Word structure in early Quechua speech: Coarticulation and inflectional morphology

Margaret Cychosz[a]

[a]University of California, Berkeley, Department of Linguistics, 1203 Dwinelle Hall, Berkeley, USA, mcychosz@berkeley.edu

Abstract

Evidence from acoustic and articulatory phonetics increasingly suggests that morphological structure is reflected in spoken language patterns. For child language, this interaction between word structure and speech production has the potential to shed considerable light on the status of children's early word forms - but the topic remains underexplored in child speech. How do children analyze the internal structure of morphologically complex words throughout childhood? To answer this, the current study measured the coarticulation patterns of bilingual Quechua-Spanish children (5-10 years) and adults. Coarticulation was measured acoustically in two word environments, within morphemes and across morpheme boundaries. Both child and adult participants distinguished coarticulatorily between the word environments, but they did so in different ways. Children differentiated between environments via a combination of duration and coarticulation while adults consistently coarticulated more in shorter duration sequences. Additionally, the children's speech patterns, but not the adults', were sensitive to prosodic structure: children produced increasingly shorter phones in words with more syllables. It was suggested that the difference between adults and children could be attributable to adults' faster speaking rate and increased dominance in Quechua. Future work is needed to determine if young children's speech patterns reflect prosodic planning, morphological planning, or both.

Keywords: coarticulation; morphophonetics; field phonetics; first language acquisition; morphology; Quechua

1 Introduction

Adult speakers can readily compose novel, morphologically complex word forms like *daxy* or *unraxlike* from morphemic subcomponents. Evidence of u-shaped learning curves in children, and the results of Berko (1958)'s Wug Test, have led developmental researchers to recognize that young children must share in adults' morphological productivity. If not, children would not be able to extend morphophonological patterns, such as the correct plural allomorphs, to novel lexical environments. Yet there remains disagreement concerning children's early linguistic representations: researchers disagree about when children begin to decompose words that are complex - at least for adults - into their component morphemes. Children's lexical forms could be fully adult-like from early toddlerhood (e.g. Wexler 1998) the forms could emerge over time with usage (e.g. Ambridge et al. 2015), or there could be some combination of the two approaches (e.g. Swingley & Aslin 2002).

There are factors, beyond the ability to flesh out a morphological paradigm with novel forms, that reveal whether children have adult-like production mechanisms for complex words. One window into the nature of lexical access and storage - at least in adult speakers - is to examine the effects of morphological structure on speech production. Adults, and even children, have been shown to produce complex words differently, arguably as a result of planning differences, or maybe practice (Cho 2001; Hay 2003; Lee-Kim et al. 2013; Plag 2014; Song, Demuth, Evans, & Shattuck-Hufnagel 2013; Song, Demuth, Shattuck-Hufnagel, & Ménard 2013; Strycharczuk 2019; Sugahara & Turk 2009; Tomaschek et al. *under review*). For example, in adult speech, morphologically complex words like *sighed* can be longer in duration than their simplex counterparts like *side* (Sugahara & Turk 2009). And children have been shown to coarticulate more between adjacent segments within mono-morphemic words like *box* than between otherwise identical segments that straddle a morpheme boundary as in *rocks* (Song, Demuth, Shattuck-Hufnagel, & Ménard 2013).

The complex relationship between speech variability and morphological structure in children is still relatively underexplored (cf. Redford 2018; Song, Demuth, Shattuck-Hufnagel, & Ménard 2013; Song, Demuth, Evans, & Shattuck-Hufnagel 2013). Yet measuring how morphological structure is reflected in children's speech could allow us to infer about the status of children's early complex word forms. Are words that, to an adult are transparently morphologically complex, treated as such by children? And when a child does recognize that a given word is morphologically complex, how does this recognition affect children's lexical storage, access, and production?

To that end, the goal of this work is to evaluate how the composition of morphologically complex words is reflected in children's speech production patterns in Quechua, a highly agglutinating language with over 200 unique, productive verbal and nominal suffixes. In doing so, this work follows a growing line of research evaluating relationships between morphology and speech production (e.g. Cho 2001; Hay 2003; Lee-Kim et al. 2013; Plag 2014; Song, Demuth, Evans, & Shattuck-Hufnagel 2013; Song, Demuth, Shattuck-Hufnagel, & Ménard 2013; Strycharczuk 2019; Tomaschek et al. *under review*). Here, coarticulation is measured across and within morpheme boundaries in adults and a cross-sectional cohort of school-aged children to examine if, and how, speakers distinguish coarticulatorily between morphological environments. It is anticipated that children, who have less experience with language and potentially less abstract linguistic categories, will coarticulate similarly between and within morpheme boundaries. This behavior would suggest that the children are not as likely as adults to analyze the internal structure of complex word forms. Adults, however, will coarticulate more within morphemes than at a morpheme boundary if adults analyze the internal structure of complex words and compose these words together from their component morphemic parts.

The choice to study these patterns by word environment in a highly agglutinating language like Quechua is an important one. Quechua may make an important typological contribution to this literature because, given its morphological complexity, there are few doubts as to children's productivity from a very young age (Courtney & Saville-Troike 2002). Furthermore, the sheer number of available inflectional forms in Quechua, compared to more analytic languages such as English, likely translates into less competition between forms accessed in their inflected form and those composed online from individual morphemes during word retrieval (Hay 2003; Pinker & Ullman 2002). In short, evidence that children are not analyzing the internal structure of complex words would be especially noteworthy in a language of this structure.

2 Background

2.1 Accessing complex words

Adult speakers' ability to compose novel, morphologically complex words (e.g. *unraxlike*) appears to be highly flexible and abstract. Children's early overgeneralization errors (e.g. *I taked**) demonstrate how even very young speakers analyze the internal structure of words and apply morphemes in new, unheard environments. Consequently, morphological productivity is a defining characteristic of the language faculty after a certain point in development.

