Language exposure predicts children's phonetic patterning:
Evidence from language shift

Margaret Cychosz

LANGUAGE EXPOSURE PREDICTS CHILDREN'S PHONETIC PATTERNING: EVIDENCE FROM LANGUAGE SHIFT

MARGARET CYCHOSZ

University of Maryland, College Park

Although understanding the role of the environment is central to language acquisition theory, rarely has this been studied for children's phonetic development, and RECEPTIVE and EXPRESSIVE language experiences in the environment are not distinguished. This last distinction may be crucial for child speech production in particular, because production requires coordination of low-level speech-motor planning with high-level linguistic knowledge. In this study, the role of the environment is evaluated in a novel way—by studying phonetic development in a bilingual community undergoing rapid language shift. This sociolinguistic context provides a naturalistic gradient of the AMOUNT of children's exposure to two languages and the RATIO of expressive to receptive experiences. A large-scale child language corpus encompassing over 500 hours of naturalistic South Bolivian Quechua and Spanish speech was efficiently annotated for children's and their caregivers' bilingual language use. These estimates were correlated with children's patterns in a series of speech production tasks. The role of the environment varied by outcome: children's expressive language experience best predicted their performance on a coarticulation-morphology measure, while their receptive experience predicted performance on a lower-level measure of vowel variability. Overall these bilingual exposure effects suggest a pathway for children's role in language change whereby language shift can result in different learning outcomes within a single speech community. Appropriate ways to model language exposure in development are discussed.*

Keywords: speech production, first language acquisition, field phonetics, morphology, language shift, Quechua, Spanish

1. INTRODUCTION. This study investigates how young children's bilingual language use and exposure predict their spoken phonetic development. Understanding the role of the language environment is foundational to child language acquisition, and to linguistics more broadly, with implications for language learning, transmission, and change (Cournane 2017, Cristia 2020, Meakins & Wigglesworth 2013, Smith, Durham, & Fortune 2009, Yang et al. 2017). Contemporary language acquisition theories ascribe different importance to the child's environment: on the one hand, theorists agree that the language-learning environment facilitates and can predict development, but on the other hand, they acknowledge discrepancies between exposure and children's observed speech-language patterns (Aslin & Newport 2009, Gagliardi & Lidz 2014).

What is the role of the language environment—the language that children are exposed to and learn from—for children's spoken phonetic development? Rarely has the environment been studied for children's phonetic outcomes (cf. Cristia 2011, Liu, Kuhl, & Tsao 2003), yet fine-grained measures of acoustic phonetics for children's speech could unearth subtle differences in environmental experience or rule them out more conclusively. Furthermore, when evaluating the role of the environment in phonetics, is it relevant to distinguish between children's receptive experience, such as the type or

quantity of child-directed speech a child hears, and expressive language experience, such as how often a child talks or the size of their expressive lexicon?

Complete answers to both of these questions require that researchers study children from different language environments—children with different receptive and expressive language experiences. In the past, the role of receptive versus expressive experience was evaluated by studying the influence of lexicon size or degree of phonological awareness on speech development (Caudrelier et al. 2019, DePaolis, Vihman, & Keren-Portnoy 2011, Noiray et al. 2019; but see Mayr, Howells, & Lewis 2015). In this study, the role of the language environment in phonetic development is evaluated in a novel way: by studying speech development in a bilingual community undergoing rapid language shift. Bilingual communities provide a gradient of child language experience because the proportion of each language used varies by child—even in situations of stable language transmission (Gathercole & Thomas 2009). The context of language shift is an even more unique study for the role of the language environment because receptive language abilities (comprehension) outlast expressive language abilities (production) during multigenerational language shift. Thus, in a cross-sectional study in a community undergoing language shift, there is an additional gradient of children's language experience: the ratio of comprehension to production experience, which may differ by language. For example, children may frequently receive input in the parents' (minoritized) language spoken in the home—which the children do learn to speak—but express themselves at school or with peers in the majority language, which dominates in the media and many sectors of public life. As a result, the roles of expressive and receptive language can be evaluated within individual children. This distinction is critical because, for child speech production outcomes, expressive experience in particular may be required to coordinate higher-level linguistic planning with low-level speech-motor coordination (DePaolis et al. 2011, Icht & Mama 2015, Zamuner et al. 2018). To acquire adult-like speech patterns, children need to practice not just hearing ambient speech but producing it: accessing the semantic content, decomposing it into words, morphemes, and phonemes, and then organizing their speech articulators to produce these linguistic units in the correct order with appropriate amounts of coarticulatory overlap. Thus, children's speech production patterns could vary by receptive language experience, like child-directed speech exposure, but they could also vary by expressive language experience, like the size of the expressive lexicon.

With regard to the potential role of the language environment for children's speech development, this article has two goals. First, it estimates the different bilingual language experiences of children simultaneously acquiring South Bolivian Quechua and Spanish: a highly ecologically valid corpus of daylong audio recordings of children's language environments, spanning over 500 hours of naturalistic, in-situ child language use and exposure, is efficiently annotated to estimate children's language dominance. Second, the children's bilingual language experiences are correlated with a series of speech production patterns, invoking phonetics, phonology, and morphology. Based on these, this article will demonstrate (i) whether individual differences in bilingual language exposure can predict variability in speech-language production in these children and (ii) what type of exposure—receptive versus expressive—best predicts this variability.

2. Background.
2.1. Environmental effects in language development. The speech and language that children are exposed to in their daily environments have been shown to predict individual differences in development. Speech input from adult caregivers predicts chil-

dren's lexical processing speed (Weisleder & Fernald 2013), syntactic complexity (Huttenlocher et al. 2002), expressive and receptive vocabulary sizes (Hoff 2003, Mahr & Edwards 2018), and phonological development (Cristia 2011, Ferjan Ramírez et al. 2019, Garcia-Sierra, Ramírez-Esparza, & Kuhl 2016, Liu et al. 2003). Infants and young children are capable of tracking statistical patterns from their input, such as phoneme or word cooccurrences, and reflecting those patterns during phoneme-discrimination and word-segmentation tasks (Maye, Werker, & Gerken 2002, Pelucchi, Hay, & Saffran 2009), consonant production (Edwards & Beckman 2008, Zamuner 2009), and early, prelexical vocalizations (de Boysson-Bardies & Vihman 1991, Ha et al. 2021).

Nevertheless, the extent of environmental effects upon speech and language development is far from clear. For one thing, there may be insufficient distributional information in children's ambient environments to derive all meaningful phonological and grammatical categories (i.e. POVERTY OF THE STIMULUS), necessitating other language-learning mechanisms (Swingley 2009, Yang 2004).[1] Another possibility is that, even if there were ample information, children may not take in all ambient cues (Lidz & Gagliardi 2015, Pearl & Lidz 2009). Instead, perceptual intake from the ambient language may interact with the child's current grammar and/or developmental stage. For example, Weisleder and Fernald (2013) found that a relationship between child-directed speech exposure and vocabulary size in two-year-olds was mediated by children's lexical processing speed. While Weisleder & Fernald 2013 is often cited in support of the idea that the environment is critical for language development, that study can likewise demonstrate how the ability to take language in at certain developmental time periods predicts later speech-language development.

A number of studies on children's artificial language learning and crosslinguistic development show similar discrepancies between input and output. Despite variable (or absent) patterns in artificial language training data, children still prefer harmonic vowels (Mintz et al. 2018), regularize word order (Culbertson & Newport 2015), and systematize determiner use (Hudson Kam & Newport 2009) during testing, suggesting that young learners bring substantial cognitive biases to the learning process (see Culbertson & Schuler 2019 for an overview). Similarly, many crosslinguistic and cross-cultural surveys of language development find that children's early speech-language milestones (e.g. the quantity of C-V vocalizations, early word combinations) are immune to exposure, at least before a certain age (Casillas, Brown, & Levinson 2020, Cychosz, Cristia, et al. 2021).

Consequently, there is strong evidence for relationships between language input and development, but there are likewise exceptions where children overregularize or have processing or domain-general biases that override statistical distributions in the ambient language. This study contributes meaningfully to our knowledge of the limitations and mechanisms of environmental influence on speech-language development by studying two topics that remain unexplored in this area. First, there has been almost no work evaluating a possible connection between the language environment and children's spoken phonetic production and variability (one exception is Foulkes, Docherty, & Watt 1999 and subsequent publications on the development of sociolinguistic variants). Yet the documented relationship between the language environment, especially caregiver input, and speech-language outcomes suggests that such a relationship between the environment

---

[1] Some characteristics of child-directed speech—hyperarticulation and an expanded vowel space, shortened sentences, isolated word production—are hypothesized to make linguistic categories more discoverable (e.g. Adriaans & Swingley 2017).

and phonetic variation is plausible. Finding a relationship between children's language exposure and phonetic patterning would suggest a mechanism for language change, especially sound change, stemming not from learning errors, as is sometimes assumed in models of child-driven language change, but instead from language contact (Kerswill & Williams 2000). Additionally, the fine-grained measures of acoustic phonetics are ideally suited to the study of environmental effects on child speech-language outcomes in a way that coarser measures, like the presence of certain grammatical features or lexicon size, are not. For one thing, phonetic patterning is highly malleable. In adults, phonetic drift is one of the first consequences of novel linguistic exposure such as second language learning (Chang 2012), and short-term phonetic change (i.e. accommodation) can occur even within individual conversations (Pardo 2006, Pardo et al. 2012) and among children (Nielsen 2014). Furthermore, in the hierarchy of language contact phenomena, sounds change and even phonologize well before grammatical categories (Thomason 2001, van Coetsem 1988). As such, phonetic measures have the potential to uncover subtler, more recent environmental effects—or rule them out more conclusively.

The second way that this work contributes to our understanding of environmental effects is by contrasting the TYPE of experience. These data come from a community undergoing language shift, meaning that receptive versus expressive experience varies by household, providing a gradient of experience within the population. The language environment is almost always instantiated as input from adult caregivers (receptive experience), which leaves open the possibility that the quantity and quality of children's OWN productions (expressive experience) could also influence speech outcomes (DePaolis et al. 2011, Icht & Mama 2015, Zamuner et al. 2018). Children's own speech production could be a vital component of the language environment because it reflects external stimuli: children's language production varies by interlocutor (parent, peer) and location (home, school). Finding an effect of children's own production experience would suggest that some of our models of the language environment are insufficient, at least for speech development. Here again phonetic production is an ideal dependent variable because it requires coordination of higher-level linguistic planning with low-level speech-motor planning. Consequently, if there were ever an effect of expressive experience, it would likely manifest in children's speech production first.

2.2. EXPOSURE EFFECTS IN BILINGUAL DEVELOPMENT. Children's bilingual development can depend on the relative amount of exposure received in the two languages. For example, Potter et al. (2019) found that bilingual Spanish-English toddlers, differing in language dominance, failed to recognize words from their nondominant language when those words were embedded in sentences spoken in the dominant language. Similar dominance effects have been documented for children's vocabulary development (Pearson et al. 1997, Place & Hoff 2016, Thordardottir 2011), while results on phonological processing are more mixed. Studies on the nonword repetition paradigm, for example, sometimes find dominance effects (Sharp & Gathercole 2013), but often do not (Brandeker & Thordardottir 2015, Core, Chaturvedi, & Martinez-Nadramia 2017, Farabolini et al. 2021) (see also Mayr et al. 2015 for phoneme accuracy).

Although evidence for the effects of bilingual dominance, and thus exposure, on children's phonetic patterning is fairly scarce (though see Bijeljac-Babic et al. 2012), these influences are well documented for adults in situations of language contact (Henriksen et al. 2019, Mooney 2019, Onosson & Stewart 2021, Simonet 2011, Yao & Chang 2016, inter alia). For example, Guion (2003) found that age of Spanish acquisition affected the ability of Quichua-Spanish speakers in highland Ecuador to partition the vowel space across the two languages.

Studies like Mayr et al. 2015 and Sharp & Gathercole 2013 evaluated effects of exposure on phonological development by studying communities with active minoritized language transmission.[2] However, language transmission as a lens into effects of exposure on phonological development can also be studied in child heritage language learners and caregivers who are second language speakers. Here there is more evidence for exposure effects on phonetic outcomes: for example, bilingual German-Dutch three- to six-year-olds follow the nonnative acoustic (voice onset time) patterns of their L2 Dutch mothers (Stoehr et al. 2019). Among child heritage language learners, Mayr and Siddika (2018) found successively larger effects of English on Sylheti stop production across generations of English-Sylheti heritage families in the UK. And Khattab (2003) hypothesized that the lack of a voicing lead in the speech of two bilingual Lebanese Arabic-English children (ages five and seven years) was caused by not receiving sufficient early input in the contrast and/or the lack of this cue in the mother's speech.

Many of the above studies established correlational relationships between bilingual language exposure and learning outcomes. However, outcomes are usually limited to coarse measures of lexical growth and phonological processing. The instantiation of the language environment was similarly limited: for the children, either bilingual language experience was quantified simply as input from caregivers via background questionnaire (e.g. Pearson et al. 1997, Place & Hoff 2011, Thordardottir 2011), or expressive and receptive experiences in the child's environment were not distinguished (e.g. Gathercole & Thomas 2009, Mayr et al. 2015, Stoehr et al. 2019).

The language environment may predict children's speech-language outcomes, but to test this, language exposure and use estimates from children's ambient environments must be quantified in a robust, reliable way from a large, naturalistic sample. The learning outcomes should also be sufficiently fine-grained to capture, or rule out, environmental influences. To that end, the current work estimates children's language exposure by annotating a large-scale, naturalistic child language corpus to estimate children's bilingual language exposure. Then, estimates derived from this corpus are compared to children's patterning on multiple measures of phonetic variability. Crucially, this research is conducted in a bilingual community undergoing language shift, where children's exposure to and use of their two languages, as well as the ratio of expressive to receptive language experiences, varies considerably by household.

3. CURRENT STUDY.

3.1. THE SPEECH COMMUNITY. Data for this study were collected from children simultaneously acquiring South Bolivian Quechua and Spanish in and around a mid-size town in the south Bolivian department of Chuquisaca. South Bolivian Quechua, henceforth 'Quechua', is a Quechua II-C language spoken by more than 1.6 million people in this region of Bolivia and northwest Argentina. (Quechuan languages have traditionally been divided into two primary genealogies: Quechua I and Quechua II, with further subdivisions within Quechua II that correlate with geographic location; Torero 1964.) The phonological inventory includes three phonemic /a, i, u/ vowels and two allophonic [e, o] vowels derived in uvular contexts (Gallagher 2016). Consonant contrasts include voiceless stops, aspirated stops, and ejectives at four places of articulation (/p, t, k, q/) and a three-way alveopalatal, stop-aspirated, stop-ejective distinction /tʃ, tʃʰ, tʃʼ/. Nasals are contrasted along three places of articulation: /m, n, ɲ/; the velar nasal [ŋ] is allophonic. Quechua is also a highly agglutinating language, with over 200 productive nominal and verbal suffixes that encode argument structure and grammatical relations.

---

[2] See also Gathercole & Thomas 2009 for morphosyntactic outcomes in these settings.

In southern Bolivia, Quechua and Spanish have been in intense contact for over four hundred years (Muysken 2019). Today, Spanish tends to dominate the media, educational, and political landscape in the country. The result is that, as in many situations of colonization, there is rampant language shift from Quechua and other Indigenous languages of the country to Spanish. Despite legislation encouraging the implementation of bilingual education in Bolivia, many public schools continue to be conducted primarily in Spanish (Hornberger 2009). Consequently, almost all school-age children who speak Quechua in the home in this region become bilingual in Spanish. Nevertheless, even if the medium in schools is Spanish, some Quechua-speaking teachers use Quechua vocabulary items with their students, and written Quechua words and vignettes are frequently introduced in Spanish-medium textbooks. In the community where I work, I have observed that most peer-to-peer interaction at school is conducted in Spanish (although this cannot be divorced from author positionality; see below). An additional consequence of these sociolinguistic and educational policies is that even though Quechua has an established writing system in Bolivia, children learn to read and write in Spanish, but many do not have the same opportunity in Quechua.

Quechua has a strong spoken presence in southern Bolivia, particularly among adults, but intergenerational transmission within families who—for reasons economic or otherwise—have relocated closer to towns or urban areas is highly variable. Consequently, the morphosyntax and phonology of the Quechua variety studied here are reported to differ significantly from varieties spoken in more rural areas of the department (Camacho Rios 2019). There are no reported statistics of intergenerational Quechua language transmission in this speech community, but on the basis of several summers of fieldwork in the area, I estimate an intergenerational transmission rate of approximately 50% (i.e. 50% of children in the area with Quechua-speaking parents learn to speak Quechua). All children in the current study were bilingual Quechua-Spanish speakers, but the ongoing language shift within the community meant that expressive and receptive language experiences, in addition to bilingual language dominance, varied by child. For example, a child could be Quechua-dominant if they speak predominantly Quechua in their everyday interactions and/or if most of their input in the home is in Quechua.

I have been conducting linguistic fieldwork annually in these communities since 2017, following a linguistic field methods course on South Bolivian (Cochabamba) Quechua at the University of California, Berkeley, in 2016, though this was halted due to COVID-19. During my fieldwork, I took on roles as language researcher and teacher/junior colleague at the primary school where I volunteered. Regarding researcher positionality, I am not a member of this speech community, but a white woman from the United States where, at the time of data collection, I was a Ph.D. student in linguistics. My foreignness was relatively uncommon in the speech community, making it all the more important to obtain naturalistic, unimpeded observations of language behavior (i.e. to avoid the observer's paradox). The overarching plan of the research program is to document ongoing synchronic and diachronic phonetic variation in these communities via controlled phonetic elicitation and naturalistic observation of child and adult behavior. The naturalistic observational data, in particular, are archived so that the community can benefit from these records of spontaneous, contextualized speech from adults and children with distinct language-learning backgrounds (see §4.2 and §6.4 for detail).

3.2. MEASURES OF PHONETIC VARIABILITY. Two measures of phonetic variability in the children's speech are studied to evaluate the role of language exposure: (i) within-

category vowel dispersion and (ii) V-C coarticulation across different word environments. These phonetic outcomes were chosen for the distinct demands they place on the developing speech apparatus and speech-planning capacities.

Infants establish early auditory-acoustic vowel categories on the basis of their ambient language during the first months of life (Werker & Curtin 2005). Thus, although vowels are some of the first sounds that infants produce (Oller 2000), acoustic-auditory perceptual categories predate vocalizations by several months in development. During the vocal exploration and babbling periods over the first months and years of life, infants compare their foundational acoustic-auditory categories to the acoustic consequences of their vocalic productions, or their AUDITORY FEEDBACK, and update their speech-motor plans accordingly. With further vocal production, feedback, and updates, infants and children can then approximate adult-like speech production and variability (Guenther 2006, Moulin-Frier, Nguyen, & Oudeyer 2014, Perkell 2012). The result is that, by early childhood, children have been honing their speech production, especially their vocalic production, for years.

In toddlerhood, the gradient nature of vowel acoustics, compared to the more quantal function that characterizes consonants (Stevens 1989), is relatively forgiving of the high acoustic and articulatory variability in child speech (Lee, Potamianos, & Narayanan 1999): a child who undershoots the articulatory target for /u/ may produce a more centralized or fronted /u/-like variant, but the same degree of undershoot for /l/ would likely result in a complete phonemic substitution like [w] or [j]. Finally, through middle childhood, vocalic development is characterized by a reduction in variability (Lee et al. 1999, Vorperian & Kent 2007), suggesting that speech-motor plans update continuously until at least puberty.

Overall, over the course of development, children get ample early practice refining their vocalic production because vowel-like vocalizations are some of the first sounds for which infants are able to incorporate auditory feedback and update speech-motor plans. Vowels are rarely a source of children's early phonemic speech errors (Stoel-Gammon & Herrington 1990), and vowel articulation appears to be mostly robust to developmental changes in the vocal tract (Ménard et al. 2007, Turner et al. 2009).

Coarticulation, or the temporal and gestural overlap of two adjacent speech segments, follows a different developmental trajectory. Infants do not begin to produce C-V transitions until the onset of canonical babbling, typically between seven and ten months (Oller et al. 1997). Like vocalic development, auditory feedback—where infants compare their acoustic output to phonemic categories—is likely at play during the development of C-V transitions. However, SOMATOSENSORY FEEDBACK, or tactile consequences of speech production like the position of the tongue and lips, likely also plays a role in coarticulatory development, given the motor demands of shifting articulation from one sound to another.

