on Speech Production*, 429–432. Retrieved from http://www.sfs.uni-tuebingen.de/~hbaayen/publications/TomaschekEtAl2014.pdf

Traunmüller, H. (1984). Articulatory and perceptual factors controlling the age-and sex-conditioned variability in formant frequencies of vowels. *Speech Communication, 3*(1), 49–61. DOI: https://doi.org/10.1016/0167-6393(84)90008-6

Traunmüller, H. (1990). Analytical expressions for the tonotopic sensory scale. *Journal of the Acoustical Society of America, 88*, 97–100. DOI: https://doi.org/10.1121/1.399849

Tsukada, K., Birdsong, D., Bialystok, E., Mack, M., Sung, H., & Flege, J. (2005). A developmental study of English vowel production and perception by native Korean adults and children. *Journal of Phonetics, 33*, 263–290. DOI: https://doi.org/10.1016/j.wocn.2004.10.002

Van Heugten, M., & Johnson, E. K. (2017). Input matters: Multi-accent language exposure affects word form recognition in infancy. *Journal of the Acoustical Society of America, 142*(2), 196–200. DOI: https://doi.org/10.1121/1.4997604

Vihman, M. M. (1993). Variable paths to early word production. *Journal of Phonetics, 21*(1–2), 61–82.

Vorperian, H. K., & Kent, R. D. (2007). Vowel acoustic space development in children: A synthesis of acoustic and anatomic data. *Journal of Speech, Language, and Hearing Research, 50*(6), 1510–1545. DOI: https://doi.org/10.1044/1092-4388(2007/104)

Weber, R., & Häuser, I. (Eds.) (2008). *Baden-Württemberg. A portrait of the German Southwest.* Landeszentrale für politische Bildung Baden Württemberg. Renningen: Pfitzer.

Wells, J. G. (1962). A study of the formants of the pure vowels of British English (Doctoral dissertation, University of London). Available from https://www.phon.ucl.ac.uk/home/wells/formants/index-uni.htm

Whitworth, N. (2000). Acquisition of VOT and vowel length by English–German bilinguals: A pilot study. *Leeds Working Papers in Linguistics and Phonetics, 8*, 15–25.

Zuur, A. F., Ieno, E. N., Walker, N. J., Saveliev, A. A., & Smith, G. M. (2009). *Mixed effects models and extensions in ecology with R.* Book series by Gail, M., Krickeberg, K., Samet, J. M., Tsiatis, A., & Wong, W. (Eds.). New York, NY: Spring Science and Business Media. DOI: https://doi.org/10.1007/978-0-387-87458-6

Zwicker, E., & Terhardt, E. (1980). Analytical expressions for critical-band rate and critical bandwidth as a function of frequency. *Journal of the Acoustical Society of America, 68*(5), 1523–1525. DOI: https://doi.org/10.1121/1.385079