Still, there is some behavioral evidence that adult speakers do not uniformly decompose complex words (Taft & Forster 1976). Using lexical frequency statistics, some work suggests that adults may instead access complex words, especially high-frequency complex words, holistically in the mental lexicon, without decomposing them (Baayen 1992; Baayen et al. 2003). For example, Colé et al. (1997) tested this idea and demonstrated a processing advantage for high-frequency base stems (e.g. *fish*) versus derived words (e.g. *fisher*). Crucially, this advantage diminished when the derived word was more frequent than the base stem.

Hay (2003) likewise demonstrated relative frequency effects. Measuring speech production (contrastive pitch accent placement and phonetic reduction at morpheme boundaries), Hay showed that when derived words are more frequent than their corresponding base form (e.g. *disentangle* vs. *entangle*), these words may be accessed holistically (see Pluymaekers et al. 2010 for an alternative explanation). Similarly, greater root frequency relative to a suffix is usually predictive of decomposition, at least in more analytic languages (R. Smith et al. 2012). And less frequent words, or words that only occur with a given suffix, do not necessarily manifest complete morphological decomposition (Kemps et al. 2005). Hay (2003)'s findings in particular imply a dual-route model of complex word access, where two different lexical mechanisms - holistic storage and online complex word formation - compete during complex word production (Baayen 1992; Koenig & Jurafsky 1995; Pinker & Prince 1994; Pinker & Ullman 2002). The method of construction that is fastest, predicted by the relative frequency of base to derived form, wins.

Finally, a number of computational models have attempted to predict decomposition versus whole-word storage. In a model of English morphology, Plag & Baayen (2009) demonstrated that affixes that are highly parsable - affixes that speakers can more easily separate from their root (e.g. English -*ness*) - predict decomposition. Affixes that are highly fused to the stem - those that speakers are less able to parse from the stem (e.g. English -*th*) - predict whole-word access. And more recently, O'Donnell (2015) proposed a probabilistic model where speakers weigh the productiveness of rules versus the storage of idiosyncratic, complex words to predict decomposition.

Given children's u-shaped learning curves, and their performance on the Wug test (Berko 1958), it is often assumed that children as young as four years share adults' morphological productivity capacities. Consequently, the above findings that cast doubt upon whether composition is always the online strategy that wins during lexical access for *adults* are highly relevant to the study of how children access morphologically complex words. Frequency ratios appear to predict word (de)composition in adults (e.g. Hay 2003). Thus, since frequency ratios change as children learn more words, stems, and suffixes, one can predict that the fastest manner for children to access a complex word - by composition or holistic access - will also change over the course of development. But lexical frequency ratios are just one method to evaluate complex word access. Another method may be to evaluate how children, particularly a cross-sectional cohort of children spanning several years of development, access complex words throughout development. This is the objective of the current study.

2.2 Morphological structure and speech production

One way to study how speakers analyze complex words, and access them in the mental lexicon, is to measure the effect of morphological structure on speech production. Morphophonetics, or how morphological structure manifests in fine, phonetic variability, is increasingly studied in adult speech (see Plag (2014) and Strycharczuk (2019) for recent overviews). Researchers now know that morphological structure and composition dictates some speech variability in adults (Cho 2001; Guy

1991; Hay 2003; Lee-Kim et al. 2013; Plag et al. 2017; R. Smith et al. 2012; Sugahara & Turk 2009; Tomaschek et al. *under review*; cf. Hanique & Ernestus 2012), but it is often unclear when, how much, or even why (Mousikou et al. 2015; Plag 2014).

Adult speech. One of the earliest studies to examine the relationship between phonetics and word structure was Losiewicz (1995) who found that frequency affected the phonetic realization of morphemes. Specifically, English past tense -*ed* was temporally longer on low-frequency verbs than high-frequency verbs.[1]. Elsewhere, Sugahara & Turk (2009) found that heteromorphemic words were longer than otherwise identical monomorphemic words (e.g. *sighed* versus *side*), while R. Smith et al. (2012) found that pseudo-prefixes (*mistake*) were shorter in duration than real prefixes (*mis-judge*). Finally, Seyfarth et al. (2018) examined paradigmatic influences on segmental duration in English, for example predicting differences between *frees* and *freeze* based on membership in the *free* paradigm. After controlling for lexical frequency, prosodic structure, and orthography, the authors found that paradigm membership can sometimes predict segment duration, though this depends upon the segment studied (fricatives versus stops).

Articulatory methods have likewise been used to evaluate the relationship between morphology and phonetic variation in adult speech. Cho (2001) found that intergestural timing in Korean, measured with electromagnetic articulagraphy and electropalatography, was more stable within a word than across a morpheme boundary. The author found a similar effect for a lexicalized compound word versus a non-lexicalized compound word (noun phrase). In addition, English /l/ darkness has been found to vary by position at morphological boundaries (Lee-Kim et al. 2013). The phone /l/ is darker in stem-final position (*coolest*) than affix–initial (*coupless*), independent of duration, because *coolest* is analogizing to word-final /l/ in its base form *cool* (See also Strycharczuk & Scobbie 2016). Most recently, Tomaschek et al. (*under review*) measured anticipatory coarticulation patterns in English. The authors used electromagnetic articulagraphy to evaluate the degree of morphological opacity in inflected verbs. Results showed that English speakers exhibited greater anticipatory coarticulation of the vowel in a verb stem (e.g. *clean*) for those verb inflections with which they had increased experience and practice.

However, the relationship between speech production and morphology is not always straightforward. For example, Plag et al. (2017) found that, after controlling for a host of other factors such as number of syllables, speaking rate, and surrounding context, the duration of English /s/ and /z/ varied systematically by morphological status. But the morphemic plural and non-morphemic /s, z/ were longer in duration than morphemic /s, z/ used as third person markers. The cause of this difference is not clear, though some authors suggest that the spontaneous speech analyzed in Plag et al. (2017) may explain the findings (Seyfarth et al. 2018). Elsewhere, others have argued that it is not morphological structure but word information load (Hanique & Ernestus 2012; Pluymaekers et al. 2010) or contextual predictability (Cohen 2014) that explains these interactions between speech production and word structure.