After the transition from babbling to word production in toddlerhood, the majority of phonological errors are consonantal—like fricative stopping (/s/ > [t]; Chiat 1989), velar fronting (/g/ > [d]; Inkelas & Rose 2007), and cluster reduction (/pl/ > [p]; Vihman 2014)—highlighting the increased motor demands of consonants over vowels. And then, even once typically developing children shed these early phonological errors, they still do not master appropriate amounts of coarticulatory overlap until puberty (Zharkova et al. 2014). Until that time, children exhibit excessive intrasyllabic coarticulation in their speech and are able to distinguish between adjacent segments only as their fine motor control develops, the lexicon grows, and phonological awareness increases (Barbier et al. 2020, Noiray et al. 2019, Popescu & Noiray 2021, Zharkova,

Hewlett, & Hardcastle 2011). Thus, children exhibit protracted coarticulatory development until early adolescence when they have acquired both sufficient linguistic AND speech-motor experience to approximate adult-like segmental overlap.

In this work, the study of children's coarticulation differs in one additional way from the study of vowel development. Here, the degree of coarticulation in child speech is measured in two distinct environments: within morphemes and across morpheme boundaries. As such, not only does this coarticulation place more motor demands upon the children than vowels do, but it also places increased linguistic demands as the children must coordinate consonant-vowel transitions and implicate their knowledge of word forms. There is evidence that children as young as two years are capable of distinguishing between word environments in their speech. For example, children (and adults) may coarticulate more across morpheme boundaries in morphologically complex words than within phonologically equivalent, morphologically simple words (Song, Demuth, Shattuck-Hufnagel, & Ménard 2013) and lengthen fricatives in morphologically complex words (e.g. *toes*) compared to simple words (e.g. *nose*) (Song, Demuth, Evans, & Shattuck-Hufnagel 2013, but see Mousikou et al. 2021). Consequently, the question is likely not whether children can distinguish between word environments in their speech, but rather how much they can do so and as a function of what kind of language experience, if any.

Comparing vowel production to coarticulation in different word environments should be especially important in order to contrast the roles of expressive and receptive experiences in the children's language environments. Transitioning between a vowel and consonant is a highly motorized skill, refined over time with practice and increased lingual, labial, and glottal control. So if there is an effect of the ambient environment on children's coarticulation, it is reasonable to predict that it would stem primarily from children's own EXPRESSIVE experiences: children who speak more in their daily lives may have more mature coarticulation patterns. By contrast, infants formulate their early acoustic-auditory vowel categories from their ambient language, establishing these categories before they produce even their earliest vowel-like vocalizations. Infants and children then continuously update their vowel articulation throughout development to match the categories. So, we may expect especially strong effects of RECEPTIVE experience on vowel development. However, this straightforward hypothesis of expressive experiences predicting coarticulation development and receptive experiences predicting vowel development is complicated by its implication that linguistic (morphological) structure must be involved in the study of coarticulation. In that case, it is possible that receptive experience will better predict coarticulation differences by word environment because children may model the degree of their coarticulation in the two environments after the bilingual input they hear in their ambient environments. In other words, we are no longer simply investigating children's ability to differentiate between adjacent, context-free phonemes during speech production. By implicating morphology, we are probing children's experience with additional levels of language, especially the lexicon and productive morphology, which develop primarily via receptive experience.

3.3. HYPOTHESES. The goal of this article is to evaluate the role of the language-learning environment and of expressive versus receptive language experiences in children's speech development. To do so, a large-scale, naturalistic child language corpus of bilingual Quechua-Spanish speech was collected. The corpus consists of daylong audio recordings made using small, lightweight recorders that children wore over the course of an entire day. Recordings were annotated for children's and caregivers' bilingual language practices. The estimates derived from the annotations, as well as infor-

mation about the caregivers' language dominance, were then used to predict variability in the children's Quechua speech.

Two hypotheses are put forward for this work, one for each phonetic outcome, with differing predictions for expressive and receptive Quechua language experience:

- Hypothesis 1: Children who use more Quechua will have tighter, more compact vowel categories in their Quechua speech.
  - 1a. Specifically, children with more RECEPTIVE Quechua language experience— those who hear more Quechua in their everyday environments and whose caregivers reported themselves to be monolingual in Quechua or Quechua-dominant bilinguals—will have more compact vowel categories in their Quechua speech.
- Hypothesis 2: Children who use more Quechua will distinguish more between word environments in morphologically complex Quechua words.
  - 2a. Specifically, children with more EXPRESSIVE Quechua language experience— those who use more Quechua in their everyday environments—will show larger coarticulation differences between word environments in their Quechua speech.

Similar hypotheses are proposed for Spanish-dominant children, who should have more dispersed vowel categories and show smaller coarticulation differences between Quechua word environments than children who use relatively more Quechua. The hypotheses for both phonetic outcomes align with findings on bilingual exposure and dominance effects for children's language development: for bilingual children, dominance in one language predicts larger vocabularies, faster speech processing, and more accurate consonant production (e.g. Mayr et al. 2015, Place & Hoff 2011, Potter et al. 2019). It is thus reasonable to propose that bilingual dominance could also predict phonetic production outcome measures.

The hypotheses are evaluated in a bilingual community undergoing language shift from the minoritized language, Quechua, to the dominant colonizing language, Spanish. This language shift has created a gradient of child language experience within the community, as bilingual language dominance, as well as receptive and expressive experiences, varies by child.

4. METHODS.

4.1. PARTICIPANTS. Families were recruited through the researcher's personal contacts in communities surrounding a mid-size town in southern Bolivia. Participants included forty children ages four years to eight years, eleven months (twenty girls, twenty boys). Participants' families all reported speaking Quechua at home. See Table 1 for further demographic information on the participants.

AGE	N	AGE RANGE	GENDER	N CAREGIVERS W/ < 6 YRS EDUCATION
4	5	4;0–4;11	2 M; 3 F	2
5	7	5;0–5;11	2 M; 5 F	6
6	8	6;1–6;8	5 M; 3 F	6
7	14	7;1–7;11	8 M; 6 F	9
8	6	8;1–8;11	3 M; 3 F	6

TABLE 1. Demographic information for child participants.

Most children had normal speech and hearing development, per parental report. The caregivers of three children (two seven-year-olds, one five-year-old) stated that their child was late to begin talking, and another three children's caregivers did not report late-talker history. These communities are medically underserved, so some language

delays or impairments may go unreported. Three children had lost one or more of their front teeth (top or bottom) at the time of recording. (The absence of front teeth could have consequences for acoustics: for example, anterior fricatives.) An attempt was made to complete a hearing test with the children. However, it became clear after trying with a few of the children that false positives were being collected (i.e. children would fail to respond to any of the hearing test stimuli) as the children were nervous about making a mistake. Consequently, it cannot be said with absolute confidence that all children would have passed a standard hearing screening.

Thirty-seven children (92.50%) were regularly attending school at the time of data collection. The other three children were four-year-olds, as pre-kindergarten education is available but not compulsory in the communities. Most children attended school in the morning for an average of four hours (range: 3–5). Three children instead attended school in the afternoon for six hours.

Socioeconomic status (SES), usually implemented as maternal education level in developmental research, is an important predictor of child language acquisition in the United States (Hoff 2003, Pace et al. 2017). However, it is not clear that SES is predictive of language outcomes in all cultural contexts, and it is unknown whether SES predicts language outcomes in Bolivia, in these speech communities, or for children learning Quechua. SES information was nevertheless collected, as it is an important predictor of developmental outcomes in many other cultural contexts.

There were thirty unique caregivers—usually the mother but the grandmother in one family—due to eight children being sibling pairs (no twins) and one three-sibling group. The distribution of maternal education in the sample of unique caregivers was: seventeen caregivers (56.67%) had attended some primary school (less than six years of education), four (13.33%) had completed primary school (six years of education), four (13.33%) had completed the equivalent of middle school (ten years of education), one (3.33%) had completed secondary/high school (thirteen years of education), and three (10%) had not received any formal schooling. One caregiver did not report.

Caregivers' language practices. An additional indicator of SES in this community may be the central caregiver's familiarity with Spanish. This is generally correlated with the mother's education level in the speech community, since usually only women who have had the opportunity to attend school learn to speak or read Spanish. All caregivers spoke Quechua as a first language, and some additionally spoke Spanish, with varying levels of fluency. To get a description of the caregivers' Quechua-Spanish bilingual language practices, the researcher walked each primary caregiver through a brief oral survey. For the thirty unique central caregivers, the level of reported Quechua-Spanish bilingualism was: seven (23.33%) were monolingual Quechua speakers, four (13.33%) were Quechua-dominant but spoke/understood some Spanish, eighteen (60%) were bilingual Quechua-Spanish speakers, and one did not report. For the fathers, the level of Quechua-Spanish bilingualism was: one (3.33%) was a monolingual Quechua speaker, four (13.33%) were Quechua-dominant but spoke/understood some Spanish, twenty-two (73.33%) were bilingual Quechua-Spanish speakers, and three did not report.

Families were additionally asked about the language practices of the central caregivers' parents, as monolingual grandparents may enhance intergenerational minoritized language transmission. Twenty-six caregivers (86.67% of the thirty unique caregivers) reported that both of their parents spoke only Quechua, two (6.66%) reported that their father spoke some Spanish but their mother was a monolingual

Quechua speaker, and two (6.66%) reported that both of their parents spoke Spanish and Quechua.

Information on the central caregivers' code-switching habits was also collected. The central caregiver responded to the following questions on code-switching behaviors:

- Do you start sentences in Spanish and finish them in Quechua? That is to say, do you use both languages in the same sentence?
- Do you start sentences in Quechua and finish them in Spanish?
- Do you use Quechua words when you speak Spanish?
- Do you use Spanish words when you speak Quechua?

Responses to these questions are listed in Table 2. The questions were generally not applicable to the seven monolingual caregivers.

QUESTION	YES	NO	NO RESPONSE/NA
Do you start sentences in Spanish and finish them in Quechua?	17 (56.67%)	5 (16.67%)	8 (26.67%)
Do you start sentences in Quechua and finish them in Spanish?	17 (56.67%)	5 (16.67%)	8 (26.67%)
Do you use Quechua words when you speak Spanish?	18 (60.00%)	4 (13.33%)	8 (26.67%)
Do you use Spanish words when you speak Quechua?	17 (56.67%)	6 (20.00%)	7 (23.33%)

TABLE 2. Primary caregiver responses to survey questions on bilingual language practice.

4.2. PROCEDURE. There were two phases in the study: the daylong recording collection to construct the corpus and the word-elicitation tasks. For both phases, families were visited in their homes or in a central area in the community to explain the experimental procedure.

DAYLONG RECORDINGS. The daylong recordings reported here come from a larger corpus of nearly 100 infants and children acquiring Quechua and Spanish. The entire corpus is housed in the HomeBank language repository (see Cychosz 2018 for access information). This study reports on forty children from the corpus. To collect the recordings, families were given a small, lightweight recorder: either a 3"×5" Language ENvironment Analysis (LENA) Digital Language Processor (Greenwood et al. 2011), or a 2"×5" Zoom H1n Handy recorder. To explain the recording procedure to participants, the researcher demonstrated how to turn the recorder on and off and how to pause the recording, among other functions. To obtain fully informed consent for the daylong recording, the researcher explained the radius that the recorder could capture and that families had the option to delete the recording after completing it. Families were encouraged to ask questions, practice using the recorder, and make sample recordings to become familiar with the technology. Per university IRB specifications, families could pause the recording whenever they wanted, and many families elected to do so at various points. Additionally, families were instructed to either remove or pause the recorder when the child attended school and when the child was sleeping. In practice, some families forgot to turn the recorder off when the child napped, so prior to annotation an additional preprocessing step was taken to identify portions of the recording where the child might be sleeping.

Children were not required to wear the recorder to school because the children's schools were led almost entirely in Spanish, and the children spoke Spanish at school (I volunteer at a primary school in the community where I have observed many school

language practices). Since all of the school-age children spent a similar amount of time at school (see §4.1), there was no need to sample school language use. Also, there was no reliable way to obtain informed consent from everyone who might appear on the recording during the school day.

After the daylong recording procedure had been explained, families were given a small cotton t-shirt. Each t-shirt had a cotton pocket sewed to the front with a Velcro or snap-button flap to close the pocket and hold the recorder inside (Figure 1). Families were told to record for at least twelve hours, at which time they could stop. These twelve hours could be nonconsecutive, since families were allowed to pause the recording and children did not wear the recorder to school.

a. Example shirt used to house recorder during data collection.

b. The LENA recorder was worn inside of the cloth pocket for the duration of the recording.

FIGURE 1. Daylong recording collection materials.[3]

Most families completed three daylong recordings. Families were visited on three different days. At each visit, the researcher checked that the families had completed the recording and exchanged the previous shirt and recorder for a clean shirt and empty recorder. In all, thirty-nine children (97.50%) successfully completed at least three daylong recordings, while one child completed only one recording because he left on a trip after the first day.

The daylong recordings were used to estimate each child's dual language exposure. Children's language exposure was estimated in this way instead of via written background questionnaire for several reasons. First, literacy levels and familiarity with behavioral research varied greatly between participant families, including in ways that created confounds with the parameters of interest (monolingual and Quechua-dominant mothers had less opportunity to attend school). Carrying out an extensive written questionnaire, such as the 'bilingual background interview' (Marchman & Martínez-

---

[3] This figure, like others in the article, is presented in color in the electronic versions of this article, but in grayscale in the print version. Color versions of the figures are also available open access along with the supplementary materials at http://muse.jhu.edu/resolve/159.

Sussmann 2002), was thus not feasible. Second, Quechua is often undervalued and frequently stigmatized in Bolivia, so there was concern that social desirability biases could cause parents to underreport their children's Quechua language exposure. Nevertheless, it is the case that one or even many daylong recordings collected at a single timepoint cannot capture the complexities of a child's language exposure, a point to which we returned in the discussion (§6).

WORD-ELICITATION TASK STIMULI. Following the daylong recording procedure, each child completed a series of picture-prompted speech production tasks: (i) Quechua real-word repetition, including a morphological extension component, (ii) Quechua nonword repetition, (iii) Spanish nonword repetition, and (iv) additional Quechua real-word repetition with morphological extension. Nonword repetition results are not discussed in this article.

Although the children were bilingual Quechua-Spanish speakers, their phonetic development was measured in Quechua, not Spanish, for three reasons. First, Quechua's agglutinating structure permitted easier manipulation of word environment. Coarticulatory differences by word environment could also be studied in a more fusional language like Spanish, but it was anticipated that there would be more variability in coarticulatory differences by word environment in Quechua. Additionally, Quechua is a much more morphologically rich and productive language than Spanish, providing a variety of suffixes with different phonological structures for creating the stimuli. Finally, there was interest in evaluating the effect of vowel frequency on vowel variability, and Quechua, with three phonemic and two less-frequent allophonic vowels, permitted that manipulation. For reasons of time, it would not have been feasible to carry out these tasks a second time in Spanish.

The real-word repetition tasks contained fifty-six high-frequency Quechua nouns (plus six training trials) familiar to children learning Spanish and Quechua. Children's recognition of the test items was confirmed via a pre-test. Female caregivers likewise confirmed that children as young as three years of age would recognize the items. An adult female bilingual Quechua-Spanish speaker recorded the real words for the experimental stimuli, and these recordings were digitized at a sampling frequency of 44.1 kHz using a portable Zoom H1n Handy recorder. Stimuli were normed for amplitude between words but not for duration, since some words had ejectives, fricatives, and so forth that are longer. The real-word picture stimuli were color photographs of the objects.

Children in these communities can have limited exposure to technology. Consequently, rather than photos presented on a screen, picture stimuli were presented on individual pages clipped into an 11"×12.4" plastic binder. For this reason, instead of randomizing the word lists between participants, two different randomized lists were created and were counterbalanced between participants. Repetitions of the same stimulus were always separated by at least two different stimuli and were presented with a novel photo of the item each time.

A subset of the lexical stimuli are analyzed in this article: twenty-four items for the vowel analysis and forty-six for coarticulation (see Tables A1–A3 in the appendix for stimuli lists). The stimuli for the coarticulation analysis were chosen because they contained the sequence [ap] or [am] either within a morpheme (e.g. *papa* 'potato') or crossing a morpheme boundary (e.g. *thapa-pi* 'prairie-LOC') in the syllable carrying primary stress. (Quechua is generally an open-syllable language, so nearly all VC syllables cross syllable boundaries.) The only exception to the stress criterion was the word *hamˈpiri* 'healer' and its inflection *hampiˈri-pi* 'healer-LOC', where the [am] sequence did not co-

incide with primary stress. This item was included to ensure sufficient items for the within-morpheme condition while still adhering to the criteria of high frequency, easily recognizable for children, and so forth.

The VC sequences [ap] and [am] were chosen for examining coarticulatory effects for several important reasons. First, Quechua nominal case-marking suffixes are conso-nant-initial (e.g. -*q* 'genitive', -*manta* 'ablative'), so it is not possible to elicit a CV se-quence that crosses a noun–case marker boundary. Also, coarticulatory measures are highly dependent upon segmentation decisions. The acoustic delimitation between vowels and voiceless stops/vowels and nasals is relatively obvious and not subjective.[4]

The two suffixes elicited for the coarticulation analysis, the locative -*pi* and the alla-tive -*man* (pronounced [maŋ]), were chosen because inflected nouns are easier to repre-sent in photos than derived word forms are (e.g. *puñu-y* 'to sleep' → *puñu-chi-y* 'to make (one) sleep'). Also, nouns are grammatical in Quechua with just one suffix, while conjugated verbs sometimes require multiple suffixes (as seen in the previous exam-ple). Finally, absent a large, fully transcribed corpus of child-directed Quechua speech, which is not available, it is reasonable to assume that the locative -*pi* and allative -*man* on high-frequency nouns would be relatively frequent in the children's input. Thirty-five unique items were used to elicit the across-morpheme-boundary condition, and eleven unique items were used in the within-boundary condition. This represents more distinct lexical items than most previous studies of morphological effects on speech production in children or adults have used (Lee-Kim, Davidson, & Hwang 2013, Song, Demuth, Shattuck-Hufnagel, & Ménard 2013).

The stimuli used for the vowel analysis were selected because the target vowels /a, i, u/ fell in stressed and, where possible, word-medial position. Taking vowels from word-medial position avoids the effect of word-final devoicing and loss of spectral en-ergy. Additionally, words were selected to avoid flanking consonants that would exert the strongest coarticulatory effects on the vowels (glides and laterals). Finally, note that the allophonic mid vowels [e] and [o] are derived only in uvular environments (see Gal-lagher 2016 for further detail), so the flanking consonant in the words used to elicit [e] and [o] was almost always uvular.

Coarticulation between vowels and neighboring sounds is a real concern for a study of vowel variability. However, recall that all of the vowel stimuli came from the same words (thus the vowel's environment and coarticulatory influences should be relatively con-stant between children). Vowels were elicited in real words instead of nonce words be-cause the objective was to elicit QUECHUA vowels, not Spanish. But since the two languages' vowel categories completely overlap (both languages have five vowels [a, e, i, o, u], though the mid vowels are allophonic in Quechua), it could be difficult to deter-mine which language system the participants were using. If the task was to repeat con-text-neutral vowels (e.g. 'say [æ] like *cat*'), there was concern that the children would default to Spanish vowels instead of Quechua. This was especially relevant since many of the children were tested in an environment where they are used to speaking Spanish (school) by someone who looks more likely to be a Spanish speaker than a Quechua speaker (the white researcher). Because the vowels were elicited within Quechua words, there was little doubt that the children were producing Quechua vowels, not Spanish.

---

[4] Much previous work on child coarticulation has studied fricative-vowel sequences (e.g. Zharkova et al. 2011). This was not possible in the current study as there are no fricative-initial nominal case markers in Quechua.

WORD-ELICITATION TASK PROCEDURE. For the word-elicitation tasks, participants sat side by side with the experimenter. The prerecorded audio stimuli were played from an iTunes playlist run on an iPhone 6. Participants wore AKG K240 binaural studio headphones; the experimenter wore Apple earpods to follow along with the experiment. Both sets of headphones were connected to the iPhone with a Belkin headphone splitter. For each trial, the participant first heard the audio stimulus (a bare noun) while presented with the accompanying photo in the binder. The participant was asked to simply repeat the bare noun. Then, the inflected form of the target word was elicited (with locative -*pi* in the first real-word repetition task and the allative -*man* in the second) by placing a large plastic toy insect on top of the visual stimulus and asking the child 'Where is the bug?'. The child would then respond with the inflected word form (e.g. *llama-pi* (llama-LOC) 'on the llama').

The participants repeated after a model speaker, instead of spontaneously naming the item in the photo, because in a previous version of this word-elicitation task, with different children, it was found that the youngest children frequently became too nervous and hesitant to follow the task when not auditorily prompted with the word. Elicited imitation paired with a visual stimulus is also a common elicitation technique in child speech research (Erskine, Munson, & Edwards 2020, Song, Demuth, Shattuck-Hufnagel, & Ménard 2013).

There were two exceptions to participant inclusion. First, four children completed a different, pilot version of the morphological extension task; consequently, only vowel data from these children are analyzed. Second, in pilot testing, four-year-olds could not reliably complete the morphological extension task (not because they showed evidence of morphological unproductivity; they just had a harder time understanding the task). So the five four-year-olds contribute only vowel data, not coarticulation data, to the analysis.

Altogether, the word-elicitation tasks took approximately thirty to forty minutes per child. For the daylong recordings and word-elicitation tasks, each family was compensated with a small monetary sum. The families also usually kept the t-shirt, and the children could pick items from a bag of toys.

4.3. CHILD LANGUAGE CORPUS ANALYSIS.

DAYLONG RECORDING SELECTION. To get the best estimate of each child's language environment, the research team annotated the longest duration recording for each child. The average duration of the daylong recordings used for bilingual language estimation was 12.12 hours (range 7.63–16 hours), with no notable durational outliers within any age group (Table 3). Supplement 1 in the online supplementary materials includes more detailed information about corpus construction and annotation.[5]

AGE	AVG RECORDING DURATION (HRS)	RANGE (HRS)	AVG # OF POTENTIAL 30 S CLIPS TO ANNOTATE
4	12.17	8.77–16	1,400
5	12.29	8.92–16	1,454
6	12.12	7.63–16	1,369
7	12.13	9.48–16	1,469
8	11.00	11.24–13.75	1,274

TABLE 3. Daylong audio recording information by age group.

[5] The supplements referenced here and throughout are available online at http://muse.jhu.edu/resolve/159.

PROCESSING DAYLONG RECORDINGS. Researchers who collect daylong recordings rarely transcribe the recordings in their entirety, instead transcribing consciously sampled portions. Some laboratories have, for example, selected audio samples from different parts of the day within the recording (e.g. Weisleder & Fernald 2013) or samples containing large numbers of words spoken by adults (e.g. Ferjan Ramírez et al. 2019). This study instead employs a general sampling-with-replacement technique to selectively annotate portions of each daylong recording. This method has been shown to result in the most efficient, representative estimation of bilingual language exposure from daylong audio recordings (Cychosz, Villanueva, & Weisleder 2021). The entire recording selection and annotation workflow used is outlined in Figure 2.

```
┌─────────────────────────────────────┐
│      Collect daylong recording      │
└─────────────────────────────────────┘
                  │
                  ▼
┌─────────────────────────────────────┐
│  Stitch recording pieces together and│
│  intersperse with 1000 ms of white   │
│            noise                     │
└─────────────────────────────────────┘
                  │
                  ▼
┌─────────────────────────────────────┐
│  Chop each recording into 30-second  │
│            clips                     │
│ (60,308 total clips across 40 children)│
└─────────────────────────────────────┘
                  │
                  ▼
┌─────────────────────────────────────┐
│  Run vocal activity detector over    │
│  clips                               │
│ (Usoltsev 2015) and report activity  │
└─────────────────────────────────────┘
                  │
                  ▼
┌─────────────────────────────────────┐
│  Manually identify portions of       │
│  recording where child is sleeping   │
└─────────────────────────────────────┘
          │                    │
          ▼                    ▼
┌──────────────────┐  ┌──────────────────────┐
│ 0% vocal activity│  │ Clips containing vocal│
│ and sleeping clips│ │ activity randomly     │
│ not drawn for    │  │ selected, with        │
│ annotation       │  │ replacement, and      │
│                  │  │ annotated for speaker,│
│                  │  │ language, and media   │
└──────────────────┘  └──────────────────────┘
                              │
                              ▼
                    ┌──────────────────────┐
                    │ Annotate until ratio │
                    │ between language     │
                    │ categories asymptotes│
                    │ (7,413 total clips   │
                    │ annotated across 40  │
                    │ children)            │
                    └──────────────────────┘
```
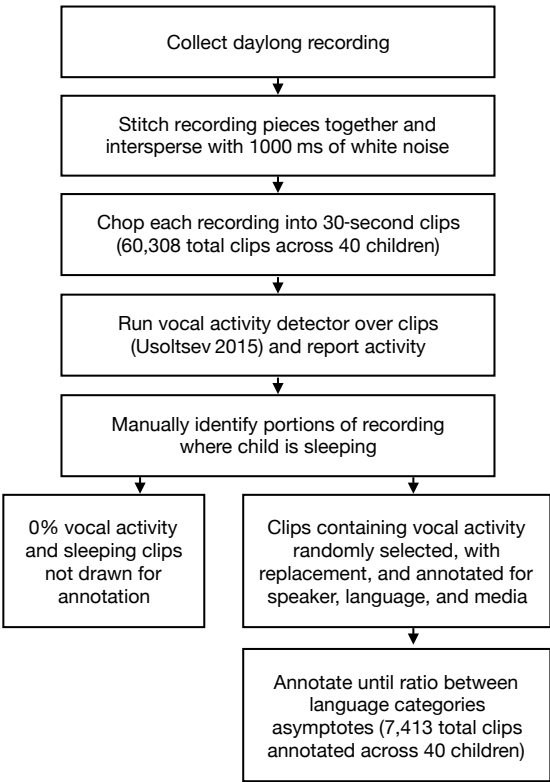
FIGURE 2. Audio clip generation, selection, and annotation workflow.

First, the recordings were chopped into thirty-second clips, as the random sampling annotation technique was validated on thirty-second clips (Cychosz, Villanueva, & Weisleder 2021). Clips were annotated using a custom graphical user interface (GUI) application that randomly selected a clip, with replacement, from a given participant's clips.[6] The researcher would listen to the drawn clip and categorize the speaker(s) and language(s) heard. Research personnel had the option to repeat the clip as many times as they would like. Clips where the child was sleeping, the researcher was present, or there was 0% reported vocal activity (determined by running a standard vocal activity

[6] The GUI application is now available open-source to use on additional annotation tasks at the project's Github repository: https://github.com/megseekosh/Categorize_app_v2.

detector; Usoltsev 2015) were drawn but not annotated. For each clip, annotators made the following decisions.

- Language?: Quechua, Spanish, mixed, no speech, personal identifying information, researcher present, or unsure
- Speaker?: target child, target child & adult, other child, other child & adult, adult, or unsure
- Media present?: yes or no

If there was no speech in the clip, annotators selected 'no speech'. If only Spanish was spoken (regardless of quantity), the researcher marked 'Spanish'. Similarly, the researcher marked 'Quechua' for monolingual Quechua clips. If the researcher heard both Quechua and Spanish in the clip—whether code-switching within a sentence or two different conversations—they marked 'mixed'. For speaker annotation, the 'target child' was the child wearing the recorder and 'other child' was anyone whose voice sounded prepubescent. Personnel were instructed to annotate 'target child and adult' if a clip contained the target child, another child, and an adult. See https://github.com/megseekosh /Categorize_app_v2/blob/master/FAQs_bilingual.MD for further details on annotation decisions, including a list of frequently asked questions used to standardize annotation between research personnel.

When the language or speaker in a clip was unclear (e.g. a caregiver singing nonce words, other nonlanguage vocalizations), research personnel could select 'unsure'. Language and speaker were coded separately, so annotators could still code for speaker or language even if the other category was unclear. The 'unsure' annotation was most often used for clips where a conversation was taking place in the background of the recording that made it difficult to determine the language and/or speaker. The team of annotators considered the possibility that it may be difficult to ascertain the speaker or language in some clips because those clips are noisier and contain multiple interlocutors. These noisy clips, with multiple interlocutors, might be more likely to contain mixed speech, so disregarding them could lead the team to inadvertently disregard clips of a certain category (i.e. mixed speech). In practice, however, the 'unsure' clips almost always contained background speech without a discernible speaker or language, so the team felt confident in excluding 'unsure' clips from further analysis.

The choice for media was binary—'present' or 'absent'—because (i) it was often difficult to determine if the media in the recording was radio or TV and (ii) almost all of the media was in Spanish, making it irrelevant to mark the language. In other words, when media was present, it was in Spanish.

As annotators drew and listened to the thirty-second clips, they were simultaneously running a Jupyter notebook to mark progress toward annotation. The notebook recorded the proportion of Quechua, Spanish, and mixed clips to total clips for each child. Human annotation was cut off when two criteria were met. First, the proportion and variance (variance measured over a moving window of sixty language-proportion estimates) between language categories had to asymptote (that is, approach but not touch a horizontal line, exemplified in Figures 3 and 4). Second, fifty language clips from each child had to be annotated (language clips include those annotated as Quechua, Spanish, or mixed, but not as 'unsure' or 'no speech'). The fifty-clip criterion was included as an additional precautionary measure to ensure sufficient transcription even if stability between language-category proportion and variance was reached. Given these predetermined criteria, the team was more confident that their annotations were accurately reflecting the child's language environment.
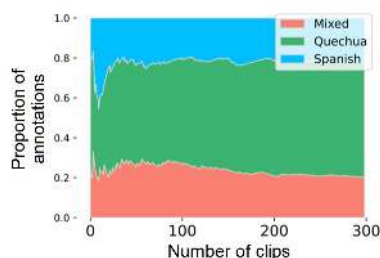
FIGURE 3. Example area plot of language proportions by number of clips annotated. Area plots were used to track progress toward language proportion stability during daylong recording annotation.
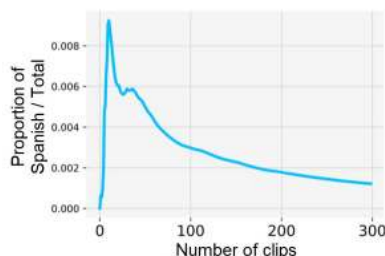

FIGURE 4. Example plot of Spanish language proportion variance by number of clips annotated. Variance was computed over a moving window of sixty clips. This plot was used to track progress toward variance stability during daylong recording annotation.

The research team was able to make stable estimates of each child's bilingual language exposure by listening to and annotating an average of 185.3 thirty-second clips from a given recording (SD = 69.72, range = 84–385), or an average of 92.66 minutes total from each recording.[7] Given that recording length varied (Table 3), the annotated clips made up an average of 13.13% of each recording (SD = 5.47, range = 4.38–29.90%). Thus, the number of clips annotated for a given child varied as a function of the unpredictability of language categories in the child's environment. But the criterion for variance between the annotated categories was the same for all children. Overall, this procedure resulted in the annotation of a total of 3,706.5 minutes, or 61.78 hours, across the forty children.

RESEARCH PERSONNEL. Three undergraduate student research assistants and the lead researcher (the author) annotated the daylong recordings. All research assistants were fluent Spanish speakers participating in a linguistics research training program. The annotation personnel underwent a stringent training procedure prior to and during annotation (see Supplement 1 for details).

Interrater reliability scores between the lead researcher and all personnel members were calculated to ensure fidelity to the coding scheme. Seventy-two clips were randomly selected from one participant's recording.[8] Each personnel member then anno-

---

[7] The figures reported in the text reflect those clips that annotators ACTUALLY listened to. A grand total of 8,974 clips were drawn, including those that were not listened to because they did not have any vocal activity, the child was sleeping, or the researcher was present. This amounted to an average of 224.35 thirty-second clips from each recording (SD = 107.06, range = 91–618), or an average of 112.15 minutes total from each recording. The number of clips that were listened to are reported in the text because those figures more accurately reflect the time commitment required for annotation.

[8] Due to a bug in the annotation script, seventy-two clips were annotated by all four team members, though seventy-five clips were originally selected.

tated the clips according to the established annotation procedure. The interrater reliability between personnel members (lead researcher and three assistants) and the remaining team was as follows: 94.44% agreement (lead researcher), 93.06%, 94.44%, and 98.61% agreement for each of the three assistants (Krippendorff's alpha = 0.87 for the entire team).

Intrarater reliability was also collected for all annotation personnel: the lead researcher had 99.17% intrarater agreement (Krippendorff's alpha = 0.99), research assistant 1 had 97.62% agreement (Krippendorff's alpha = 0.93), research assistant 2 had 99.29% (Krippendorff's alpha = 0.93), and research assistant 3 had 100% (Krippendorff's alpha = 1.0). In all, these inter- and intrarater agreement scores were satisfactory to conclude that raters were calibrated and annotating uniformly.

4.4. ACOUSTIC ANALYSIS. Participants' audio files from the word-elicitation tasks were first manually aligned to the word level in Praat (Boersma & Weenink 2020) and to the phone level using a Quechua forced aligner trained on the participants' data (McAuliffe et al. 2017). The phone-level alignment was hand-corrected by one of two trained phoneticians. Alignment was conducted auditorily and by reviewing the acoustic waveform and broadband spectrogram. These acoustic analyses are sensitive to alignment decisions, so a number of parameters were set prior to alignment. Word-initial plosive, affricate, and ejective onset corresponded to burst onset. Onset of periodicity and formant structure in the waveform and spectrogram marked vowel onset. Nasals were identified by anti-formants and dampened amplitude. Glide-vowel sequences were delimited visually, or, when this was not possible, half of the vowel-glide sequence was attributed to the vowel and half to the glide. There is some variability in the realization of mid vowels in Quechua speakers; vowels were transcribed phonemically.

Interrater agreement between the phoneticians aligning the files was evaluated. Both phoneticians aligned two randomly selected recordings, one from a child aged five years, nine months and another from a child aged seven years, four months. The difference between the aligners' average consonant duration was 4 ms, and vowel duration was 2 ms for the first child. Pearson correlations between the aligners for this child were significant for consonants: $r = 0.86$, $p < 0.001$, 95% CI [0.83, 0.89], and for vowels: $r = 0.94$, $p < 0.001$, 95% CI [0.93, 0.96]. For the second child, the difference between aligners' average consonant duration was 2 ms, and vowel duration was 2 ms. Pearson correlations between the aligners were significant for consonants: $r = 0.98$, $p < 0.001$, 95% CI [0.97, 0.98] and for vowels: $r = 0.95$, $p < 0.001$, 95% CI [0.94, 0.96], suggesting fidelity to the alignment protocol.

COARTICULATION ANALYSIS. The high frequencies and breathiness of child speech can make it difficult to implement some traditional acoustic measures of coarticulation such as peak equivalent rectangular bandwidth (ERB_N) (Reidy et al. 2017), center of gravity, or formant transitions and/or spectral peaks (Lehiste & Shockey 1972, Öhman 1966). To circumvent these issues, this study measures coarticulation as the spectral distance between two phones, a technique that has been validated for children's speech and a variety of consonants (Cychosz et al. 2019, Gerosa et al. 2006). For this measure, a custom Python script running librosa functions (McFee et al. 2015) (available in the project's Github repository) computed the mel-frequency log spectra over the middle third of two adjacent phones (e.g. [a] and [p]) that fell within morpheme boundaries (e.g. *api* 'corn/citrus drink') and across morpheme boundaries (e.g. *llama-pi* 'llama-LOC'). Then, the average spectrum for each phone was computed, and the Euclidean distance between those averages was measured. Finally, to compute the coarticulatory

differences by word environment, the difference in coarticulation between the across-morpheme environment and the within-morpheme environment was calculated. This computed difference is a measure of how different the degree of coarticulation is, so a larger difference indicates that a speaker differentiates more between the two word environments.

Vowel analysis. Because vowel formant frequencies can be difficult to track reliably in children's speech, a triple formant tracker running three trackers (inverse filter-control (Watanabe 2001), Entropic Signal Processing System's (ESPS) 'covariance', and ESPS's 'autocorrelation') was built.[9] The tracker produced three measurements (one from each tracker) for the first two formant frequencies (F1 and F2) at the vowel midpoint. The median formant measurement was then used in the analysis. In this way, anomalous measurements from any single tracker did not have an outsize influence. Formant frequencies were Lobanov-normalized to control for between-child anatomical differences (Lobanov 1971). Additional steps taken to clean and standardize the formant measures are outlined in Supplement 2.

Vowel dispersion was implemented as the average Euclidean distance in F2/F1 space from the vowel category mean, resulting in a single coefficient per vowel category.[10] This category dispersion coefficient reflects both the mean value of each vowel category and its variability along the F1 and F2 dimensions. A larger dispersion coefficient indicates that the acoustic vowel category is more disperse.

5. Results. The first section of the results (§5.1) presents descriptive analyses of the proportion of Quechua, Spanish, and mixed Quechua-Spanish speech clips (henceforth 'mixed') in the daylong recordings. These analyses quantify the variation in bilingual language exposure between children, as well as how this exposure varies by child age and the primary caregiver's language dominance. The second section (§5.2) examines how individual differences in language exposure predict the phonetic outcomes measured. It is expected that children with more Quechua-dominant caregivers—who hear more Quechua—will have tighter vowel categories (a smaller vowel dispersion coefficient). Likewise, it is expected that children who use more Quechua will be more likely to distinguish via coarticulation between word environments in their speech production.

All analyses were conducted in the RStudio computing environment (version: 1.2.5033; RStudio Team 2020). Data visualizations were created with ggplot2 (Wickham 2016). Modeling was conducted using the lme4 (Bates et al. 2015) and lmerTest (Kuznetsova, Brockhoff, & Christensen 2017) packages, and summaries were presented with Stargazer (Hlavac 2018). The significance of potential model parameters was determined using a combination of log-likelihood comparisons between models, Akaike information criterion (AIC) estimations, and *p*-values procured from model summaries. Scripts to replicate these results are available in the project's Github repository.

5.1. Descriptive analyses of bilingual language exposure. Figure 5 shows the distribution of the language categories Quechua, Spanish, and mixed, as well as clips annotated 'unsure' or 'no speech', by the central caregiver's language profile (a table

---

[9] The tracker is now available open-source for the broader speech research community to use at https://github.com/megseekosh/vocal_tract_vowel.

[10] Previous work has used the coefficient of variation (CoV) to measure vowel category dispersion in children (e.g. Lee et al. 1999). The CoV is the ratio of the standard deviation of the mean to the mean of each phoneme category (Bradlow 1995, Eguchi & Hirsh 1969, Lee et al. 1999). One disadvantage of the CoV is that, unlike the category dispersion technique, separate coefficients must be computed for each acoustic dimension, so the result is the CoV of F1, CoV of F2, and so forth.

listing the number of clips in each language by maternal profile is included in Supplement 3). Henceforth the central caregiver is referred to as the mother, though one of the caregivers was the child's grandmother. The maternal language profile was determined from the brief background survey completed during testing. Three maternal language profiles were compared: mothers who were monolingual Quechua speakers ($n = 10$), Quechua-dominant speakers ($n = 6$), and bilingual Quechua-Spanish speakers ($n = 23$). One family did not report the mother's bilingual language profile. Unsurprisingly, a larger number of monolingual Quechua clips were found in the recordings of the children with monolingual mothers ($n = 381$, 18.07%) than of the children with Quechua-dominant mothers ($n = 163$, 11.71%) or bilingual Quechua-Spanish mothers ($n = 353$, 9.29%). Children with monolingual mothers are exposed to more Quechua than children with Quechua-dominant or bilingual mothers. The percentage of Spanish clips did not vary greatly by language profile. See appendix Table A4 and Figure A1 for a distribution of clip annotation categories and percentages for each individual child.[11]
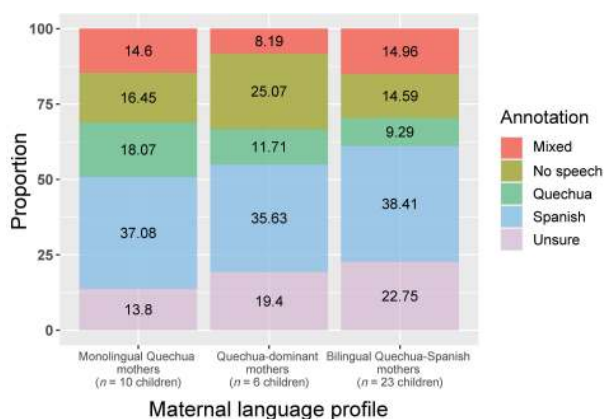

FIGURE 5. Proportion of language categories, by maternal language profile. Numbers on barplot reflect percentages of each category. Note: one family did not report maternal language profile.