A complete review on the consequences of morphological structure for adult speech production is beyond the scope of a paper on children's speech development. What is important, however, is that the previous two decades of morphophonetics research has converged somewhat on the fact that words with productive morphology are processed and produced differently than simplex words (Kemps et al. 2005; Plag 2014; Tomaschek et al. *under review*). For child speech, these findings present a new method to detect when and whether children are decomposing words. Furthermore, the current work extends the study of morphophonetics to Quechua. This could be an important typological contribution because Quechua's morphology is highly productive, more so than in other more analytic languages that tend to be the object of study. As a result, children's productivity from a young age is not in doubt (Courtney & Saville-Troike 2002). And, furthermore, Quechua's agglutinating structure means that there should be less relative frequency effects between forms composed online and those accessed in the inflected form (Baayen 1992; Hay 2003; Pinker & Ullman 2002).

Consequently, the goal of this paper is twofold. First, it extends the acknowledged relationship between speech production and word structure from adults to a cross-sectional sample of children. Second, it examines production and word structure in a language with a vastly different morphological structure from the western European languages that have been the focus of previous research.

Child speech. While we understand some of the explanatory mechanisms behind morphological structure and phonetic variation in adult speech, studies have rarely examined this relationship in children (cf. Song, Demuth, Evans, & Shattuck-Hufnagel 2013; Song, Demuth, Shattuck-Hufnagel, & Ménard 2013). For the current study, examining child morphophonetics also contributes to a line of research evaluating the status of children's early word forms: some theorists posit that children's lexical forms are fully adult-like from early toddlerhood (e.g. Wexler 1998), while others posit that they emerge over time with usage (e.g. Ambridge et al. 2015), or a combination of the two approches (e.g. Swingley & Aslin 2002).

A handful of studies have used morphophonetics to address the question of how children represent complex words. Song, Demuth, Shattuck-Hufnagel, & Ménard (2013) studied morphophonetic development in English-learning children. They employed acoustic and ultrasound analyses to study adult and 2;0 (year;month) children's CC syllable (/ks/) coarticulation in the coda of bimorphemic (*rocks*) and monomorphemic (*box*) (These, along with *rock* and the nonword *das*, were the items tested.). The articulatory data showed that children reliably raised their tongue more for *rocks* than *box*, suggesting that they could distinguish between morphemic and non-morphemic coda consonant clusters. While there was no evidence of anticipatory coarticulation for *box* in the adults or children studied, there was a strong perseveratory coarticulation effect of /k/ on /s/ in *box* and anticipatory coarticulation effect of /s/ on /k/ in *rocks*. This finding led the authors to conclude that the primary articulatory target for monomorphemic *box* was the C_1 of the consonant cluster (/ks/), while the articulatory target for bimorphemic *rocks* was C_2 or the plural morpheme /s/. The articulatory target differs by morphological role, suggesting that semantic information may be encoded in articulatory gestures. This applied for both adults and children.

Elsewhere, coda position American English fricatives /s, z/ were studied in the naturalistic speech of children and their caretakers (Song, Demuth, Evans, & Shattuck-Hufnagel 2013). Children's morphemic /z/ was longer in duration than non-morphemic /z/ in word-final position. This acoustic evidence suggests that, as early as two years of age, child speakers of American English reliably distinguish between otherwise identical morphemic and non-morphemic segments. The authors used this finding to argue that even at this young age children may not rote-memorize morphologically complex forms, but may instead compose the words online.

Finally, Redford (2018) studied the relationship between word structure and phonetic variation across word boundaries in prosodic words (e.g. *the bat*). The results showed that schwas in children's *the* productions were reliably louder and longer than in adults' productions. According to Redford, this result was strong evidence that children exhibit immature timing control, but the results were inconclusive concerning children's speech organization - whether children accessed the prosodic words holistically or in parts.

These developmental studies aside, the relationship between speech production and morphological structure has been almost entirely unexplored in children. This is to the detriment of the field because a parallel line of research on children's coarticulation has shed considerable light on early phonological representations. Conflicting evidence from this research suggests that children may store speech and word forms in larger chunks than adults, at the syllabic level or higher (e.g. Nittrouer et al. 1989; Noiray, Wieling, et al. 2019; Zharkova et al. 2011). The following section summarizes this work on child coarticulation.

2.3 Children's phonological storage: Evidence from coarticulation

Numerous works have found that children coarticulate more than adults, often until early puberty, with the degree and variability of coarticulation decreasing with age (Goodell & Studdert-Kennedy 1992; Nittrouer et al. 1989; Nittrouer et al. 1996; Noiray, Wieling, et al. 2019; Zharkova et al. 2014; Zharkova et al. 2018 *inter alia*). Many authors have interpreted this relationship between coarticulation and age as evidence that young children represent speech in syllable-sized units, which gradually individuate into phoneme-sized units over the course of development.

While one may expect children to coarticulate less than adults – children speak slower and with less coordinated movement (**?**) – many studies have concluded that children coarticulate *more* than adults (Goodell & Studdert-Kennedy 1992; Nittrouer et al. 1989; Nittrouer et al. 1996; Zharkova et al. 2011). Citing this evidence, some have argued that the phenomenon of "coarticulation" in child speech reflects children's phonological representations (this will be outlined in further detail in the following section). This approach thus argues that unlike adult coarticulation, child coarticulation does not necessarily reflect efficiency in speech planning or the clear anticipation of upcoming segments (Bradlow 2002; Whalen 1990), but instead reflects a larger representational unit.

Studying children's coarticulation, Nittrouer et al. (1989) concluded that children may have more holistic, syllable-sized representations. The authors measured coarticulation in fricative-vowel sequences within nonce words to test the anticipatory effect of vowels on /s/ and /ʃ/. Their results from children 3;0-8;0 showed that children coarticulated more than adults – vowels affected the children's fricative production more than the adults' production. The authors concluded that this pattern manifested in the fricative-vowel sequences because the children did not distinguish between /s/ and /ʃ/ as reliably as the adults. It was not, the authors argued, because children differed from adults in their degree of anticipatory lip rounding or even the lingual constriction shape (see also Nittrouer et al. 1996).

Nittrouer et al. (1989)'s findings partially supported conclusions from the two children studied in Repp (1986). There, the younger child (4;8) showed greater anticipatory lip rounding for /s/ before /u/ than the older child (9;5). More recently, the coarticulation patterns of children with and without apraxia of speech and adults without apraxia were studied (Nijland et al. 2002). Participants produced nonce sequences of ə-V-C where V was /a, ɪ, or u/ and C was /s, x, b, or d/. The typically-developing children once again showed more intra- and inter-syllabic anticipatory coarticulation than the adults.

Most recently, Zharkova and colleagues (Zharkova et al. 2011; Zharkova et al. 2014; Zharkova et al. 2018) and Noiray and colleagues (Noiray et al. 2018; Noiray, Wieling, et al. 2019; Rubertus & Noiray 2018) incorporated articulatory ultrasound data and have frequently observed that children coarticulate more than adults. The authors frequently infer that children's propensity to coarticulate reflects their larger representational units. In Zharkova et al. (2011), children aged (6;3-9;9) altered their productions of /ʃ/ based on the following vowel (/a, i, or u/) more than adults. However, Zharkova et al. (2014) conducted a similar experiment with older children (10;0-12;4) and did not find any differences between child and adult coarticulation patterns. The authors suggest that children may approximate adult-like phonological representations by preadolescence.

Also using ultrasound imaging, Noiray, Wieling, et al. (2019) measured vocalic anticipation in children (3;05-7;06) and adults. They found that in V#CV strings, children anticipated the upcoming vowel sooner than adults at age 3;0 and 5;0, but progressively less by 7;0. (See Noiray et al. (2018) for segment-specific explanations for coarticulation development.) Most recently, Noiray, Popescu, et al. (2019) correlated children's (4;06-7;02) coarticulation, measured articulatorily with ultrasound imaging, with measures of expressive vocabulary and phonological awareness. The amount of children's intrasyllabic coarticulation and gestural individuation was negatively correlated with phonological awareness (a measure of speaker awareness of segmental units), which the authors suggest reflects a lack of segmental individuation in children.

Authors of the above acoustic and articulatory studies often argue that children's coarticulation patterns reflect holistic, syllable-sized representations. According to such an interpretation, representations progressively individuate and become more adult-like as children age. It is nevertheless important to note that results in the child coarticulation literature are mixed. Numerous studies have instead found that adults coarticulate more than children. These works often argue that children's phonological representations are just as or more discretely organized into individual segments than adult phonology. As a result, children learn to coordinate and exhibit appropriate coarticulatory overlap as part of standard phonetic/phonological development (Barbier et al. 2013; Barbier et al. 2015; Katz et al. 1991; Kent 1983; Whiteside & Hodgson 2000). Other works have found no differences in coarticulatory patterns between adults and children (Flege 1988; Goffman et al. 2008; Noiray et al. 2013; Sereno & Lieberman 1987; Sereno et al. 1987).

Methodological differences between the studies summarized above may account for some of the different conclusions concerning children's coarticulation patterns. There are innumerable ways to measure coarticulation - even in child populations where data collection techniques are often more limited. There are acoustic techniques (e.g. Nittrouer et al. 1996) and articulatory approaches (e.g. Zharkova et al. 2014). But even within the acoustic realm there are different techniques. Most frequently, for fricatives, coarticulation is quantified as the change in centroid (average) based on vocalic context while coarticulation within vowels is typically quantified as the midpoint or steady-state formant value (Nittrouer et al. 1989; Nittrouer et al. 1996). Elsewhere, coarticulation has been measured acoustically as the spectral change between adjacent segments (Gerosa et al. 2006).

The current study employs one of these measures of spectral change, outlined in Gerosa et al. (2006) and validated in Cychosz et al. (2019). This measure takes into account change occurring over the timecourse of segments, instead of static measurements. Children speak slower than adults, so taking measurements from the steady-state or midpoint of a vowel is not ideal. Children's slower speaking rate permits them greater opportunity to achieve a steady acoustic signature. Measuring coarticulation from a larger portion of the overall segment, which is what is proposed here, mitigates the effect of speaking rate to a certain degree.

3 Current study

3.1 The language

The current study measures coarticulation in morphologically complex words in South Bolivian Quechua, henceforth Quechua, a Quechua-II/C language with over 1.6 million speakers in southwest Bolivia and northwest Argentina (Torero 1964). This variety of Quechua is spoken in the Chuquisaca department of southern Bolivia. Like many Quechuan varieties, Quechua in southern Bolivia has been in intense contact with Spanish for hundreds of years resulting in large amounts of language mixing and borrowing (Muysken 2012a; Muysken 2012b). Furthermore, public schools in Bolivia are conducted primarily in Spanish. Consequently, almost all school-age children who speak Quechua in the home, such as the children in the present study, are bilingual in Spanish. And because of Spanish schooling, children in Bolivia learn to read in Spanish. Though Quechua has an established writing system, children generally do not learn to read or write in Quechua and few Quechua speakers use the language in its written form.

Quechua is a highly-agglutinating language with over 200 productive nominal and verbal suffixes that encode argument structure and grammatical relations (for comparison, English has around 35 productive suffixes). The morphology is nevertheless highly regular and not subject to significant morphophonological processes or fusion. The phonological inventory includes three phonemic vowels /i, a, u/ and two allophonic vowels derived in uvular contexts [e, o] (Gallagher 2016). The consonantal inventory contrasts voiceless stops, aspirated stops, and ejectives along four places of articulation, /p, t, k, q/, as well as a three-way alveopalatal stop-aspirated stop-ejective distinction /ʧ, ʧʰ, ʧ'/. Nasals are contrasted along three places of articulation, /m, n, ɲ/ with an allophonic velar nasal. See Appendix A for a complete consonant inventory.