Figure 6 presents the distribution of clip annotations by child age (four to eight years; corresponding table is given in Supplement 3). Given the ongoing language shift in this community, we might anticipate more Spanish or mixed speech among the youngest children. Instead, the seven- and eight-year-olds have the largest percentage of Spanish clips, suggesting that factors other than age may predict the proportion of each language in the input. The data also suggest that language environments become more verbal as children age, since far fewer clips in the seven- and eight-year-old groups contained no speech.

While there are too few children within each age group to reliably evaluate the effects of maternal language profile by age, a larger proportion of the seven-year-old group had bilingual Quechua-Spanish mothers (Table 4): nine of thirteen seven-year-olds had bilingual mothers compared to three of seven five-year-olds, for example. Nevertheless, maternal language profile cannot entirely explain the differences between age groups: four of six eight-year-olds had monolingual mothers, but a large percentage of their clips (39.3%) were still Spanish.

---

[11] Clips containing personal identifying information accounted for between 0 and 0.04% of the total clips from each recording and are not further reported in the analyses.
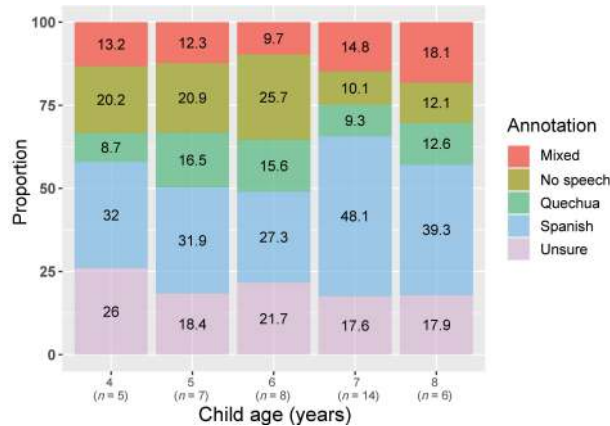
FIGURE 6. Proportion of language categories, by child age (years). Numbers on barplot reflect percentages of each category.

AGE	MONOLINGUAL QUECHUA	QUECHUA-DOMINANT	BILINGUAL QUECHUA-SPANISH	TOTAL
4	0	1	4	5
5	2	2	3	7
6	2	1	5	8
7	2	2	9	13
8	4	0	2	6

TABLE 4. Maternal language profiles by child age (in years).

5.2. CORRELATING LANGUAGE DOMINANCE AND SPEECH PRODUCTION. Five parameters, reflecting the child's expressive and receptive language experiences in Quechua and Spanish, were correlated with the children's phonetic outcomes. Maternal language profile modeled the mothers' language dominance from the background survey conducted during testing. The four remaining parameters were calculated from the annotated corpus. See Table 5.

VOWEL CATEGORY DISPERSION. The first parameter studied is the effect of maternal language profile on the children's three phonemic /a, i, u/ and two allophonic [e, o] Quechua vowels (Figure 7; see Supplement 4 for individual vowel plots). Maternal language profile predicted the children's vowel dispersion. Children with monolingual mothers appear to have tighter, less variable Quechua vowel categories: their average vowel category dispersion coefficient was consistently smaller than the average coefficient of the children with Quechua-dominant or bilingual mothers, especially for the peripheral, phonemic /a, i, u/ (summary statistics in Table 6). The only exception was for [o], which had an average category dispersion coefficient of 0.60 ($SD = 0.39$) for the children with Quechua-dominant mothers and 0.63 ($SD = 0.37$) for the children with monolingual Quechua mothers. The standard deviation of the coefficients also tended to be smaller for the children with monolingual mothers, suggesting that, as a group, they had more uniform vowel variability.

There were also differences between the children with Quechua-dominant mothers and those with bilingual mothers. Children with bilingual mothers had more expansive [e] and [i] categories (larger coefficient), while [a] dispersion was larger for the Quechua-dominant group. When considering differences between language profile groups, it is important to note that the group with bilingual mothers had almost four times as many

PARAMETER	TYPE	DESCRIPTION
Maternal language profile	Receptive	Three levels: monolingual Quechua, Quechua-dominant, and bilingual Quechua-Spanish
% of child's monolingual Spanish clips	Expressive	Calculated by dividing the number of monolingual Spanish clips where the target child was speaking by the total number of language clips where the child was speaking*
% of child's mixed language clips	Expressive	Calculated by adding the total number of Quechua clips and mixed clips where the target child was speaking and then dividing it by the total number of clips where the child was speaking[†]
% of other speakers' monolingual Spanish clips	Receptive	The percentage of monolingual Spanish clips in the recording where an adult or other child—not the target child—was speaking; calculated by dividing the number of Spanish clips where an adult or other child (i.e. sibling) was speaking by the total number of language clips containing an adult or other child
% of other speakers' monolingual Quechua clips	Receptive	Calculated by dividing the number of monolingual Quechua clips where an adult or other child was speaking by the total number of language clips containing an adult or other child

TABLE 5. Language exposure parameters used in modeling. Notes: * Language clips were defined as those annotated as Spanish, Quechua, or mixed, and not clips annotated as unsure, no speech, or containing personal identifying information. [†] Both monolingual Quechua clips and mixed Quechua/Spanish clips were used, instead of just monolingual Quechua clips, because some children had very few monolingual Quechua clips in which they were speaking. Figures representing the relationship between both speech production outcomes and just the percentage of monolingual Quechua clips, not including mixed clips, are included in Supplement 3. Similar results were found across the samples.
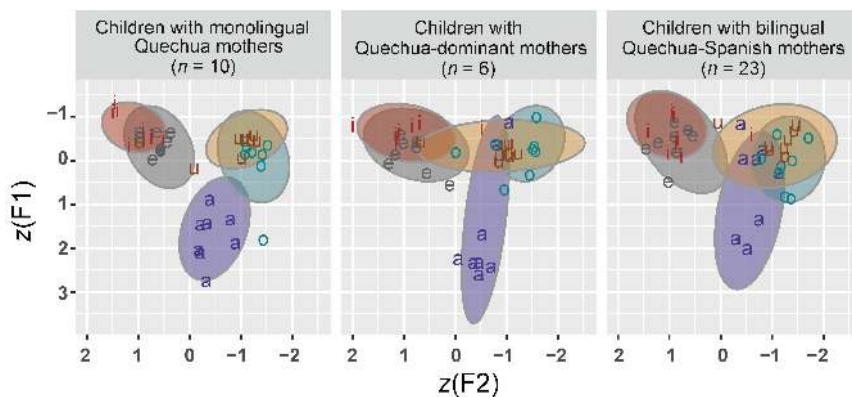
FIGURE 7. Children's vowel spaces by maternal language profile. Ellipses represent 95% CIs, or approximately 2 *SD*s of all data, assuming a normal *t*-distribution. Individual points represent a random subset of eight tokens per vowel category.

	[a]	[e]	[i]	[o]	[u]
Monolingual Quechua	0.69 (0.47)	0.61 (0.36)	0.44 (0.27)	0.63 (0.37)	0.58 (0.41)
Quechua-dominant	1.03 (0.59)	0.70 (0.30)	0.60 (0.31)	0.60 (0.39)	0.85 (0.55)
Bilingual Quechua-Spanish	0.91 (0.57)	0.78 (0.51)	0.63 (0.50)	0.70 (0.40)	0.84 (0.60)

TABLE 6. Average and standard deviation (in parentheses) of vowel category dispersion by phone and maternal language profile.

children (*n* = 23) as the group with Quechua-dominant mothers (*n* = 6), and more than twice as many children as the group with monolingual Quechua mothers (*n* = 10). Still, the differences in vowel patterning suggest that mothers' language dominance affects the children's vowel dispersion.

As the vowel plots in Fig. 7 show, there was considerable overlap between the allophonic vowels [e] and [o] and their underlying phonemic forms /i/ and /u/, respectively. This overlap does not appear to qualitatively differ by maternal language profile, even though one could expect children with bilingual mothers to have more distinct [e] and [o] categories—those vowels are phonemic in Spanish, which may be the children's dominant language. Instead, there are similar amounts of overlap between the allophonic and phonemic vowels across the three maternal language profiles, suggesting that all children employed a Quechua vowel system during the task.

To further evaluate the effect of maternal language profile on the children's vowel variability, a series of linear mixed-effects models were fit. The dependent variable was the dispersion coefficient of each child's vowel category or, ideally, five coefficient estimations (one for each vowel) per child, though some vowel categories were removed due to a low number of observations (see tables in Supplement 2 for removal statistics, as well as the number of tokens by vowel category and maternal language profile used in modeling).

First, a baseline model with a random effect of SPEAKER was fit to predict the dispersion coefficients. The parameter PHONE was then added, which unsurprisingly improved model fit. Next, the parameter MATERNAL LANGUAGE PROFILE, with the levels Monolingual Quechua, Quechua-dominant, and Bilingual Quechua-Spanish, was added. Maternal language profile significantly improved upon a model fit containing Phone and the random effect of Speaker, under an alpha level of 0.10 (model summary presented in Table 7). There was no effect of CHILD AGE on vowel dispersion. This modeling shows a trend that the mother's language profile predicts vowel variability in these bilingual children. More specifically, the positive coefficients for the levels of Maternal language profile, with a reference level of Monolingual Quechua, show a trend that children with Quechua-dominant and bilingual Quechua-Spanish caregivers have more variable vowels.

	ESTIMATE β [CI]	t	p	
(intercept)	0.63 [0.47, 0.79]	7.88	< 0.001	***
Phone: [e]	−0.25 [−0.40, −0.09]	−3.16	0.002	**
Phone: [i]	−0.29 [−0.43, −0.15]	−4.10	< 0.001	***
Phone: [o]	−0.35 [−0.52, −0.18]	−4.14	< 0.001	***
Phone: [u]	−0.14 [−0.29, 0.02]	−1.68	0.10	+
Lang. profile: Quechua-dominant	0.20 [−0.01, 0.41]	1.86	0.07	+
Lang. profile: Bilingual	0.15 [−0.01, 0.30]	1.87	0.07	+

TABLE 7. Model predicting vowel category variability. Note: $+ p < 0.1$, $** p < 0.01$, $*** p < 0.001$.

After the effects of maternal language profile on vowel category dispersion were evaluated, the effects of the ambient language measures derived from the daylong audio recordings were correlated with vowel dispersion. These environmental predictors are necessarily correlated. That is, the more monolingual Spanish clips where the child is speaking there are in a recording, the fewer Quechua/mixed clips containing the child

there are likely to be. A Pearson correlation coefficient assessing the relationship between the percentage of Spanish clips and the percentage of Quechua/mixed clips containing the target child in the recordings demonstrates that these two predictors are indeed significantly negatively correlated ($r(37) = -0.46$, $p = 0.004$). The percentage of Spanish clips and percentage of Quechua clips containing adults and other children are also significantly negatively correlated ($r(37) = -0.93$, $p < 0.001$).

According to the experimental hypotheses, children who are exposed to and use more Spanish, at the expense of Quechua, should have more variable Quechua categories. There is some limited evidence for this idea, at least for /a/: children who have a smaller percentage of Spanish clips in their recording have less variable [a] categories (Figure 8). However, this relationship is not consistent across all vowel categories, so vowel variability does not seem to vary by the percentage of Quechua or Spanish clips in the children's recordings. (See Supplement 3 for a plot of vowel dispersion by the percentage of monolingual Quechua clips.)
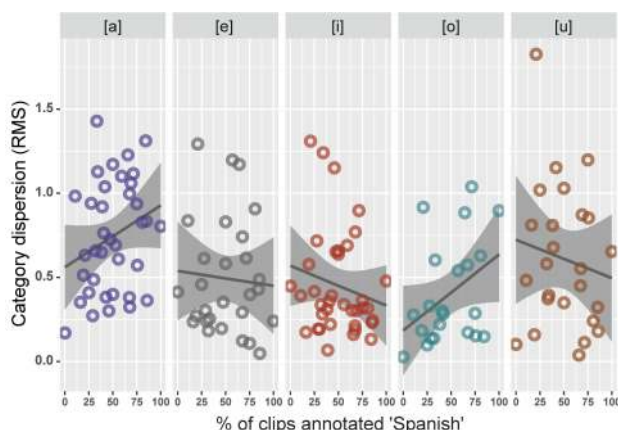

FIGURE 8. Vowel category dispersion by percentage of Spanish clips containing target child. Each point represents one child. Ribbons represent 95% confidence intervals.

A series of linear mixed-effects models were fit to evaluate whether the parameters from the daylong recordings predicted children's vowel variability. As with the modeling for mother's bilingual language profile, the outcome variable was the vowel dispersion coefficient, so five categories per speaker. None of the parameters estimated from the recordings, PERCENTAGE OF SPANISH CLIPS CONTAINING TARGET CHILD, PERCENTAGE OF QUECHUA/MIXED CLIPS CONTAINING TARGET CHILD, PERCENTAGE OF SPANISH CLIPS CONTAINING ADULT/OTHER CHILD, or PERCENTAGE OF QUECHUA CLIPS CONTAINING ADULT/OTHER CHILD, improved upon a model fit that included a random effect of Speaker and fixed effect of Phone, suggesting that there was no relationship between environmental effects estimated from the recordings and vowel category variability.

To conclude, the observed relationship between category variability and maternal language profile suggests that a receptive measure, the caregivers' bilingual language practices, may influence children's spoken vowel patterning, but only when implemented as the discrete maternal language profile. There was little or no relationship between receptive Quechua language exposure or children's expressive Quechua experience, garnered from the daylong recordings, and Quechua vowel variability.

Speech production by word environment. The final section of the results evaluates the relationship between the children's dual language environments and how they distinguish between word environments via coarticulation. Recall that the outcome measure is the difference in the amount of coarticulation between the across-morpheme environment and the within-morpheme environment. A larger difference between the two environments indicates that the child differentiates between them more. This outcome is referred to as the coarticulation difference. The overall hypothesis is that children who hear, but especially use, more Quechua will be more likely to differentiate their coarticulation patterns by word environment.

Figure 9 presents the analysis of the coarticulation difference by the discrete receptive parameter Maternal language profile. The coarticulation difference does not appear to vary by maternal language profile.
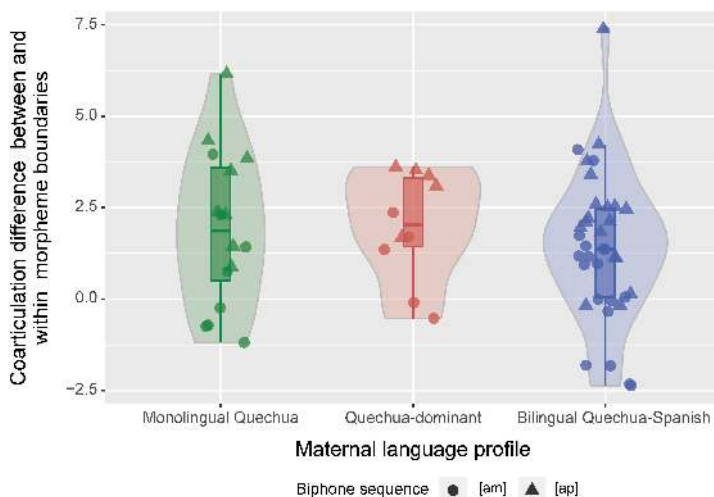

Figure 9. Coarticulation difference by maternal language profile and biphone sequence: there was no reliable effect of profile on coarticulation. Circles and triangles represent each child's coarticulatory difference across -*man* 'allative' and -*pi* 'locative' morpheme boundaries, respectively. Boxplot hinges represent the interquartile range. Points are jittered horizontally to avoid overlap.

Next, the bilingual language characteristics from the annotated corpus were evaluated against the children's correlation difference. Children who spoke more Spanish exhibited a smaller coarticulation difference: they tended to distinguish less between across-morpheme and within-morpheme word environments (Figure 10). A similar result was found for the amount of receptive experience with Spanish: children who heard more Spanish from siblings, caregivers, and other speakers in the environment also tended to distinguish less between word environments (Figure 11).

To model these expressive and receptive experiences in the children's environments, a series of linear mixed-effects models were fit to predict coarticulation difference. The baseline model included a random effect for Speaker and a fixed effect of biphone sequence. Each environmental effect was added separately to the model, given the strong correlations between variables (see §5.2).

The parameter Percentage of Spanish clips containing target child improved upon model fit (model summary in Table 8). The negative coefficient for this parameter suggests that the more Spanish clips containing the child in the recording, the smaller the coarticulation difference, and the less the child was distinguishing between the two word
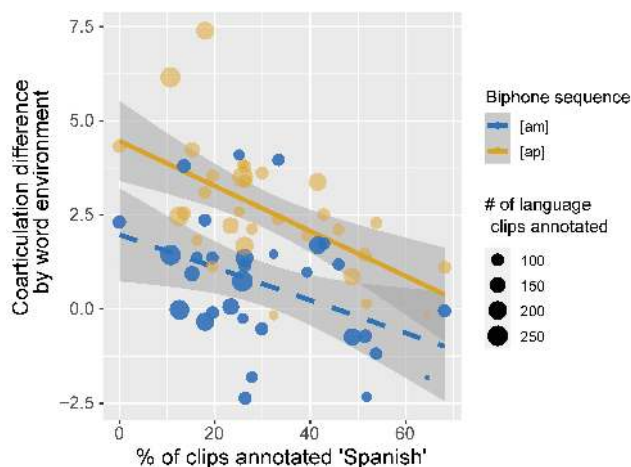
FIGURE 10. Coarticulation difference by percentage of Spanish clips containing the target child. Data points represent each child's average coarticulatory difference by word environment for *-man* 'allative' (blue; dashed line) and *-pi* 'locative' (yellow; solid line). Ribbons represent 95% confidence intervals.
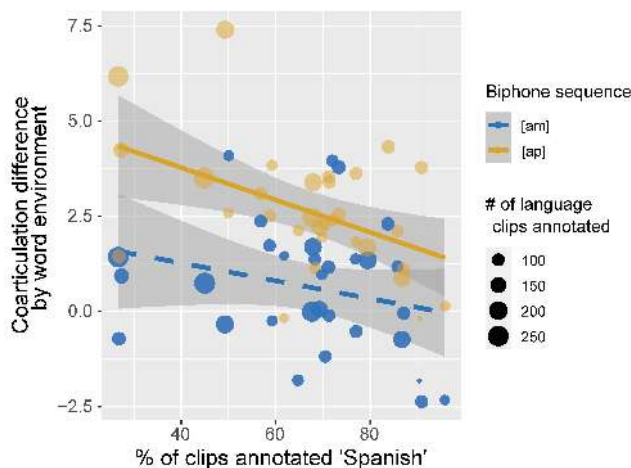
FIGURE 11. Coarticulation difference by percentage of Spanish clips containing an adult or nontarget child. Data points represent each child's average coarticulatory difference by word environment for *-man* 'allative' (blue; dashed line) and *-pi* 'locative' (yellow; solid line). Ribbons represent 95% confidence intervals.

environments. The environmental effect Percentage of Quechua/mixed clips containing target child also improved upon the baseline model fit, with a positive coefficient showing that the more Quechua/mixed speech the child uses, the larger the coarticulation difference between the two word environments.

For the measures of receptive experience, Maternal language profile, representing the bilingual language profile of the mother (monolingual Quechua, Quechua-dominant, or bilingual Quechua-Spanish), did not improve upon model fit (Table 9). Percentage of Spanish clips containing adult/other child and Percentage of Quechua clips containing adult/other child both improved (separately) upon the baseline fit. However, the AIC values were reliably lower for the models containing the expressive parameters (percentage of the target child's Spanish and Quechua/mixed speech use), indicating

	SPANISH MODEL				QUECHUA MODEL		
	ESTIMATE β [CI]	t	p		ESTIMATE β [CI]	t	p
(intercept)	2.22 [1.37, 3.08]	5.08	< 0.001	***	0.12 [−0.54, 0.78]	0.36	0.72
Biphone sequence: [ap]	1.99 [1.26, 2.71]	5.38	< 0.001	***	1.99 [1.22, 2.76]	5.06	< 0.001	***
Percentage of Spanish clips	−0.05 [−0.07, −0.03]	−4.51	< 0.001	***			
Percentage of Quechua /mixed clips					0.11 [0.03, 0.19]	2.83	0.005	**
Log likelihood	−114.05				−117.75		
Akaike inf. crit.	238.11				245.50		
Bayesian inf. crit.	248.74				256.13		

TABLE 8. Models predicting coarticulation difference by expressive language measures
(containing target child). Note: ** $p < 0.01$, *** $p < 0.001$.

that expressive language use fit the data better and might mediate the effect of receptive language on the coarticulation outcome; this hypothesis was subsequently evaluated in a mediation model. There was no effect of Child age on the differentiation by morphological environment.

	MODEL 1				MODEL 2		
	ESTIMATE β [CI]	t	p		ESTIMATE β [CI]	t	p
(intercept)	0.91 [−0.06, 1.87]	1.84	0.07		2.84 [1.26, 4.43]	3.51	< 0.001	***
Biphone sequence: [ap]	1.99 [1.22, 2.76]	5.06	< 0.001	***	1.99 [1.22, 2.76]	5.06	< 0.001	***
Mat. lang. profile: Quechua-dominant	0.11 [−1.32, 1.54]	0.15	0.89				
Mat. lang. profile: Bilingual	−0.50 [−1.56, 0.57]	−0.91	0.37				
Percentage of Spanish clips					−0.03 [−0.06, −0.01]	−2.91	0.004	**

	MODEL 3			
	ESTIMATE β [CI]	t	p	
(intercept)	−0.01 [−0.78, 0.76]	−0.04	0.98	
Biphone sequence: [ap]	1.99 [1.22, 2.76]	5.06	< 0.001	***
Percentage of Quechua/ mixed clips	0.03 [0.01, 0.06]	2.43	0.02	*

	MODEL 1	MODEL 2	MODEL 3
Log likelihood	−117.64	−118.81	−119.66
Akaike inf. crit.	247.28	247.61	249.32
Bayesian inf. crit.	260.05	258.25	259.96

TABLE 9. Models predicting coarticulation difference by receptive language measures.
Note: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

To determine whether the receptive or expressive measures derived from the corpus resulted in a better model fit, a mediation model was fit to predict the children's coarticulation difference (Figure 12). A basic linear model showed that receptive Spanish experience positively predicted children's own Spanish language use ($\beta = 0.37$, $p = 0.009$). When simultaneously modeling the effects of receptive and expressive experiences on

children's coarticulation in the same mixed-effects model as above, the effect of receptive experience (Spanish language input) was weaker and less reliable (β = −0.02, $t = −1.9$, $p = 0.07$; dotted line in mediation plot). Thus, expressive language experience directly predicts children's coarticulation difference by word environment. The effect of Spanish input from others is indirect: children who hear more Spanish use more Spanish and, in turn, distinguish less between Quechua word environments (and vice versa for Quechua). The modeling thus suggests that to develop robust acoustic representations of morpheme boundaries—especially coarticulatory, although the result was replicated for duration as well—it is insufficient to hear Quechua in the ambient environment. Children appear to need the tactile and somatosensory feedback experience associated with speaking Quechua to differentiate coarticulatorily between word environments.
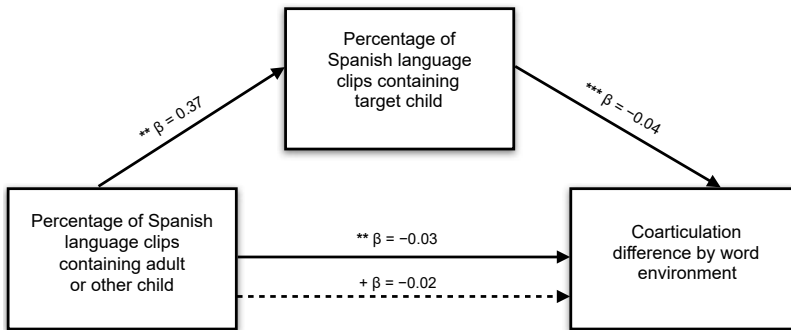

FIGURE 12. Mediation model predicting children's coarticulation difference between word environments. Solid lines denote direct relationships and the dotted line an indirect relationship. The effect of Spanish language input on coarticulation is mediated by the children's own Spanish language use.

As a final step in the model-fitting procedure, a parameter estimating each child's 'talkativeness' was added to the models. This parameter, PERCENTAGE OF CHILD SPEAKING, was included to control for the possibility that children who simply talk more—in whatever language—may exhibit different speech production patterns by word environment. Percentage of child speaking was calculated by dividing the number of clips where the target child was speaking by the total number of language clips, irrespective of speaker. The parameter Percentage of child speaking did not improve upon a model containing Biphone sequence, Percentage of Spanish clips containing target child or Percentage of Quechua/mixed clips containing target child, and the random effect of Speaker. Furthermore, neither the magnitude nor direction of the effect of Percentage of Spanish clips/Percentage of Quechua clips changed with the addition of Percentage of child speaking. Though the Percentage of child speaking is just an estimation of the children's talkativeness, these results suggest that it is the effect of speaking a particular language that predicts the children's speech patterns, and not simply how frequently they talk.

Overall, this modeling demonstrates that the proportion of children's Spanish and Quechua/mixed language use predicted their coarticulation differences between word environments. Although some of the children's receptive experiences—the proportion of Spanish and Quechua that other speakers in their environment used—likewise predicted their coarticulation by word environment, the children's own language use was the stronger predictor. The receptive parameter Maternal language profile did not predict the children's coarticulation patterns. Unlike vowel variability then, coarticulation patterns by word environment are best predicted by the children's own speech practices.

6. Discussion. This work asked if and how children's language environments predict their speech-language development by studying phonetic development in a bilingual community undergoing language shift. A large-scale child language corpus of over 500 hours of naturalistic Quechua-Spanish speech was efficiently annotated to estimate children's dual language exposure. These dual language estimates, and the central caregiver's language dominance, were then used to predict the children's patterns on a series of speech-production tasks.

Exposure did predict the children's patterning, with different results depending on the phonetic outcome measure. Vowel dispersion was most contingent upon the mother's language dominance, a receptive measure, and not the children's own language production. Children's coarticulation by word environment, however, depended most on the proportion of each language used—a measure of expressive language experience. Thus, this showed that in addition to predicting developmental outcomes like babbling and lexical growth, children's language environments also impact their phonetic development. The language environment is also not a static construct. Different components like receptive versus expressive speech interact, highlighting the need to operationalize the environment in different ways before concluding or ruling out environmental influences.

6.1. Receptive versus expressive experiences. Current theoretical debates in child speech-language acquisition concern the EXTENT and MECHANISMS of environmental influence on children's development. This work contributes to both lines of inquiry, because children in the community studied have a wide range of bilingual, and receptive and expressive, language experiences. Concerning the extent of environmental influence, children's own language use was the best predictor of coarticulation patterns by word environment, but it did not predict vowel dispersion; caregiver language use best predicted vowel dispersion. One takeaway is thus that the impact of caregiver speech models depends on the outcome measure, at least for this age range, even for acoustic-phonetic measures that have as much in common as vowel dispersion and coarticulation (versus outcomes that are more distinct, like vowel dispersion and lexical growth). More critically, this result demonstrates the importance of distinguishing between different types of environmental experiences (receptive versus expressive), especially for outcomes in phonetics and phonology that draw heavily on speech-motor coordination. Finding a null effect of one type of exposure metric, like quantity of adult language input, cannot preclude all exposure-related effects.

Concerning the mechanisms of environmental influence, results from the two phonetic outcomes are informative. Why do we see a mild effect of receptive experience on vowels but a stronger effect of expressive experience on coarticulation by word environment? The hypotheses laid out predictions for the effect of receptive and expressive experience on the children's development. It was thought that receptive experience would predict vocalic development, since foundational phonemic categories are formulated from ambient (caregiver) language models before the vocal exploration or babbling periods begin in infancy. Vocalic development progresses as children update speech-motor plans to approximate these phonemic categories, so it logically followed that receptive experience—the quantity of caregiver input in Quechua/Spanish—should predict children's vocalic development.

The predictions for coarticulatory development were less straightforward. On the one hand, coarticulation between phonemes is a highly motorized skill, acquired through years of practice and somatosensory feedback experience. On the other hand, this study measured coarticulation in two word environments, potentially implicating children's

morphological awareness, which might be expected to vary with the quantity of Quechua input. Given that the modeling showed a mediating effect of children's language use (Fig. 12), it seems likely that we are witnessing two environmental pressures at work on the children's speech. The first is the adult language model, because children who hear more Quechua may process and segment Quechua words from their ambient language faster and go on to have larger vocabularies (Newman et al. 2006, Weisleder & Fernald 2013). These larger vocabularies may increase the children's morphological awareness (Diamanti et al. 2017), allowing them to analyze morphologically complex words and to distinguish more between word environments via coarticulation in their spoken language (Song, Demuth, Shattuck-Hufnagel, & Ménard 2013).

The second environmental pressure is the frequency of children's own language use. The percentage of Quechua/Spanish clips containing the target child is a proxy, likely standing in for a number of experiential variables relating to speech production. For example, children who use more Quechua have more practice accessing and, as the expressive parameter suggests, ARTICULATING Quechua words. In this way, children's own productions constitute a different form of input, not from caregivers, but from the children themselves via self-auditory and somatosensory feedback systems (see Goldin-Meadow, Mylander, & Franklin 2017 for an additional example of self-input for children's development of morphological structure in emerging signed systems). Note that this variable 'frequency of language use' is independent of children's domain-general fine motor control. Children with increased fine motor control might also be able to distinguish more between word environments via coarticulation. But if fine motor control explained the results, then children who simply talked more—in whichever language—should differentiate more between environments; modeling did not show such a relationship.

Children clearly learn speech and language from their environments. Yet input-output models of child language acquisition are rife with mediating variables, which often require novel analyses to uncover. For example, it was long known that children in North America who heard more child-directed speech grew larger vocabularies (Hart & Risley 1995, Hoff 2003). But only recently have speech perception (eye-tracking) and brain imaging (MRI) analyses uncovered lexical processing speed and Broca's area activation as mediating variables between child-directed speech and lexical outcomes (Romeo et al. 2018, Weisleder & Fernald 2013). Overall, the unique sociolinguistic context of this speech community in Bolivia, where the overall amount of language input and ratio of receptive to expressive experience varied by household, demonstrated the importance of not simply modeling environmental influences on development, but also distinguishing between the type of experience, at least for phonetic and phonological outcomes—something that is almost never done. It is essential to model multiple components of the environment, to rule out (or confirm) environmental influence, to disentangle correlation and causation, and to better explain the factors underlying speech-language development.

6.2. ENVIRONMENTAL EFFECTS ON VOWEL VARIABILITY. Children in this study with monolingual Quechua mothers tended to produce smaller, tighter vowel categories than children with Quechua-dominant mothers or bilingual Quechua-Spanish mothers. Likewise, children with Quechua-dominant mothers tended to have smaller, tighter categories than children with bilingual mothers. While acknowledging that these maternal language profile groups were unbalanced—the monolingual mother group had just ten children compared to twenty-three children in the bilingual group—this finding still suggests a

relationship between a receptive predictor, maternal language patterns, and the children's own speech variation, but not an effect of the children's own language use on their vowel variability: there was no effect of children speaking more Spanish or more Quechua/ mixed speech on their vowel dispersion. The absence of a correlation between child speech production and vowel variation does not mean that the children's language use or expressive language patterns do not affect their vowel variability. It is possible that the language-use estimates made from a single daylong recording were not sufficiently sensitive to capture those individual differences, or a potential role of expressive language, for vowel variation—recall that receptive parameters from the recordings did not explain the children's vowel variability either. Instead, the more distal measure of maternal language profile emerged as the only relevant parameter, suggesting that mediators of vowel variation were established well before daylong recordings of the children were made (i.e. in infancy and toddlerhood), as predicted in §3.2.

Nevertheless, the reasons for the direction of receptive experience require explanation: why do children with bilingual mothers have more variable vowel categories? After all, children of bilingual mothers are likely exposed to more Spanish, a language with five phonemic vowels. This five-vowel contrast could put pressure on the phonological system to REDUCE variability, in an effort to maintain contrast (Bradlow 1995).

There are two potential explanations for why children with bilingual mothers tended to have more variable Quechua vowel categories. First, the categories could be more variable due to a general effect of bilingualism: all of the children receive input in two languages, potentially rendering their phonological targets in both languages more variable. Alternatively, the vowel categories could be more variable due to the specific language combination studied here: Quechua and Spanish. The goal of this discussion is not to distinguish between the effect of bilingual input and the specific effect of Spanish and Quechua; the current study is not designed to make this distinction. Rather, this discussion highlights two reasons for the direction of the effect of maternal language profile on vowel variation.

Children with bilingual mothers may have more variable vowel categories because of the type of language those children are exposed to. Although all of the children tested were bilingual Quechua-Spanish speakers, the primary difference between the maternal language profile groups was, presumably, the quantity of Quechua (and Spanish) that the children heard. The children with bilingual mothers, while probably not exposed to equal amounts of Quechua and Spanish, could have a more mixed linguistic environment than the children with monolingual mothers. Receiving, as an estimation, 50% of their input in one phonological system and 50% in the other could render the children's phonological targets—in both languages—more variable, since those children receive less overall input in each language.

These types of 'bilingual exposure effects' have been found in other bilingual populations, for other outcome measures, as summarized in §2.2. For example, infants bilingual in French and another language who received a larger proportion of their input in the second language had larger receptive vocabularies in that language (Carbajal & Peperkamp 2019). (See also the summaries of Bijeljac-Babic et al. 2012 and Potter et al. 2019 in §2.2.) To my knowledge, a relationship between bilingual exposure and children's vowel production has not previously been quantified. However, it is reasonable to propose that the relative proportion of two languages in a bilingual environment—whatever those languages are—could predict variability in vowel patterning: the more input the child receives in one of their languages, the sharper their phonemic representations may become, and the less variable their spoken vowel categories may be through auditory feedback-generated updates.

Alternatively, vowel categories may be more variable in the bilingual maternal language profile group because of the specific language combination of Quechua and Spanish studied here. Recall that Quechua has three phonemic vowels /a, i, u/ and two allophonic mid vowels [e, o], derived only in uvular contexts (Gallagher 2016). Spanish has the same vowel system, /a, e, i, o, u/, but all of the vowels are phonemic.

Because mid vowels in Quechua are derived only in limited (uvular) contexts, they are less frequent than the peripheral phonemic vowels. An explanation based on language-specific vowel frequencies would predict that the children with monolingual mothers should have smaller, less variable categories. This would hold for all of the tested vowels in Quechua, but ESPECIALLY for the more frequent, peripheral Quechua /a, i, u/. For children with monolingual mothers, these peripheral vowels are more frequent because (i) the children are exposed to more Quechua in their environment and (ii) /a, i, u/ are more frequent than [e] and [o] within Quechua.

The data tend to bear out a prediction based on language-specific vowel frequencies. Consider Table 6 in §5.2 above, which displays each vowel category's dispersion by maternal language profile. The largest difference in vowel category dispersion between the monolingual Quechua group and the bilingual Quechua-Spanish group is for /u/ (0.26), followed by /a/ (0.22). The dispersion coefficient differences between the maternal language profile groups for [e] and [o] are less notable: for [o], it is just 0.03 between the monolingual and Quechua-dominant groups and 0.07 between the monolingual and bilingual groups. In fact, [o] is actually MORE variable in the monolingual Quechua group than in the Quechua-dominant group.

These results suggest that language-specific vowel frequencies may explain some of the variability between maternal language profile groups. Children with monolingual mothers have tighter, more compact vowel categories in Quechua because they receive more input in Quechua, allowing them to more quickly home in on phonemic categories and make speech-motor updates based on auditory feedback in infancy and early childhood. The result is less dispersed vowel categories in their speech production. This exposure-based explanation is further supported by the differences in variability between the Quechua peripheral and mid vowels. Even the children with monolingual Quechua mothers receive relatively less input for the less frequent [e] and [o]; consequently, the differences in variability of [e] and [o] between maternal language profile groups tend to be less notable.[12]

The discussion above leaves us with several conclusions. First, these bilingual children's vowel variability is best predicted by a receptive language measure, maternal bilingual language profile, as exhibited by differences between monolingual Quechua, Quechua-dominant, and bilingual Quechua-Spanish caregiver groups. Differences in the children's own expressive language did not predict any individual vowel variability. A second important conclusion addresses WHY the children with monolingual mothers have tighter, less variable vowel categories: the children with bilingual Quechua-Spanish mothers receive less overall input in Quechua, rendering their phonemic categorization of that language more variable, as reflected in their speech production.

---

[12] One could note that the vowel dispersion coefficient is roughly equivalent between the mid vowels and the peripheral vowels. However, there are numerous reasons—independent of frequency or exposure effects—that some vowels may have greater variability than others. For example, /i/ is known to be a highly stable vowel, consistently exhibiting less within-category variability than low vowels such as /a/. This is often attributed to the inflexible lingual posturing required to approximate /i/ without, for example, articulating a fricative when the tongue is that close to the palate. As a result of these articulatory configuration differences between vowels, it does not NECESSARILY seem reasonable to attribute within-category differences between vowels to external influences such as frequency of exposure.

6.3. ENVIRONMENTAL EFFECTS ON COARTICULATION BY WORD ENVIRONMENT. The children in this speech community who produced more Spanish throughout the day distinguished less between within-morpheme and across-morpheme environments of Quechua words. In line with previous work, phonetic patterning by word environment is assumed to indicate morphological analysis and the parsing of complex words, in adults and children (Cho 2001, Lee-Kim et al. 2013, Song, Demuth, Evans, & Shattuck-Hufnagel 2013, Song, Demuth, Shattuck-Hufnagel, & Ménard 2013). Consequently, the following section outlines what environmental effects on morphological parsing mean for children's word learning and their role in language change.

IMPLICATIONS FOR WORD LEARNING. The relationship between children's expressive language use and morphological patterning by word environment has clear implications for word learning. The current study, for example, showed that children who used more Quechua/mixed speech were better able to analyze the internal morphological structure of complex words. The bilingual children studied here thus demonstrated how different amounts of language experience can result in different word-learning outcomes. This result also predicts that bilingual children who use less of one of their languages could demonstrate a protracted learning period in analyzing the morphological structure of word forms WHEN COMPARED TO THEIR AGE-MATCHED PEERS.[13] If phonetic differentiation by word environment indicates morphological analysis, as work in adult and child morphophonetics suggests (Cho 2001, Lee-Kim et al. 2013, Song, Demuth, Evans, & Shattuck-Hufnagel 2013, Song, Demuth, Shattuck-Hufnagel, & Ménard 2013), then less differentiation by word environment could indicate less robust analysis of morphologically complex word forms. Children who distinguish less between word environments may maintain more unanalyzed lexical forms in their lexicons, affecting the abstraction of sublexical categories like morphemes and phonemes. For example, a child who predominately uses Quechua may have sufficient experience with the language to reliably parse most Quechua suffixes from roots (even monolingual adult speakers do not always parse morphologically complex words—consider words like *illuminate*, *increase*, or *uncouth* in English). For this child with more Quechua experience, the suffixes then become increasingly abstract and independent (from the root morpheme). The result is children's heightened metalinguistic and morphological awareness (Diamanti et al. 2017). The child with relatively less Quechua experience, however, may continue to leave a great many morphologically complex word forms in their lexicon unanalyzed, at least for a longer period of time in development.