3.2 Research objectives

The primary objective of this chapter is to measure how Quechua word structure affects speakers' production of morphologically complex words. To accomplish this, a relatively novel acoustic measure of coarticulation is employed, which has been validated for young children's voices and a variety of consonants (Cychosz et al. 2019). Using this measure, the coarticulation patterns in adult and child Quechua speakers are evaluated to see if, and how, the patterns differ between children and adults, and across development, in a cross-sectional sample of school-aged children.

As an agglutinating language, Quechua provides unique insight into interactions of morphology and phonetics. Quechua speakers appear to have highly flexible inflectional and derivational lexicons: suffixes and roots are abstracted away from the original lexical contexts and are easily rearranged for novel stem+suffix pairings. This process is similar to how speakers of more analytic languages, such as English, effortlessly arrange novel noun-adjective pairings. It is this morphological structure that makes Quechua an interesting typological contribution to morphophonetics. For one thing, as noted in the previous sections, given the number of possible complex word forms, is difficult to doubt the the morphological productivity of young Quechua speakers. Quechua is understudied, both in child language research and linguistics more broadly. Still, much like well-studied, less synthetic languages, there is undeniable evidence of morphological productivity in children's Quechua (and other highly agglutinating languages). Child speakers of Cuzco Quechua appear to have a productive morphology by age 5;0, if not earlier (Courtney & Saville-Troike 2002).[2]

Another reason why Quechua makes an important typological contribution again concerns its morphological structure. In Quechua, frequency ratios between words that might be compiled online and those that might be accessed holistically are likely to be much smaller. This also means that the type frequency of each word is lower than in moderately synthetic languages, like English and Dutch, which are more commonly studied in morphophonetics. So concerns about a race between online compilation and holistic access may be less relevant.

One final reason why Quechua may be an interesting addition to the morphophonetics literature concerns its applied use in Bolivia. As described in section 3.1, very few Quechua speakers read or write in the language (speakers who attend school instead become literate in Spanish). Disengaging morphological effects from orthographic effects has been immensely difficult for morphophonetics, and the study of phonetics in general (Seyfarth et al. 2018; Warner et al. 2006). Quechua removes this problematic variable because speakers do not have strong orthographic representations.[3] Thus, while the primary typological interest of Quechua concerns its morphological productivity and what that means for the structure of the lexicon, there are additional, applied considerations that make Quechua a unique language to examine the effects of word structure on speech production.

Given these characteristics of Quechua, this study makes two predictions. First, relying on morphophonetic studies on adults and coarticulation studies on children, it predicts that children will produce complex words differently from adults. The hypothesis is tested in a tightly-controlled experimental paradigm where coarticulation between the phones [a] and [p], or [a] and [m], is measured. When [ap] or [am] straddle a morpheme boundary (e.g. sunkh**a-p**i 'beard-LOC'), the prediction is that adults, and potentially older children, will coarticulate less between the phones than when [ap] or [am] fall inside of a root morpheme (e.g. p**ap**a 'potato'). The second prediction of this study is that children will coarticulate more than adults at morpheme boundaries meaning that overall, adults will show a larger differential between within- and across-morpheme coarticulation than children.

If older speakers' phonetic realization of the exact same [ap] or [am] sequence differs within a morpheme versus across a boundary, then we can infer that the adults are breaking words down at constituent boundaries: they have learned to parse speech segmentally and store it morphophonemically. The relative lack of distinction in the children's speech, however, infers that children may be accessing complex words in larger, inflected chunks whose internal structure the children may not have analyzed.

The predicted outcomes of this experiment do not suggest that Quechua-speaking children are morphologically unproductive. Quechua is understudied, both in child language research and linguistics more broadly. Still, much like well-studied, less synthetic languages, such as English and French, there is undeniable evidence of morphological productivity in children's Quechua (and other highly agglutinating languages). Child speakers of Cuzco Quechua appear to have a productive morphology by age 5;0, if not earlier (Courtney & Saville-Troike 2002).[4]

Instead of suggesting a lack of productivity, the predictions in experiment two reflect recent proposals in child speech and language research that children may initially analyze language differently than adults and, as a result, children may store language more holistically (Davis & Redford 2019; Lieven et al. 1997; Redford 2019). Such proposals argue that children do not always recognize complex forms as having an internal structure, and thus do not deconstruct words into morphemes, or morphemes into phones. Instead, children may be more prone to represent and access language in chunks, particularly frequently-occurring combinations or phonotactically-probable "subwords." You could imagine that chunks in English, like *Immagonna* or *lookatit*, could be stored in this way, in addition to their deconstructed forms (individual words, morphemes, and phonemes). Thus, child language may be redundant, with representation at multiple levels in the grammar.

This approach to speech and language development may surprise some audiences – after all the birth of modern linguistics marked a decided turn away from memorized, holistic chunks (Chomsky 1959; Skinner 1957). This is because Behaviorism made untenable demands on psycholinguistic representation and instigated logical fallacies such as the Poverty of the Stimulus. However, abstraction is not undone by the "chunking" approach to language acquisition outlined here. In a chunking approach, children learn language chunks such as words and syllables which then individuate into smaller linguistic units with time. This chunking approach could postulate that abstraction is postponed until later in development: abstraction of morphemes and phonemes could emerge piecemeal via diffusion in the child's lexicon and grammar. Alternatively, a chunking approach could argue that children (and adults) have redundant representations at multiple levels (lexical, syllabic, phonemic) (Arnon & Christiansen 2017; Arnon & Cohen Priva 2013; Arnon 2010; Bannard & Matthews 2008). Some recent exemplar-based models of adult phonology propose exactly this organization, where both speaker-specific and abstracted categories co-exist with the latter emerging out of the former (Pierrehumbert 2016).

So the proposal that children do not reliably analyze the internal structure of word forms, thus storing some forms more holistically, does not claim that children are not morphologically productive language users. It also does not claim that speakers do not develop abstract language. Finally, a concept of more holistic storage in childhood does not argue that speakers must store every experienced multisyllabic/multiword sequence in perpetuity. More holistic storage *does* argue that linguistic representations are probabilistic, emerge with experience, and exist at multiple levels in the grammar. If we ignore biases like literacy and linguistic concepts like words and phonemes, it actually makes sense that young children, who are not exposed to written language and have limited experience with spoken language, might analyze and store words differently than adults. This work empirically tests this theory.

3.3 Hypotheses

The primary objective of this experiment is to compare adult and child Quechua speakers' coarticulatory patterns within and between morpheme boundaries.

1. Do adult Quechua speakers coarticulate more between biphone sequences within morpheme boundaries (e.g. p**ap**a 'potato') versus across morpheme boundaries (e.g. pap**a-p**i 'potato-

LOC')?

It is predicted that adult speakers will coarticulate more between the same biphone sequence *within* morphemes than *across* morphemes. This speech pattern between morphological environments may reflect the decomposition of complex words into component parts, as previous work on both adult (e.g. Cho 2001; Tomaschek et al. *under review*) and child coarticulation (e.g. Song, Demuth, Shattuck-Hufnagel, & Ménard 2013) has demonstrated.

2. Do child speakers of Quechua, aged 5;0-10;0, differentiate their coarticulatory patterns across versus within morpheme boundaries? If so, does this change throughout development in this cross-sectional sample?

On the other hand, it is predicted that child Quechua speakers will not necessarily differentiate their coarticulatory patterns between the two morphological environments. Instead, children may coarticulate equally within and across morphemes. This would suggest, as previous work on child speech has suggested (e.g. Noiray, Popescu, et al. 2019; Redford 2018; Zharkova et al. 2011), that increased coarticulation reflects more holistic representations. Specifically, the children in the current study may coarticulate similarly in the two morphological environments. If this occurs, it may indicate that they are not always breaking down morphologically complex words in the same manner as adults, but are instead storing items more holistically (Redford 2018).

In addition to measuring coarticulation in the two morphological environments, the duration of the biphone sequences will also be measured. There are acknowledged interactions between duration and coarticulation; for example, speakers tend to coarticulate more when they speak faster (Gay 1981; Matthies et al. 2001). Thus, measuring how coarticulation interacts with speaking rate could be an important component to the speech patterns evaluated here. Furthermore, given that adults speak faster than children, it may be especially important to measure, and control for, the duration of the biphone sequences when measuring differences in coarticulation between different age groups.

4 Methods

4.1 Participants

Fifty-one children, aged 5;0-10;11, and ten female adults (adult μ_{age}=23, σ=5.46, three did not report age) participated in this study. Children's distribution by age was as follows: 10 five-year-olds (μ=5;7, σ=0;4; 6 girls, 4 boys, one did not report), 10 six-year-olds (μ=6;5, σ=0;2; 5 girls, 8 boys), 13 seven-year-olds (μ=7;7, σ=0;4; 3 girls, 8 boys), 8 eight-year-olds (μ=8;8, σ=0;2; 4 girls, 5 boys, one did not report), 5 nine-year-olds (μ=9;4, σ=0;3; 2 girls, 3 boys, two did not report), and 5 ten-year-olds (μ=10;8, σ=0;4; 1 girl, 4 boys, two did not report). The recording for one of the adult participants contained significant wind interference and was removed from analysis leaving n=9 adult participants in the final sample. All participants were bilingual Spanish-Quechua speakers living in or around a mid-size town in southern Bolivia. The child participants were either recruited at a local primary school where the author was volunteering (n=13), or through personal contacts in the surrounding communities (n=38). The adult participants were recruited through local contacts. This research was approved by the UC Berkeley Institutional Review Board.

Most children had typical speech and hearing development, per parental/teacher self-report. The caregivers of n=3 children (2 seven-year-olds, 1 five-year-old) stated that their child was late to begin talking.[5] Note that these communities are medically under-served so some language delays/impairments may go unreported. Additionally, n=3 children had lost one or more of their top/bottom front teeth at the time of recording.[6] An attempt to conduct a hearing screening for the children was made. However, it became clear after attempting with a few of the children during pilot testing that false positives were being collected during the test as some children were nervous/afraid of making a mistake. Consequently, it cannot be said with absolute confidence that all children would have passed a standard hearing screening. The adult participants did not report any speech or language disorders.

Socioeconomic status (SES), usually implemented as mother's level of education in child development research, is an important predictor of child language development in the United States (Hoff 2003; Pace et al. 2017). However, it is not clear if SES is predictive of language outcomes in all cultural or linguistic contexts. Specifically, it is unknown if SES predicts language development in Bolivia as a whole, in these speech communities specifically, or for children learning Quechua. Still, SES information is reported here as it is an important predictor in many other cultural contexts.

Information on the central caregiver's education level (usually the mother, occasionally the grandmother) was collected from those families recruited from the surrounding community, but not those recruited at the school. There is no a priori reason to believe that the distribution of socioeconomic strata of the children recruited at the school would differ from those who were recruited from elsewhere in the community. That is to say, the children from the surrounding communities likewise attended school, just in a different location from where the school children were recruited and tested. There were 7 sibling pairs (no twins), and 1 three-sibling pair, in the child sample resulting in 43 unique caregivers in the sample. For the 31 caregivers of children recruited from the surrounding community, the caregivers' education levels were: n=17 caregivers (59% of caregivers from the community) completed some primary school (less than six years of education), n=5 (17%) completed primary school (6 years of education), n=3 (10%) completed the equivalent of a middle school (10 years of education), n=3 (10%) completed secondary/high school (13 years of education) and n=2 (7%) had not received any formal schooling. Two caregivers did not report.