Children who use Quechua less almost certainly have different representations of suffixes and word forms than children who use relatively more Quechua. Children who use Quechua less do not unilaterally represent words in an unanalyzed form. Rather, for these children with less experience, the suffixes may simply not be as concrete, parsed away from the original inflected form, as they are for children who use Quechua more. This analysis suggests that the structure of the lexicon, and connections between items within it, varies as a function of language exposure, specifically language PRODUCTION. As such, this conclusion supports previous work that established how the structure of the lexicon differs as a function of language experience (e.g. vocabulary size) in monolingual children (Edwards, Beckman, & Munson 2004, Storkel 2002).

---

[13] Peers matched for lexicon size or number of morphemes per word would likely show exceptions to these predictions. For example, a child who did not use Quechua very frequently in their recording, but has a very large Quechua vocabulary, may still be able to differentiate between word environments more than their language use alone would predict.

Note that these results do not suggest, or demonstrate, protracted morphological productivity in any of the participants. All of the children, regardless of language exposure or use, were capable of completing the morphological extension task. Thus the task results showed, as was assumed, that all of the children were morphologically productive Quechua speakers. The results do, however, suggest something more fine-grained: bilingual children who use less Quechua may be morphologically productive and capable of generalizing affixes to novel lexical environments while SIMULTANEOUSLY leaving some morphologically complex forms unanalyzed. Redundant, conflicting representations of this form, at multiple levels in the grammar, are anticipated in models that assume that linguistic categories, and thus grammars, are constructed on the basis of perceptual and articulatory experience (Hayes 1999, Pierrehumbert 2003, Steriade 1997). Thus, even Spanish-dominant bilingual children could have abstract morpheme categories, though the morpheme categories would not necessarily be as abstract and parseable from the root as those of children who use relatively more Quechua.

IMPLICATIONS FOR LANGUAGE CHANGE. Finding that children's morphological analysis is correlated with their language use and exposure also has implications for historical language change. It is well known that morphosyntactic change can occur as affixes undergo phonetic reduction and fuse to roots during running speech. Change can then occur as speakers (and learners) subsequently fail to analyze the internal structure of complex words. Some have proposed that this change can occur during first language acquisition (Roberts & Roussou 1999, van Gelderen 2004; see Cournane 2017, 2019 for recent arguments). But this assumption of the role of children in language change continues to be the subject of some debate, with some authors arguing that the similarities between language diachrony and ontogeny merely reflect shared cognitive biases between adults and children (Bybee & Slobin 1982, Diessel 2011).

The current study is consistent with the idea that morphological reanalysis could occur during first language acquisition. However, contrary to the viewpoints of children as innovators of language change at other levels of language like syntax or semantics, the current analysis does not maintain (but also does not rule out) that change occurs because child language learners receive insufficient input. Nor does this analysis suggest that change occurs because child learners regularize variable input or overgeneralize from the data they are given (Hudson Kam & Newport 2009, Kerswill & Williams 2000). Instead, particular language-learning environments—such as bilingual environments—could facilitate different LEVELS of morphological analysis (Meakins & Wigglesworth 2013). This graded reanalysis, dependent on the learner's exposure and experience, could lead to change over time (Cournane 2017, Gathercole & Thomas 2009, O'Shannessy 2012). The idea of gradient morphological analysis is fairly consistent with recent research on morphological productivity in adults which has demonstrated that affix productivity depends upon speaker experience—like the frequency ratio of an affix to a stem (Hay 2003, Plag & Baayen 2009).

A graded reanalysis of morphologically complex words might be particularly common during sociolinguistic situations of language shift. For example, in southern Bolivia, rapid language shift to Spanish is occurring in Quechua-speaking communities. But as the current study has demonstrated, this shift does not always occur in a single generation, where the adults speak Quechua and the children of the following generation switch entirely to Spanish. And since receptive language abilities tend to outlast expressive abilities during multigenerational language shift, the ratio of expressive to receptive experiences varies between children in this sample. The result is that children who use, but

specifically PRODUCE, proportionately more Spanish analyze and produce Quechua word forms differently, potentially igniting morphological change.[14] Crucially, these more Spanish-dominant children do not distinguish between the word environments in their spoken language patterns. It is possible that the more Spanish-dominant children will, with time and experience, differentiate more between word environments in their speech. And it is worth stressing that even if they do not, lack of phonetic differentiation by word environment does not indicate pathological or incomplete acquisition: all of these children are fluent, bilingual speakers with native competency. Rather, if the Spanish-dominant children do not eventually distinguish between the word environments in their speech, it is worth considering the structure and acoustic signal of the language input that they might one day provide to their own children—a potential source of language change.

6.4. Long-form recordings and community linguistics. Comprehensive, contextualized speech corpora like those employed in this work could have unique value to speech communities interested in documentation and revitalization efforts. For an extensive discussion of this topic, the reader is referred to Cychosz & Cristia 2021, but some of these ideas are also summarized here. (See also Cychosz et al. 2020 for discussion of the ethics surrounding this data format.) Here the term ‘long-form recordings’ is used instead of ‘daylong’ since the recordings may or may not transpire over an entire day.

First, since naturalistic, long-form recordings take place in and around the home, they have the potential to help communities understand the mechanisms and predictors of language maintenance, be they related to media exposure or intergenerational transmission. The data also constitute records of speech registers, such as youth language or child-directed speech, that are often ignored in documentation projects, despite the great potential for those registers to elucidate linguistic phenomena (Hellwig & Jung 2020). And they record phonological processes, especially those common to child phonology, that are almost never included in formal, grammatical descriptions but, again, often straightforwardly reflect aspects of phonological contrast, frequency, and saliency that are harder to observe in adults (Demuth 1996, Pye, Ingram, & List 1987).

For communities interested in language revitalization efforts, the potential of these long-form recordings for language teaching may be even more promising. While researchers can annotate or gloss portions of a long-form recording in order to ensure its accessibility to those who need explicit documentation of language structure, there are few specialist barriers to reusing long-form recordings for other purposes, because they merely capture individuals going about their daily lives. Cychosz and Cristia (2021) highlight one such example for a speech community undergoing language shift. In that context, long-form recordings could be used to demonstrate which lexical items adults continue to use in the target language but children have started to use in the contact language. This type of data is infrequently documented during traditional one-on-one (researcher-adult) linguistic elicitation. Yet the data have straightforward translational impacts for the speech community. For example, coursebooks written in the target language would then include only words that children are actively using and not those that are considered antiquated by or are simply unknown among younger speakers.

6.5. Future directions.

Maternal speech. In this work, the receptive measure of maternal language profile predicted the children’s vowel variability better than the children’s own language use

---

[14] Adult second language learners could also exhibit some of these same patterns, though they clearly come to the learning problem with distinct learning biases and experiences (Bergmann, Dale, & Lupyan 2016, Lupyan & Dale 2010).

and the proportion of receptive Quechua experience in the child's environment. Given this relationship between variation and maternal language usage, an important next step is to determine the roles of directed versus overheard maternal language input for the construction of children's phonological categories and their phonetic production. We may see an effect of receptive experience in the ambient environment if we computed the quantity of speech in the corpus ONLY from the child's mother, and only in a child-directed register. Examining the differences between directed and overheard speech patterns in bilingual environments would have implications for recent work evaluating the importance of child-directed speech crosslinguistically (Casillas et al. 2020, Cristia 2020, Cristia et al. 2017). It is well known that many cultures do not speak directly to their children until toddlerhood or later (Lieven 1994). Yet it is equally acknowledged that the children socialized into these cultures acquire their native language(s) and reach core language development milestones (Casillas et al. 2020, Cychosz, Cristia, et al. 2021). Here again the fine-grained measures of acoustic phonetics have the potential to make a substantial contribution. If children who receive more direct Quechua speech input, rather than overheard, have tighter, less variable (i.e. more adult-like) vowel categories, this would suggest some benefit for child-directed speech. If, however, the distinction between directed versus overheard speech had no effect on the children's vowel variability, this would suggest that the children were forming phonological categories from all ambient (maternal) language.

Evaluating this question of overheard versus directed speech in the context of language shift, or bilingualism more generally, could be of further interest because the proportion of directed versus overheard speech likely varies by language. For example, most of a child's Spanish exposure could be child-directed, but their Quechua exposure could primarily be overheard (as would be common in many situations of generational language shift). It is plausible that the child's vowel patterning in each language would reflect this difference in input, with Spanish vowels (learned through directed speech) being less variable than Quechua vowels (learned through overheard speech).[15] It is also possible that the distinction between overheard and directed speech is either meaningless for vowel variability or, alternatively, highly dependent on the cultural context. The point here is that sociolinguistic situations of language shift and bilingualism uniquely permit us to evaluate the roles of directed versus overheard speech in speech development. This inquiry is a crucial next step for this line of research.

Another interesting line of research would be to analyze the acoustic phonetics of the mothers' speech and compare it to the children's phonetic outcomes. In the current study, maternal input was instantiated only in language categories (proportion of Spanish, Quechua, etc.). But looking at the mothers' acoustic speech patterns could unearth different, perhaps more subtle, environmental influences. For example, children's expressive experience predicted their coarticulation patterns by word environment, but it is likewise possible that the acoustic-phonetic coarticulation patterns of caregiver speech predict children's coarticulation, an idea that future work could evaluate.

---

[15] It is important to acknowledge that Quechua and Spanish vowel systems, while very similar, differ in the phonemic status of [e] and [o]. This could affect the variability of those vowels, independent of learning context. It is likewise important to acknowledge that comprehensive examinations of vowel contrasts in infant-directed speech actually find a slight tendency for mothers to hypoarticulate in the register (Cristia & Seidl 2014, Martin et al. 2015). This means that perhaps the Spanish vowel categories are more reduced, even in a child-directed register.

ESTIMATING BILINGUAL ENVIRONMENTS OVER MULTIPLE DAYS. This work made use of a large-scale child language corpus consisting of over 500 hours of naturalistic Quechua-Spanish language use captured on daylong audio recordings. The data samples from daylong recordings are robust, and likely more representative than short recording samples made in the lab. Combined with smart sampling and annotation procedures, they offer new insight into children's everyday environments. Nevertheless, the environmental factors were computed from a single daylong recording that does not capture the complexities of a child's learning experience (Orena, Byers-Heinlein, & Polka 2020). For example, when the child participants from this study travel to more rural areas, where more monolingual Quechua speakers live, the children are more likely to speak Quechua. Similarly, the children use almost exclusively Spanish at school.[16] Going forward, more thorough evaluation of bilingual measures derived from daylong recordings might be necessary to ascertain if and how the bilingual environment varies across or within recording days (e.g. morning versus night).

7. CONCLUSION. A current goal of language development research is to discern the extent and mechanisms of environmental influence upon children's speech-language outcomes. Here the influence of the environment was evaluated by studying speech development in a unique context: a bilingual community undergoing language shift. A large-scale, naturalistic child language corpus of South Bolivian-Quechua speech was annotated for children's and their caregivers' bilingual language use. Corpus analyses established that the children's language dominance, and receptive versus expressive language experience, varied by household. Then, language dominance estimates derived from the corpus and background questionnaires were used to predict children's speech patterns. Children with monolingual Quechua mothers had tighter, more compact Quechua vowel categories than children with bilingual or Quechua-dominant mothers. But children's own Quechua language use best predicted their coarticulation patterns by word environment: children who used more Quechua were more likely to differentiate between within- and across-morpheme environments in Quechua words.

On the basis of these relationships between children's exposure and their phonetic production, several arguments pertaining to language development and change were made. First, Quechua-dominant children have more stable phonological categories and differentiate between word environments more, suggesting that increased language exposure predicts more abstract speech-language outcomes, independent of children's age. Second, graded phonetic categorization and morphological analysis within a single bilingual community was proposed to be the cradle of much child-initiated language and sound change over successive generations. Bilingual children who use more of the majority language could ignite change within the minoritized language, especially as language is transmitted to their children. Taken together, these results show how distal and proximal properties of the environment impact phonetic outcomes differently, with implications for language learning and change, demonstrating the need to (i) model exposure effects in speech development, (ii) differentiate between receptive and expressive experience in language development, and (iii) incorporate diverse sociolinguistic contexts in evaluating long-standing questions in linguistic theory.

---

[16] Neither of these contexts was captured in the recordings annotated for this project. These examples simply demonstrate how certain linguistic contexts may be unrepresented in one or even a few daylong recordings.

APPENDIX: STIMULI LISTS AND CLIP ANNOTATION INFORMATION

REAL WORD	TRANSLATION	MORPHEME ENVIRONMENT
chiˈta-pi	'sheep-LOC'	across
cuˈca-pi	'coca (leaves)-LOC'	across
hatunmaˈma-pi	'grandma-LOC'	across
imiˈlla-pi	'girl-LOC'	across
juk'uˈcha-pi	'mouse-LOC'	across
llaˈma-pi	'llama-LOC'	across
llaˈpa-pi	'lightning-LOC'	across
maˈma-pi	'mom-LOC'	across
pamˈpa-pi	'prairie-LOC'	across
paˈpa-pi	'potato-LOC'	across
q'aˈpa-pi	'palm of hand-LOC'	across
sunˈkha-pi	'beard-LOC'	across
t'iˈka-pi	'flower-LOC'	across
thaˈpa-pi	'nest-LOC'	across
uhuˈt'a-pi	'sandal-LOC'	across
waˈka-pi	'cow-LOC'	across
wallˈpa-pi	'chicken-LOC'	across
waˈwa-pi	'baby/child-LOC'	across
ˈpapa	'potato'	within
ˈllapa	'lightning'	within
ˈapi	'corn/citrus drink'	within
ˈthapa	'nest'	within
ˈq'apa	'palm of hand'	within

TABLE A1. Real-word repetition stimuli to elicit [ap]. Note: ˈ indicates stress, ' indicates ejective. Each 'across' item additionally inflected with -*man* (allative); see Table A2.

REAL WORD	TRANSLATION	MORPHEME ENVIRONMENT
chiˈta-man	'sheep-ALL'	across
cuˈca-man	'coca (leaves)-ALL'	across
hatunmaˈma-man	'grandma-ALL'	across
imiˈlla-man	'girl-ALL'	across
juk'uˈcha-man	'mouse-ALL'	across
llaˈma-man	'llama-ALL'	across
llaˈpa-man	'lightning-ALL'	across
maˈma-man	'mom-ALL'	across
pamˈpa-man	'prairie-ALL'	across
paˈpa-man	'potato-ALL'	across
q'aˈpa-man	'palm of hand-ALL'	across
sunˈkha-man	'beard-ALL'	across
t'iˈka-man	'flower-ALL'	across
thaˈpa-man	'nest-ALL'	across
waˈka-man	'cow-ALL'	across
wallˈpa-man	'chicken-ALL'	across
waˈwa-man	'baby/child-ALL'	across
ˈmama	'mom'	within
ˈllama	'llama'	within
hamˈpiri	'healer'	within
hampiˈri-pi	'healer-LOC'	within
ˈpampa	'prairie'	within
hatunˈmama	'grandma'	within

TABLE A2. Real-word repetition stimuli to elicit [am]. Note: ˈ indicates stress, ' indicates ejective.

VOWEL	SYLLABIFIED LEXICAL ITEM	TRANSLATION
[a]	ˈ**a**.pi	'corn/citrus drink'
[a]	ˈp**a**m.pa	'prairie'
[a]	ˈp**a**.pa	'potato'
[a]	ˈth**a**.pa	'nest'
[a]	ˈm**a**.ma	'mom'
[a]	hatunˈm**a**.ma	'grandmother'
[i]	ˈch**i**.ta	'sheep'
[i]	ˈt'**i**.ka	'flower'
[i]	ham.ˈp**i**.ri	'healer'
[i]	ham.pi.ˈr**i**-pi	'healer-LOC'
[i]	a.ˈp**i**-pi	'corn/citrus drink-LOC'
[i]	q'e.ˈp**i**-pi	'bundle-LOC'
[u]	ˈp**u**n.ku	'door'
[u]	ˈp**u**n.chu	'poncho'
[u]	ju.ˈk'**u**.cha	'mouse'
[u]	ˈr**u**n.tu	'egg'
[u]	ˈs**u**n.kha	'beard'
[u]	u.ˈh**u**.t'a	'sandal'
[e]	ˈp'**e**s.qo	'bird'
[e]	ˈq'**e**.pi	'bundle'
[e]	qol.ˈq**e**-pi	'money-LOC'
[o]	ˈq**o**l.qe	'money'
[o]	al.ˈq**o**-pi	'dog-LOC'
[o]	p'es.ˈq**o**-pi	'bird-LOC'

TABLE A3. Vowels analyzed in current study (in bold) and their lexical context. Note: ˈ indicates stress, '
indicates ejective, '.' indicates syllable boundary, '-' indicates morpheme boundary.

SPKR ID	MIXED	UNSURE	NO SPEECH	QUECHUA	SPANISH	TOTAL # OF CLIPS
1003	30 (13.57%)	17 (7.69%)	9 (4.07%)	7 (3.17%)	158 (71.49%)	221
1006	16 (14.29%)	27 (24.11%)	5 (4.46%)	1 (0.89%)	63 (56.25%)	112
1008	19 (10.73%)	23 (12.99%)	54 (30.51%)	42 (23.73%)	39 (22.03%)	177
1018	26 (25.24%)	21 (20.39%)	2 (1.94%)	41 (39.81%)	13 (12.62%)	103
1029	18 (12.68%)	39 (27.46%)	23 (16.20%)	8 (5.63%)	54 (38.03%)	142
1033	31 (12.70%)	25 (10.25%)	26 (10.66%)	25 (10.25%)	137 (56.15%)	244
1034	10 (5.65%)	41 (23.16%)	31 (17.51%)	12 (6.78%)	83 (46.89%)	177
1037	8 (5.19%)	22 (14.29%)	28 (18.18%)	4 (2.60%)	92 (59.74%)	154
1039	19 (8.02%)	73 (30.80%)	28 (11.81%)	68 (28.69%)	49 (20.68%)	237
1042	50 (16.95%)	31 (10.51%)	16 (5.42%)	144 (48.81%)	54 (18.31%)	295
1043	81 (25.47%)	48 (15.09%)	10 (3.14%)	76 (23.90%)	103 (32.39%)	318
1045	29 (10.21%)	68 (23.94%)	105 (36.97%)	25 (8.80%)	57 (20.07%)	284
1049	32 (22.86%)	35 (25.00%)	9 (6.43%)	5 (3.57%)	59 (42.14%)	140
1050	16 (8.94%)	45 (25.14%)	47 (26.26%)	25 (13.97%)	46 (25.70%)	179
1054	24 (6.23%)	116 (30.13%)	83 (21.56%)	33 (8.57%)	129 (33.51%)	385
1055	22 (16.30%)	21 (15.56%)	12 (8.89%)	18 (13.33%)	62 (45.93%)	135
1057	2 (2.38%)	18 (21.43%)	14 (16.67%)	0 (0.00%)	50 (59.52%)	84
1058	62 (24.03%)	42 (16.28%)	26 (10.08%)	43 (16.67%)	85 (32.95%)	258
1062	15 (7.77%)	48 (24.87%)	45 (23.32%)	22 (11.40%)	63 (32.64%)	193
1063	22 (19.13%)	17 (14.78%)	2 (1.74%)	12 (10.43%)	62 (53.91%)	115
1064	9 (4.07%)	30 (13.57%)	15 (6.79%)	19 (8.60%)	148 (66.97%)	221
1065	11 (4.04%)	22 (8.09%)	150 (55.15%)	50 (18.38%)	39 (14.34%)	272
1070	9 (6.62%)	49 (36.03%)	12 (8.82%)	1 (0.74%)	65 (47.79%)	136
1071	16 (10.13%)	33 (20.89%)	29 (18.35%)	14 (8.86%)	66 (41.77%)	158
1076	15 (8.33%)	39 (21.67%)	48 (26.67%)	5 (2.78%)	73 (40.56%)	180
1078	52 (15.07%)	89 (25.80%)	27 (7.83%)	43 (12.46%)	134 (38.84%)	345
1079	33 (21.43%)	21 (13.64%)	22 (14.29%)	13 (8.44%)	65 (42.21%)	154

(TABLE A4. *Continues*)