An additional indicator of SES in these indigenous communities in Bolivia may be the central caregiver's familiarity with Spanish, which indicates that the caregiver had the opportunity to attend school (conducted in Spanish) for a longer period of time. A coarse estimation of the central caregiver's level of Spanish-Quechua bilingualism was also collected from those families recruited from the surrounding community: n=6 (21% of caregivers from the community) were monolingual Quechua speakers, n=5 (17%) were Quechua-dominant but spoke/understood some Spanish, n=17 (59%) were bilingual Quechua-Spanish speakers, and n=1 did not report.

4.2 Tasks

The child participants completed four tasks, all prompted with pictures, in the following order: 1) real word repetition, including a morphological extension component (to be explained in the following sections), 2) Quechua nonword repetition, 3) Spanish nonword repetition, and 4) additional real word repetition with morphological extension. For all tasks, children repeated the real words or nonwords after a pre-recorded model speaker. Nonword repetition tasks are not further discussed.

The adult participants only completed the two real word repetition tasks because even the eldest children approached ceiling on the nonword repetition tasks. Because much of the children's data was collected during the school day, the entire testing procedure had to be relatively short and executable in one sitting. The entire experimental procedure took approximately 30-40 minutes per child and 20 minutes per adult. For their participation, all children could choose an item from a toy bag. Children at the school additionally received academic assistance and lessons on English and Spanish from the author who was volunteering at the school. Additional donations such as school supplies and books were made to the school. The adult participants and caregivers of children from the surrounding communities who did not attend the school instead received a small monetary sum.

4.3 Stimuli

The real word repetition tasks consisted of 56 high-frequency Quechua nouns (plus 6 training trials for 62 total lexical items) that are familiar to children learning Spanish and Quechua in southern Bolivia (stimuli listed in Appendix B). Neither Bolivian Spanish nor any Quechuan language has an equivalent to the *Macarthur Bates Communicative Development Inventory* (Fenson et al. 2007), which reports stages of age-normed vocabulary development. Nor do these languages have any large, transcribed child-directed speech corpus from which to infer vocabulary development. For these reasons, children's knowledge of the test items was confirmed via a pre-test that demonstrated that children as young as 3;0 recognized all items. Female caregivers also confirmed that children as young as 3;0 should recognize the items.

In addition to selecting high-frequency lexical items, likely to be recognized by the children and easily represented in a photo, these particular lexical stimuli were also chosen because they contained the sequence [ap] or [am] within a morpheme (e.g. p**ap**a 'potato') or crossing a morpheme boundary (e.g. thap**a-p**i 'prairie-LOC').[7] And to control for acoustic correlates of stress, the phone [a] in the biphone sequence had to fall in the syllable carrying primary stress.[8] Thus, of the original 56 real word test items, this study measured coarticulation on a subset that contained the biphone sequence [ap] (n=23) or [am] (n=23) (see Table 1 for list of stimuli that elicited [ap] and Appendix C for the stimuli that elicited [am]).[9] The experimental hypotheses remain the same for both the [ap] and [am] biphone sequences. However, given the acoustic measure of coarticulation employed here - spectral distance - overall there is less "coarticulation" between the phones in [ap] than the phones in [am] due to the increased acoustic similarity of the phones in [am] (voiced, sonorous).

The VC sequences [ap] and [am] were chosen to examine coarticulatory effects for several important reasons. First, instead of CV sequences, VC sequences were chosen because all Quechua nominal case-marking suffixes are consonant-initial (e.g. '-*q*' GENITIVE, '-*manta*' ABLATIVE). Consequently, it is not possible to elicit a CV sequence that crosses a noun-case marker boundary in Quechua. Then, of the possible VC sequences, [ap] and [am] were chosen because 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.[10]

The two suffixes -*pi* and -*man* (pronounced [maŋ]) were also chosen for several specific reasons. First, nominal case markers were chosen because nouns are easier to represent in picture prompts than derived word forms (e.g. puñu-y 'to sleep' puñu-chi-y 'to make (one) sleep') or verb conjugations. Second, nouns are grammatical in Quechua with just one suffix. Some conjugated verbs require multiple suffixes (see previous 'sleep' example), which would make elicitation and tight control of the experimental stimuli more difficult. By using nominal suffixes, elicitation could be isolated to a single stem+suffix combination. Finally, the locative and allative markers were used because, absent a large corpus of child-directed Quechua speech, it is reasonable to assume that the locative -*pi* and allative -*man* on high-frequency nouns, such as those elicited, will be relatively frequent in a child's input.

Given all of these considerations - the need for high-frequency, child-friendly words, the correct prosodic environment, frequent suffixes, and segments that were easily segmented - it was challenging to identify lexical items for use in the experiment. Still, with (n=35) unique items in the across morpheme boundary condition and (n=11) unique items in the within boundary condition, this study uses more distinct lexical items than most previous studies of morphological effects on speech production in children or adults (Lee-Kim et al. 2013; Song, Demuth, Shattuck-Hufnagel, & Ménard 2013).

Table 1

Real word repetition stimuli to elicit [ap]

Real word*	Translation	Morpheme environment†
chi'**ta-p**i	'sheep-LOC'	across
cu'c**a-p**i	'coca (leaves)-LOC'	across
hatunma'm**a-p**i	'grandma-LOC'	across
imi'll**a-p**i	'girl-LOC'	across
juk'u'ch**a-p**i	'mouse-LOC'	across
lla'm**a-p**i	'llama-LOC'	across
lla'p**a-p**i	'lightening-LOC'	across
ma'm**a-p**i	'mom-LOC'	across
pam'p**a-p**i	'prairie-LOC'	across
pa'p**a-p**i	'potato-LOC'	across
q'a'p**a-p**i	'palm of hand-LOC'	across
sun'kh**a-p**i	'beard-LOC'	across
t'i'k**a-p**i	'flower-LOC'	across
tha'p**a-p**i	'nest-LOC'	across
uhu't'**a-p**i	'sandal-LOC'	across
wa'k**a-p**i	'cow-LOC'	across
wall'p**a-p**i	'chicken-LOC'	across
wa'w**a-p**i	'baby/child-LOC'	across
'p**ap**a	'potato'	within
'll**ap**a	'lightening'	within
'**ap**i	'corn/citrus drink'	within
'th**ap**a	'nest'	within
'q'**ap**a	'palm of hand'	within

* ' indicates stress, ' indicates ejective

† Each "across" item additionally inflected with -*man* (ALLATIVE) (See Appendix C).