SPKR ID	MIXED	UNSURE	NO SPEECH	QUECHUA	SPANISH	TOTAL # OF CLIPS
1080	17 (11.56%)	33 (22.45%)	8 (5.44%)	2 (1.36%)	87 (59.18%)	147
1083	13 (7.43%)	70 (40.00%)	37 (21.14%)	11 (6.29%)	44 (25.14%)	175
1085	46 (22.89%)	22 (10.95%)	15 (7.46%)	26 (12.94%)	92 (45.77%)	201
1086	36 (24.66%)	38 (26.03%)	9 (6.16%)	17 (11.64%)	46 (31.51%)	146
1087	5 (5.05%)	23 (23.23%)	1 (1.01%)	1 (1.01%)	69 (69.70%)	99
1088	29 (26.36%)	18 (16.36%)	5 (4.55%)	11 (10.00%)	47 (42.73%)	110
1089	23 (13.45%)	30 (17.54%)	65 (38.01%)	16 (9.36%)	37 (21.64%)	171
1090	16 (8.79%)	10 (5.49%)	92 (50.55%)	26 (14.29%)	38 (20.88%)	182
1091	36 (26.09%)	14 (10.14%)	15 (10.87%)	16 (11.59%)	57 (41.30%)	138
1092	27 (22.50%)	18 (15.00%)	20 (16.67%)	16 (13.33%)	39 (32.50%)	120
1094	19 (8.96%)	55 (25.94%)	76 (35.85%)	11 (5.19%)	51 (24.06%)	212
1095	31 (25.62%)	18 (14.88%)	23 (19.01%)	18 (14.88%)	31 (25.62%)	121
1097	41 (25.47%)	37 (22.98%)	8 (4.97%)	11 (6.83%)	64 (39.75%)	161

TABLE A4. Clip annotation category counts and percentages for each child participant.
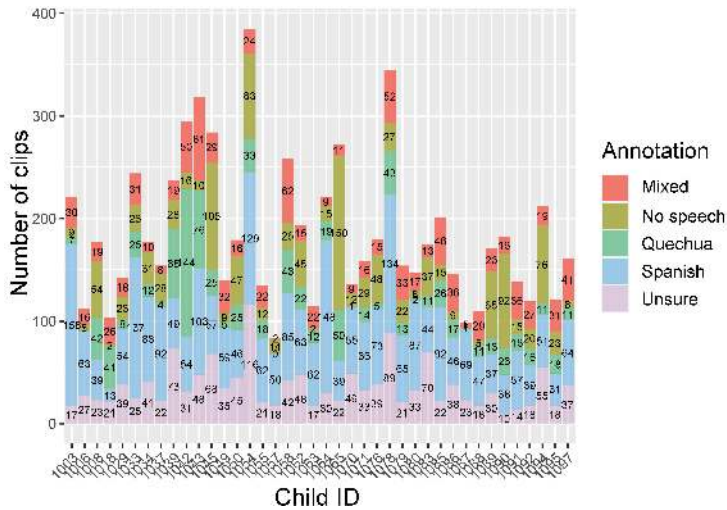

FIGURE A1. Clip annotation category counts for each child. Numbers on barplot reflect the number of clips from each category

REFERENCES

ADRIAANS, FRANS, and DANIEL SWINGLEY. 2017. Prosodic exaggeration within infant-directed speech: Consequences for vowel learnability. *The Journal of the Acoustical Society of America* 141(5).3070–78. DOI: 10.1121/1.4982246.

ASLIN, RICHARD N., and ELISSA L. NEWPORT. 2009. What statistical learning can and can't tell us about language acquisition. *Infant pathways to language: Methods, models, and research directions*, ed. by John Colombo, Peggy McCardle, and Lisa Freund, 15–29. New York: Psychology Press.

BARBIER, GUILLAUME; PASCAL PERRIER; YOHAN PAYAN; MARK K. TIEDE; SILVAIN GERBER; JOSEPH S. PERKELL; and LUCIE MÉNARD. 2020. What anticipatory coarticulation in children tells us about speech motor control maturity. *PLOS ONE* 15(4):e0231484. DOI: 10.1371/journal.pone.0231484.

BATES, DOUGLAS; MARTIN MÄCHLER; BEN BOLKER; and STEVE WALKER. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67(1).1–48. DOI: 10.18637/jss.v067.i01.

BERGMANN, TILL; RICK DALE; and GARY LUPYAN. 2016. Socio-demographic influences on language structure and change: Not all learners are the same. *Behavioral and Brain Sciences* 39:e66. DOI: 10.1017/S0140525X15000710.

BIJELJAC-BABIC, RANKA; JOSETTE SERRES; BARBARA HÖHLE; and THIERRY NAZZI. 2012. Effect of bilingualism on lexical stress pattern discrimination in French-learning infants. *PLOS ONE* 7(2):e30843. DOI: 10.1371/journal.pone.0030843.

BOERSMA, PAUL, and DAVID WEENINK. 2020. Praat: Doing phonetics by computer. Online: http://www.praat.org/.

BRADLOW, ANN R. 1995. A comparative acoustic study of English and Spanish vowels. *The Journal of the Acoustical Society of America* 97(3).1916–24. DOI: 10.1121/1.412 064.

BRANDEKER, MYRTO, and ELIN THORDARDOTTIR. 2015. Language exposure in bilingual toddlers: Performance on nonword repetition and lexical tasks. *American Journal of Speech-Language Pathology* 24(2).126–38. DOI: 10.1044/2015_AJSLP-13-0106.

BYBEE, JOAN L., and DAN I. SLOBIN. 1982. Why small children cannot change language on their own: Suggestions from the English past tense. *Papers from the Fifth International Conference on Historical Linguistics*, ed. by Anders Alqvist, 29–37. Amsterdam: John Benjamins.

CAMACHO RIOS, GLADYS. 2019. Verb morphology in South Bolivian Quechua: A case study of the Uma Piwra rural variety. Austin: University of Texas, Austin master's thesis.

CARBAJAL, MARIA JULIA, and SHARON PEPERKAMP. 2019. Dual language input and the impact of language separation on early lexical development. *Infancy* 25(1).22–45. DOI: 10.1111/infa.12315.

CASILLAS, MARISSA; PENELOPE BROWN; and STEPHEN C. LEVINSON. 2020. Early language experience in a Tseltal Mayan village. *Child Development* 91(5).1819–35. DOI: 10 .1111/cdev.13349.

CAUDRELIER, TIPHAINE; LUCIE MÉNARD; PASCAL PERRIER; JEAN-LUC SCHWARTZ; SILVAIN GERBER; CAMILLE VIDOU; and AMÉLIE ROCHET-CAPELLAN. 2019. Transfer of sensori-motor learning reveals phoneme representations in preliterate children. *Cognition* 192:103973. DOI: 10.1016/j.cognition.2019.05.010.

CHANG, CHARLES B. 2012. Rapid and multifaceted effects of second-language learning on first-language speech production. *Journal of Phonetics* 40(2).249–68. DOI: 10.1016/j .wocn.2011.10.007.

CHIAT, SHULA. 1989. The relation between prosodic structure, syllabification and segmental realization: Evidence from a child with fricative stopping. *Clinical Linguistics & Phonetics* 3(3).223–42. DOI: 10.3109/02699208908985287.

CHO, TAEHONG. 2001. Effects of morpheme boundaries on intergestural timing: Evidence from Korean. *Phonetica* 58(3).129–62. DOI: 10.1159/000056196.

CORE, CYNTHIA; SHREYA CHATURVEDI; and DIEGO MARTINEZ-NADRAMIA. 2017. The role of language experience in nonword repetition tasks in young bilingual Spanish-English speaking children. *Proceedings of the Boston University Conference on Language Development (BUCLD)* 41.179–85.

COURNANE, AILÍS. 2017. In defence of the child innovator. *Micro-change and macro-change in diachronic syntax*, ed. by Eric Mathieu and Robert Truswell, 10–24. Oxford: Oxford University Press. DOI: 10.1093/oso/9780198747840.003.0002.

COURNANE, AILÍS. 2019. A developmental view on incrementation in language change. *Theoretical Linguistics* 45(3–4).127–50. DOI: 10.1515/tl-2019-0010.

CRISTIA, ALEJANDRINA. 2011. Fine-grained variation in caregivers' /s/ predicts their infants' /s/ category. *The Journal of the Acoustical Society of America* 129(5).3271–80. DOI: 10.1121/1.3562562.

CRISTIA, ALEJANDRINA. 2020. Language input and outcome variation as a test of theory plausibility: The case of early phonological acquisition. *Developmental Review* 57:100914. DOI: 10.1016/j.dr.2020.100914.

CRISTIA, ALEJANDRINA; EMMANUEL DUPOUX; MICHAEL GURVEN; and JONATHAN STIEG-LITZ. 2017. Child-directed speech is infrequent in a forager-farmer population: A time allocation study. *Child Development* 90(3).759–73. DOI: 10.1111/cdev.12974.

CRISTIA, ALEJANDRINA, and AMANDA SEIDL. 2014. The hyperarticulation hypothesis of infant-directed speech. *Journal of Child Language* 41(4).913–34. DOI: 10.1017/S030500 0914000105.

CULBERTSON, JENNIFER, and ELISSA L. NEWPORT. 2015. Harmonic biases in child learners: In support of language universals. *Cognition* 139.71–82. DOI: 10.1016/j.cognition .2015.02.007.

CULBERTSON, JENNIFER, and KATHRYN SCHULER. 2019. Artificial language learning in children. *Annual Review of Linguistics* 5(1).353–73. DOI: 10.1146/annurev-linguistics-011718-012329.

CYCHOSZ, MARGARET. 2018. Cychosz HomeBank corpus. DOI: 10.21415/YFYW-HE74.

CYCHOSZ, MARGARET, and ALEJANDRINA CRISTIA. 2021. Using big data from long-form recordings to study development and optimize societal impact. Open Science Framework preprint. Online: https://osf.io/ybqfw/.

CYCHOSZ, MARGARET; ALEJANDRINA CRISTIA; ELIKA BERGELSON; MARISA CASILLAS; GLADYS BAUDET; ANNE S. WARLAUMONT; CAMILA SCAFF; LISA YANKOWITZ; and AMANDA SEIDL. 2021. Vocal development in a large-scale crosslinguistic corpus. *Developmental Science* 24(5):e13090. DOI: 10.1111/desc.13090.

CYCHOSZ, MARGARET; JAN R. EDWARDS; BENJAMIN MUNSON; and KEITH JOHNSON. 2019. Spectral and temporal measures of coarticulation in child speech. *The Journal of the Acoustical Society of America* 146(6).EL516–EL522. DOI: 10.1121/1.5139201.

CYCHOSZ, MARGARET; RACHEL ROMEO; MELANIE SODERSTROM; CAMILA SCAFF; HILLARY GANEK; ALEJANDRINA CRISTIA; MARISA CASILLAS; KAYA DE BARBARO; JANET Y. BANG; and ADRIANA WEISLEDER. 2020. Longform recordings of everyday life: Ethics for best practices. *Behavior Research Methods* 52.1951–69. DOI: 10.3758/s13428-020-01365-9.

CYCHOSZ, MARGARET; ANELE VILLANUEVA; and ADRIANA WEISLEDER. 2021. Efficient estimation of children's language exposure in two bilingual communities. *Journal of Speech, Language and Hearing Research* 64(10).3843–66. DOI: 10.1044/2021_JSLHR-20-00755.

DE BOYSSON-BARDIES, BÉNÉDICTE, and MARILYN MAY VIHMAN. 1991. Adaptation to language: Evidence from babbling and first words in four languages. *Language* 67(2).297–319. DOI: 10.2307/415108.

DEMUTH, KATHERINE. 1996. Alignment, stress, and parsing in early phonological words. *Proceedings of the UBC International Conference on Phonological Acquisition*, 113–25.

DEPAOLIS, RORY A.; MARILYN M. VIHMAN; and TAMAR KEREN-PORTNOY. 2011. Do production patterns influence the processing of speech in prelinguistic infants? *Infant Behavior and Development* 34(4).590–601. DOI: 10.1016/j.infbeh.2011.06.005.

DIAMANTI, VASSILIKI; ANGELIKI MOUZAKI; ASIMINA RALLI; FAYE ANTONIOU; SOFIA PAPAIOANNOU; and ATHANASSIOS PROTOPAPAS. 2017. Preschool phonological and morphological awareness as longitudinal predictors of early reading and spelling development in Greek. *Frontiers in Psychology* 8:2039. DOI: 10.3389/fpsyg.2017.02039.

DIESSEL, HOLGER. 2011. Grammaticalization and language acquisition. *The Oxford handbook of grammaticalization*, ed. by Bernd Heine and Heiko Narrog, 130–41. Oxford: Oxford University Press. DOI: 10.1093/oxfordhb/9780199586783.013.0011.

EDWARDS, JAN R., and MARY E. BECKMAN. 2008. Some cross-linguistic evidence for modulation of implicational universals by language-specific frequency effects in phonological development. *Language Learning and Development* 4(2).122–56. DOI: 10.1080/15475440801922115.

EDWARDS, JAN R.; MARY E. BECKMAN; and BENJAMIN MUNSON. 2004. The interaction between vocabulary size and phonotactic probability effects on children's production accuracy and fluency in novel word repetition. *Journal of Speech, Language, and Hearing Research* 47(2).421–36. DOI: 10.1044/1092-4388(2004/034).

EGUCHI, SUCO, and IRA J. HIRSH. 1969. Development of speech sounds in children. *Acta Otolaryngolica* (Supplement 257), 21–51.

ERSKINE, MICHELLE E.; BENJAMIN MUNSON; and JAN R. EDWARDS. 2020. Relationship between early phonological processing and later phonological awareness: Evidence from nonword repetition. *Applied Psycholinguistics* 41(2).319–46. DOI: 10.1017/S0142716419000547.

FARABOLINI, GIANMATTEO; PASQUALE RINALDI; MARIA CRISTINA CASELLI; and ALEJANDRINA CRISTIA. 2021. Non-word repetition in bilingual children: The role of language exposure, vocabulary scores and environmental factors. *Speech, Language and Hearing*. DOI: 10.1080/2050571X.2021.1879609.

FERJAN RAMÍREZ, NAJA; SARAH ROSEBERRY LYTLE; MELANIE FISH; and PATRICIA K. KUHL. 2019. Parent coaching at 6 and 10 months improves language outcomes at 14

months: A randomized controlled trial. *Developmental Science* 22(3):e12762. DOI: 10 .1111/desc.12762.

FOULKES, PAUL; GERRY DOCHERTY; and DOMINIC WATT. 1999. Tracking the emergence of structured variation. *Leeds Working Papers in Linguistics* 7:23.

GAGLIARDI, ANNIE, and JEFFREY LIDZ. 2014. Statistical insensitivity in the acquisition of Tsez noun classes. *Language* 90(1).58–89. DOI: 10.1353/lan.2014.0013.

GALLAGHER, GILLIAN. 2016. Vowel height allophony and dorsal place contrasts in Cochabamba Quechua. *Phonetica* 73(2).101–19. DOI: 10.1159/000443651.

GARCIA-SIERRA, ADRIAN; NAIRAN RAMÍREZ-ESPARZA; and PATRICIA K. KUHL. 2016. Relationships between quantity of language input and brain responses in bilingual and monolingual infants. *International Journal of Psychophysiology* 110.1–17. DOI: 10 .1016/j.ijpsycho.2016.10.004.

GATHERCOLE, VIRGINIA C. MUELLER, and ENLLI MÔN THOMAS. 2009. Bilingual first-language development: Dominant language takeover, threatened minority language take-up. *Bilingualism: Language and Cognition* 12(2).213–37. DOI: 10.1017/S1366728909 004015.

GEROSA, M.; SUNGBOK LEE; D. GIULIANI; and SHRIKANTH NARAYANAN. 2006. Analyzing children's speech: An acoustic study of consonants and consonant-vowel transition. *2006 IEEE International Conference on Acoustics Speed and Signal Processing Proceedings*, vol. 1, 393–96. DOI: 10.1109/ICASSP.2006.1660040.

GOLDIN-MEADOW, SUSAN; CAROLYN MYLANDER; and AMY FRANKLIN. 2007. How children make language out of gesture: Morphological structure in gesture systems developed by American and Chinese deaf children. *Cognitive Psychology* 55(2).87–135. DOI: 10 .1016/j.cogpsych.2006.08.001.

GREENWOOD, CHARLES R.; KATHY THIEMANN-BOURQUE; DALE WALKER; JAY BUZHARDT; and JILL GILKERSON. 2011. Assessing children's home language environments using automatic speech recognition technology. *Communication Disorders Quarterly* 32(2). 83–92. DOI: 10.1177/1525740110367826.

GUENTHER, FRANK H. 2006. Cortical interactions underlying the production of speech sounds. *Journal of Communication Disorders* 39(5).350–65. DOI: 10.1016/j.jcomdis .2006.06.013.

GUION, SUSAN G. 2003. The vowel systems of Quichua-Spanish bilinguals. *Phonetica* 60(2). 98–128. DOI: 10.1159/000071449.

HA, SEUNGHEE; CYNTHIA J. JOHNSON; D. KIMBROUGH OLLER; and HYUNJOO YOO. 2021. Cross-linguistic comparison of utterance shapes in Korean- and English-learning children: An ambient language effect. *Infant Behavior and Development* 62:101528. DOI: 10.1016/j.infbeh.2021.101528.

HART, BETTY, and TODD RISLEY. 1995. *Meaningful differences in the everyday experience of young American children*. Baltimore: Paul H. Brookes.

HAY, JENNIFER. 2003. *Causes and consequences of word structure*. London: Routledge.

HAYES, BRUCE P. 1999. Phonetically driven phonology: The role of optimality theory and inductive grounding. *Functionalism and formalism in linguistics, vol. 1: General papers*, ed. by Michael Darnell, Edith A. Moravscik, Michael Noonan, Frederick J. Newmeyer, and Kathleen Wheatley, 243–85. Amsterdam: John Benjamins.

HELLWIG, BIRGIT, and DAGMAR JUNG. 2020. Child-directed language—and how it informs the documentation and description of the adult language. *Language Documentation & Conservation* 14.188–214. Online: http://hdl.handle.net/10125/24920.

HENRIKSEN, NICHOLAS; LORENZO GARCÍA-AMAYA; ANDRIES W. COETZEE; and DAAN WISSING. 2019. Language contact in Patagonia: Durational control in the acquisition of Spanish and Afrikaans phonology. *The Routledge handbook of Spanish phonology*, ed. by Sonia Colina and Fernando Martínez-Gil, 416–38. New York: Routledge. DOI: 10.4324/9781315228112-23.

HLAVAC, MAREK. 2018. stargazer: Well-formatted regression and summary statistics tables. R package version 5.2.2. Online: https://cran.r-project.org/web/packages/stargazer /index.html.

HOFF, ERIKA. 2003. The specificity of environmental influence: Socioeconomic status affects early vocabulary development via maternal speech. *Child Development* 74(5). 1368–78. DOI: 10.1111/1467-8624.00612.

Hornberger, Nancy H. 2009. Multilingual education policy and practice: Ten certainties (grounded in Indigenous experience). *Language Teaching* 42(2).197–211. DOI: 10.1017/S0261444808005491.

Hudson Kam, Carla L., and Elissa L. Newport. 2009. Getting it right by getting it wrong: When learners change languages. *Cognitive Psychology* 59(1).30–66. DOI: 10.1016/j.cogpsych.2009.01.001.

Huttenlocher, Janellen; Marina Vasilyeva; Elina Cymerman; and Susan Levine. 2002. Language input and child syntax. *Cognitive Psychology* 45(3).337–74. DOI: 10.1016/S0010-0285(02)00500-5.

Icht, Michal, and Yaniv Mama. 2015. The production effect in memory: A prominent mnemonic in children. *Journal of Child Language* 42(5).1102–24. DOI: 10.1017/S0305000914000713.

Inkelas, Sharon, and Yvan Rose. 2007. Positional neutralization: A case study from child language. *Language* 83(4).707–36. DOI: 10.1353/lan.2008.0000.

Kerswill, Paul, and Ann Williams. 2000. Creating a New Town koine: Children and language change in Milton Keynes. *Language in Society* 29(1).65–115. DOI: 10.1017/S0047404500001020.

Khattab, Ghada. 2003. Age, input, and language mode factors in the acquisition of VOT by English-Arabic bilingual children. *Proceedings of the 15th International Congress of Phonetic Sciences (ICPhS)*, Barcelona, 3213–16. Online: https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2003/p15_3213.html.

Kuznetsova, Alexandra; Per B. Brockhoff; and Rune H. B. Christensen. 2017. lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software* 82(13).1–26. DOI: 10.18637/jss.v082.i13.

Lee, Sungbok; Alexandros Potamianos; and Shrikanth Narayanan. 1999. Acoustics of children's speech: Developmental changes of temporal and spectral parameters. *The Journal of the Acoustical Society of America* 105(3).1455–68. DOI: 10.1121/1.426686.

Lee-Kim, Sang-Im; Lisa Davidson; and Sangjin Hwang. 2013. Morphological effects on the darkness of English intervocalic /l/. *Laboratory Phonology* 4(2).475–511. DOI: 10.1515/lp-2013-0015.

Lehiste, Ilse, and Linda Shockey. 1972. On the perception of coarticulation effects in English VCV syllables. *Journal of Speech and Hearing Research* 15(3).500–506. DOI: 10.1044/jshr.1503.500.

Lidz, Jeffrey, and Annie Gagliardi. 2015. How nature meets nurture: Universal grammar and statistical learning. *Annual Review of Linguistics* 1.333–53. DOI: 10.1146/annurev-linguist-030514-125236.

Lieven, Elena V. M. 1994. Crosslinguistic and crosscultural aspects of language addressed to children. *Input and interaction in language acquisition*, ed. by Clare Gallaway and Brian J. Richards, 56–73. Cambridge: Cambridge University Press. DOI: 10.1017/CBO9780511620690.005.

Liu, Huei-Mei; Patricia K. Kuhl; and Feng-Ming Tsao. 2003. An association between mothers' speech clarity and infants' speech discrimination skills. *Developmental Science* 6(3).F1–F10. DOI: 10.1111/1467-7687.00275.

Lobanov, Boris M. 1971. Classification of Russian vowels spoken by different speakers. *The Journal of the Acoustical Society of America* 49(2B).606–8. DOI: 10.1121/1.1912396.

Lupyan, Gary, and Rick Dale. 2010. Language structure is partly determined by social structure. *PLOS ONE* 5:e8559. DOI: 10.1371/journal.pone.0008559.

Mahr, Tristan, and Jan R. Edwards. 2018. Using language input and lexical processing to predict vocabulary size. *Developmental Science* 21(6):e12685. DOI: 10.1111/desc.12685.

Marchman, Virginia A., and Carmen Martínez-Sussmann. 2002. Concurrent validity of caregiver/parent report measures of language for children who are learning both English and Spanish. *Journal of Speech, Language, and Hearing Research* 45(5).983–97. DOI: 10.1044/1092-4388(2002/080).

Martin, Andrew; Thomas Schatz; Maarten Versteegh; Kouki Miyazawa; Reiko Mazuka; Emmanuel Dupoux; and Alejandrina Cristia. 2015. Mothers speak less

clearly to infants than to adults: A comprehensive test of the hyperarticulation hypothesis. *Psychological Science* 26(3).341–47. DOI: 10.1177/0956797614562453.

Maye, Jessica; Janet F. Werker; and LouAnn Gerken. 2002. Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition* 82(3).B101–B111. DOI: 10.1016/S0010-0277(01)00157-3.

Mayr, Robert; Gwennan Howells; and Rhonwen Lewis. 2015. Asymmetries in phonological development: The case of word-final cluster acquisition in Welsh–English bilingual children. *Journal of Child Language* 42(1).146–79. DOI: 10.1017/S0305000913 000603.

Mayr, Robert, and Aysha Siddika. 2018. Inter-generational transmission in a minority language setting: Stop consonant production by Bangladeshi heritage children and adults. *International Journal of Bilingualism* 22(3).255–84. DOI: 10.1177/13670069 16672590.

McAuliffe, Michael; Michaela Socolof; Sarah Mihuc; Michael Wagner; and Morgan Sonderegger. 2017. Montreal Forced Aligner: Trainable text-speech alignment using Kaldi. *Proceedings of Interspeech 2017*, 498–502. DOI: 10.21437/Inter speech.2017-1386.

McFee, Brian; Colin Raffel; Dawen Liang; Daniel P. W. Ellis; Matt McVicar; Eric Battenberg; and Oriel Nieto. 2015. librosa: Audio and music signal analysis in Python. *Proceedings of the 14th Python in Science Conference*, 18–24. DOI: 10.25080 /Majora-7b98e3ed-003.

Meakins, Felicity, and Gillian Wigglesworth. 2013. How much input is enough? Correlating comprehension and child language input in an endangered language. *Journal of Multilingual and Multicultural Development* 34(2).171–88. DOI: 10.1080/01434632 .2012.733010.

Ménard, Lucie; Jean-Luc Schwartz; Louis-Jean Boë; and Jérôme Aubin. 2007. Articulatory–acoustic relationships during vocal tract growth for French vowels: Analysis of real data and simulations with an articulatory model. *Journal of Phonetics* 35(1).1–19. DOI: 10.1016/j.wocn.2006.01.003.

Mintz, Toben H.; Rachel L. Walker; Ashlee Welday; and Celeste Kidd. 2018. Infants' sensitivity to vowel harmony and its role in segmenting speech. *Cognition* 171.95–107. DOI: 10.1016/j.cognition.2017.10.020.

Mooney, Damien. 2019. Phonetic transfer in language contact: Evidence for equivalence classification in the mid-vowels of Occitan-French bilinguals. *Journal of the International Phonetic Association* 49(1).53–85. DOI: 10.1017/S0025100317000366.

Moulin-Frier, Clément; Sao M. Nguyen; and Pierre-Yves Oudeyer. 2014. Self-organization of early vocal development in infants and machines: The role of intrinsic motivation. *Frontiers in Psychology* 4:1006. DOI: 10.3389/fpsyg.2013.01006.

Mousikou, Petroula; Patrycja Strycharczuk; Alice Turk; and James M. Scobbie. 2021. Coarticulation across morpheme boundaries: An ultrasound study of past-tense inflection in Scottish English. *Journal of Phonetics* 88:101101. DOI: 10.1016/j.wocn .2021.101101.

Muysken, Pieter. 2019. Multilingualism and mixed language in the mines of Potosí (Bolivia). *International Journal of the Sociology of Language* 2019(258).121–42. DOI: 10 .1515/ijsl-2019-2031.

Newman, Rochelle; Nan Bernstein Ratner; Ann Marie Jusczyk; Peter W. Jusczyk; and Kathy Ayala Dow. 2006. Infants' early ability to segment the conversational speech signal predicts later language development: A retrospective analysis. *Developmental Psychology* 42(4).643–55. DOI: 10.1037/0012-1649.42.4.643.

Nielsen, Kuniko. 2014. Phonetic imitation by young children and its developmental changes. *Journal of Speech, Language, and Hearing Research* 57(6).2065–75. DOI: 10.1044/2014_JSLHR-S-13-0093.

Noiray, Aude; Anisia Popescu; Helene Killmer; Elina Rubertus; Stella Krüger; and Lisa Hintermeier. 2019. Spoken language development and the challenge of skill integration. *Frontiers in Psychology* 10:2777. DOI: 10.3389/fpsyg.2019.02777.

Öhman, S. E. G. 1966. Coarticulation in VCV utterances: Spectrographic measurements. *The Journal of the Acoustical Society of America* 39(1).151–68. DOI: 10.1121/1.1909 864.

OLLER, D. KIMBROUGH. 2000. *The emergence of the speech capacity*. Mahwah, NJ: Lawrence Erlbaum.

OLLER, D. KIMBROUGH; REBECCA E. EILERS; RICHARD URBANO; and ALAN B. COBO-LEWIS. 1997. Development of precursors to speech in infants exposed to two languages. *Journal of Child Language* 24(2).407–25. DOI: 10.1017/S0305000997003097.

ONOSSON, SKY, and JESSE STEWART. 2021. The effects of language contact on non-native vowel sequences in lexical borrowings: The case of Media Lengua. *Language and Speech*, OnlineFirst. DOI: 10.1177/00238309211014911.

ORENA, ADRIEL JOHN; KRISTA BYERS-HEINLEIN; and LINDA POLKA. 2020. What do bilingual infants actually hear? Evaluating measures of language input to bilingual-learning 10-month-olds. *Developmental Science* 23(2):e12901. DOI: 10.1111/desc.12901.

O'SHANNESSY, CARMEL. 2012. The role of codeswitched input to children in the origin of a new mixed language. *Linguistics* 50(2).305–40. DOI: 10.1515/ling-2012-0011.

PACE, AMY; RUFAN LUO; KATHY HIRSH-PASEK; and ROBERTA MICHNICK GOLINKOFF. 2017. Identifying pathways between socioeconomic status and language development. *Annual Review of Linguistics* 3.285–308. DOI: 10.1146/annurev-linguistics-011516-034226.

PARDO, JENNIFER S. 2006. On phonetic convergence during conversational interaction. *The Journal of the Acoustical Society of America* 119(4).2382–93. DOI: 10.1121/1.2178720.

PARDO, JENNIFER S.; RACHEL GIBBONS; ALEXANDRA SUPPES; and ROBERT M. KRAUSS. 2012. Phonetic convergence in college roommates. *Journal of Phonetics* 40(1).190–97. DOI: 10.1016/j.wocn.2011.10.001.

PEARL, LISA, and JEFFREY LIDZ. 2009. When domain-general learning fails and when it succeeds: Identifying the contribution of domain specificity. *Language Learning and Development* 5(4).235–65. DOI: 10.1080/15475440902979907.

PEARSON, BARBARA Z.; SYLVIA C. FERNANDEZ; VANESSA LEWEDEG; and D. KIMBROUGH OLLER. 1997. The relation of input factors to lexical learning by bilingual infants. *Applied Psycholinguistics* 18(1).41–58. DOI: 10.1017/S0142716400009863.

PELUCCHI, BRUNA; JESSICA F. HAY; and JENNY R. SAFFRAN. 2009. Statistical learning in a natural language by 8-month-old infants. *Child Development* 80(3).674–85. DOI: 10.1111/j.1467-8624.2009.01290.x.

PERKELL, JOSEPH S. 2012. Movement goals and feedback and feedforward control mechanisms in speech production. *Journal of Neurolinguistics* 25(5).382–407. DOI: 10.1016/j.jneuroling.2010.02.011.

PIERREHUMBERT, JANET B. 2003. Phonetic diversity, statistical learning, and acquisition of phonology. *Language and Speech* 46(2–3).115–54. DOI: 10.1177/00238309030460020501.

PLACE, SILVIA, and ERIKA HOFF. 2011. Properties of dual language exposure that influence 2-year-olds' bilingual proficiency. *Child Development* 82(6).1834–49. DOI: 10.1111/j.1467-8624.2011.01660.x.

PLACE, SILVIA, and ERIKA HOFF. 2016. Effects and noneffects of input in bilingual environments on dual language skills in 2½-year-olds. *Bilingualism: Language and Cognition* 19(5).1023–41. DOI: 10.1017/S1366728915000322.

PLAG, INGO, and HARALD BAAYEN. 2009. Suffix ordering and morphological processing. *Language* 85(1).109–52. DOI: 10.1353/lan.0.0087.

POPESCU, ANISIA, and AUDE NOIRAY. 2021. Learning to read interacts with children's spoken language fluency. *Language Learning and Development* 18(2).151–70. DOI: 10.1080/15475441.2021.1941032.

POTTER, CHRISTINE E.; EVA FOURAKIS; ELIZABETH MORIN-LESSARD; KRISTA BYERS-HEINLEIN; and CASEY LEW-WILLIAMS. 2019. Bilingual toddlers' comprehension of mixed sentences is asymmetrical across their two languages. *Developmental Science* 22(4):e12794. DOI: 10.1111/desc.12794.

PYE, CLIFTON; DAVID INGRAM; and HELEN LIST. 1987. A comparison of initial consonant acquisition in English and Quiché. *Children's language*, vol. 6, ed. by Keith E. Nelson and Anne van Kleeck, 175–90. Hillsdale, NJ: Lawrence Erlbaum.

REIDY, PATRICK F.; KAYLA KRISTENSEN; MATTHEW B. WINN; RUTH Y. LITOVSKY; and JAN R. EDWARDS. 2017. The acoustics of word-initial fricatives and their effect on word-

level intelligibility in children with bilateral cochlear implants. *Ear and Hearing* 38(1). 42–56. DOI: 10.1097/AUD.0000000000000349.

ROBERTS, IAN, and ANNA ROUSSOU. 1999. A formal approach to 'grammaticalization'. *Linguistics* 37.1011– 41. DOI: 10.1515/ling.37.6.1011.

ROMEO, RACHEL R.; JULIA A. LEONARD; SYDNEY T. ROBINSON; MARTIN R. WEST; ALLYSON P. MACKEY; MEREDITH L. ROWE; and JOHN D. E. GABRIELI. 2018. Beyond the 30-million-word gap: Children's conversational exposure is associated with language-related brain function. *Psychological Science* 29(5).700–710. DOI: 10.1177/095 6797617742725.

RSTUDIO TEAM. 2020. RStudio: Integrated development for R. Boston: RStudio. Online: http://www.rstudio.com/.

SHARP, KATHRYN M., and VIRGINIA C. M. GATHERCOLE. 2013. Can a novel word repetition task be a language-neutral assessment tool? Evidence from Welsh–English bilingual children. *Child Language Teaching and Therapy* 29(1).77–89. DOI: 10.1177/0265659 012465208.

SIMONET, MIQUEL. 2011. Intonational convergence in language contact: Utterance-final F0 contours in Catalan-Spanish early bilinguals. *Journal of the International Phonetic Association* 41(2).157–84. DOI: 10.1017/S0025100311000120.

SMITH, JENNIFER; MERCEDES DURHAM; and LIANE FORTUNE. 2009. Universal and dialect-specific pathways of acquisition: Caregivers, children, and t/d deletion. *Language Variation and Change* 21(1).69–95. DOI: 10.1017/S0954394509000039.

SONG, JAE YUNG; KATHERINE DEMUTH; KAREN EVANS; and STEFANIE SHATTUCK-HUFNAGEL. 2013. Durational cues to fricative codas in 2-year-olds' American English: Voicing and morphemic factors. *Journal of the Acoustical Society of America* 133(5).2931–46. DOI: 10.1121/1.4795772.

SONG, JAE YUNG; KATHERINE DEMUTH; STEFANIE SHATTUCK-HUFNAGEL; and LUCIE MÉNARD. 2013. The effects of coarticulation and morphological complexity on the production of English coda clusters: Acoustic and articulatory evidence from 2-year-olds and adults using ultrasound. *Journal of Phonetics* 41(3–4).281–95. DOI: 10.1016/j.wocn .2013.03.004.

STERIADE, DONCA. 1997. Phonetics in phonology: The case of laryngeal neutralization. Los Angeles: University of California, Los Angeles, MS. Online: https://linguistics.ucla.edu /people/steriade/papers/PhoneticsInPhonology.pdf.

STEVENS, KENNETH N. 1989. On the quantal nature of speech. *Journal of Phonetics* 17(1– 2).3–45. DOI: 10.1016/S0095-4470(19)31520-7.

STOEHR, ANTJE; TITIA BENDERS; JANET G. VAN HELL; and PAULA FIKKERT. 2019. Bilingual preschoolers' speech is associated with non-native maternal language input. *Language Learning and Development* 15(1).75–100. DOI: 10.1080/15475441.2018.1533473.

STOEL-GAMMON, CAROL, and PAULA BECKETT HERRINGTON. 1990. Vowel systems of normally developing and phonologically disordered children. *Clinical Linguistics & Phonetics* 4(2).145–60. DOI: 10.3109/02699209008985478.

STORKEL, HOLLY L. 2002. Restructuring of similarity neighbourhoods in the developing mental lexicon. *Journal of Child Language* 29(2).251–74. DOI: 10.1017/S0305000902 005032.

SWINGLEY, DANIEL. 2009. Contributions of infant word learning to language development. *Philosophical Transactions of the Royal Society B: Biological Sciences* 364(1536). 3617–32. DOI: 10.1098/rstb.2009.0107.

THOMASON, SARAH G. 2001. *Language contact*. Edinburgh: Edinburgh University Press.

THORDARDOTTIR, ELIN. 2011. The relationship between bilingual exposure and vocabulary development. *International Journal of Bilingualism* 15(4).426–45. DOI: 10.1177/1367 006911403202.

TORERO, ALFREDO. 1964. Los dialectos Quechuas. *Anales Científicos de la Universidad Agraria* 2.446–78.

TURNER, RICHARD E.; THOMAS C. WALTERS; JESSICA J. M. MONAGHAN; and ROY D. PATTERSON. 2009. A statistical, formant-pattern model for segregating vowel type and vocal-tract length in developmental formant data. *The Journal of the Acoustical Society of America* 125(4).2374–86. DOI: 10.1121/1.3079772.

USOLTSEV, ALEXANDER. 2015. Voice activity detector—Python. Online: https://github.com /marsbroshok/VAD-python.

van Coetsem, Frans. 1988. *Loan phonology and the two transfer types in language contact*. Dordrecht: Foris.

van Gelderen, Elly. 2004. *Grammaticalization as economy*. Amsterdam: John Benjamins.

Vihman, Marilyn M. 2014. *Phonological development: The first two years*. 2nd edn. Boston: Wiley-Blackwell.

Vorperian, Houri K., and Ray D. Kent. 2007. Vowel acoustic space development in children: A synthesis of acoustic and anatomic data. *Journal of Speech, Language, and Hearing Research* 50(6).1510–45. DOI: 10.1044/1092-4388(2007/104).

Watanabe, Akira. 2001. Formant estimation method using inverse-filter control. *IEEE Transactions on Speech and Audio Processing* 9(4).317–26. DOI: 10.1109/89.917677.

Weisleder, Adriana, and Anne Fernald. 2013. Talking to children matters: Early language experience strengthens processing and builds vocabulary. *Psychological Science* 24(11).2143–52. DOI: 10.1177/0956797613488145.

Werker, Janet F., and Suzanne Curtin. 2005. PRIMIR: A developmental framework of infant speech processing. *Language Learning and Development* 1(2).197–234. DOI: 10.1080/15475441.2005.9684216.

Wickham, Hadley. 2016. *ggplot2: Elegant graphics for data analysis*. New York: Springer.

Yang, Charles D. 2004. Universal grammar, statistics or both? *Trends in Cognitive Sciences* 8(10).451–56. DOI: 10.1016/j.tics.2004.08.006.

Yang, Charles D.; Stephen Crain; Robert C. Berwick; Noam Chomsky; and Johan J. Bolhuis. 2017. The growth of language: Universal grammar, experience, and principles of computation. *Neuroscience & Biobehavioral Reviews* 81.103–19. DOI: 10.1016/j.neubiorev.2016.12.023.

Yao, Yao, and Charles B. Chang. 2016. On the cognitive basis of contact-induced sound change: Vowel merger reversal in Shanghainese. *Language* 92(2).433–67. DOI: 10.1353/lan.2016.0031.

Zamuner, Tania S. 2009. Phonotactic probabilities at the onset of language development: Speech production and word position. *Journal of Speech, Language, and Hearing Research* 52(1).49–60. DOI: 10.1044/1092-4388(2008/07-0138).

Zamuner, Tania S.; Stephanie Strahm; Elizabeth Morin-Lessard; and Michael P. A. Page. 2018. Reverse production effect: Children recognize novel words better when they are heard rather than produced. *Developmental Science* 21(4):e12636. DOI: 10.1111/desc.12636.

Zharkova, Natalia; Nigel Hewlett; and William J. Hardcastle. 2011. Coarticulation as an indicator of speech motor control development in children: An ultrasound study. *Motor Control* 15(1).118–40. DOI: 10.1123/mcj.15.1.118.

Zharkova, Natalia; Nigel Hewlett; William J. Hardcastle; and Robin J. Lickley. 2014. Spatial and temporal lingual coarticulation and motor control in preadolescents. *Journal of Speech, Language, and Hearing Research* 57(2).374–88. DOI: 10.1044/2014_JSLHR-S-11-0350.

0100 Samuel J. LeFrak Hall
University of Maryland, College Park
7251 Preinkert Dr.
College Park, MD 20742
[mcychosz@umd.edu]