The real word stimuli came from recordings of an adult female bilingual Quechua-Spanish speaker. These recordings were digitized at a sampling frequency of 44.1 kHz using a portable Zoom H1 Handy Recorder. Stimuli were normed for amplitude between words, but not duration, since some words had ejectives, fricatives, etc. that are temporally longer. The real word picture stimuli were color photographs of the objects. These picture stimuli are available for viewing and reuse in the Open Science Framework project repository affiliated with this work (Cychosz 2020).

Children in these communities have limited exposure to technology (some mothers have flip phones but most children are unfamiliar with computing devices). Consequently, instead of presenting each picture stimulus on a screen, which could have been culturally inappropriate, pictures were presented on individual pages clipped into an 11 x 12.4” plastic binder. For this reason, the words were not entirely randomized for each participant. Instead, two different randomized lists were created with approximately half of the children and adults completing the first list and half completing the second. Since there were more across-morpheme stimuli than within-morpheme stimuli (it was much more difficult to find stimuli for the within-morpheme condition), participants repeated the within-morpheme stimuli three times in the experiment and the across-morpheme stimuli two times. 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.

4.4 Data collection

For the experimental phase, participants were seated on the ground or on a stool, side-by-side with the experimenter. Audio stimuli were were played for the experimenter and participant from an iTunes playlist run on an iPhone 6. Each participant wore AKG K240 binaural studio headphones and the experimenter wore Apple earpods to follow along with the experiment; both headphones were connected to the iPhone with a Belkin headphone splitter.

For data collection, the participant first heard the audio stimulus (a bare noun) and was simultaneously presented with the accompanying photo in the binder. The participant was instructed to simply repeat the bare noun after the model speaker. The participant was then instructed to produce the word again in inflected form. In this way, the researcher could be confident that the children independently knew how to inflect each of the nouns with the tested suffixes and were not simply copying a prompt. For the inflected form, the two morphemes described in section 4.3 were elicited. The locative marker -*pi* was elicited in the first real word repetition task and the allative marker -*man* was elicited in the second task. Ideally, all of the participants' productions would have been spontaneous instead of repeated. However, in a previous version of a similar 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 prompted with the word. Elicited imitation paired with a visual stimulus is also the method used in the previous study on coarticulatory effects within and between morpheme boundaries (Song, Demuth, Shattuck-Hufnagel, & Ménard 2013).

For the children, the inflected forms were elicited using a large plastic toy insect. For the locative marker -*pi*, the toy insect was placed on top of the picture stimulus and the child was prompted, "Where is the bug?" In response, the child produced the word with the correct suffixal carrier, e.g. *llama-pi (kasan)* (llama-LOC COP-3PS, "(It is) on the llama."). For the allative marker -*man*, the researcher wiggled and moved the toy insect on the page towards the noun in question and prompted the child, "Where is the bug going to?," to which the child would produce the word with the suffixal carrier, e.g. *llama-man (risan)* (llama-ALL go-3PS, "(It is going) to/towards the llama."). The adult participants were merely told to add the relevant morpheme to each prompted word in a carrier phrase. For the first real word task, the carrier phrase was, "I say in the _____ two times" (*Noqa nini _____-pi iskay kutita*). For the second task, the phrase was, "I say to/towards the _____ two times" (*Noqa nini _____-man iskay kutita*). The experimenter would then manually advance to the next stimulus item.

Eliciting participant responses in quasi-sentential contexts such as these may disguise the contrasts between inflected and base word forms since "phonetic variation between orthographically-distinct homophones increases when the target homophones are dictated in isolated word list or in contrastive sentences" (Seyfarth et al., 2018:35). The elicitation methods used here thus discourage metalinguistic awareness to the degree possible given the number of lexical items needed for the experiment.

Participant responses were recorded with a portable Zoom H1 Handy Recorder at a 44.1 kHz sampling rate. Children were rewarded with stickers throughout the task and many additionally chose to help the experimenter flip through the pages of the binder.

4.5 Data analysis

Each participant's audio file was first hand-aligned to the word level in Praat (Boersma & Weenink 2018). A Quechua forced aligner was trained on all of the participants' data using the Montreal Forced Aligner (McAuliffe et al. 2017) to align the words to the phone level (The trained Quechua aligner used in the study is available at `https://github.com/megseekosh/kid_align`). The phone-level alignment was subsequently hand-checked by one of two trained phoneticians. One of these phoneticians was blind to the hypothesis of the current experiment. Alignment was conducted auditorily and by reviewing the associated acoustic waveform and broadband spectrogram in Praat.

Coarticulation measures can be sensitive to alignment decisions, so a number of parameters were set prior to alignment to ensure segmentation reliability. Word-initial plosive, affricate, and ejective onset corresponded to the burst. The start of the vowel [a] corresponded to the onset of periodicity and formant structure in the waveform and spectrogram. Nasals were differentiated from vowels by the presence of anti-formants in the spectrogram and a reduction in intensity in the waveform. Additional parameters were set (e.g. for glides, rhotics) but are irrelevant for the segments under analysis in the current study.

To evaluate agreement between the phoneticians conducting the alignment, both phoneticians aligned two randomly-selected word lists, one from a child aged 5;9 and another from a child aged 7;4. For the 5;9 child's list, the difference between the aligners' average consonant duration was 4ms and the average difference in vowel duration was 2ms. Pearson correlations between the aligners for the 5;9 child's list were significant for consonants: r=0.86 p<.001, 95% CI=[0.83, 0.89] and vowels: r=0.94 p<.001, 95% CI=[0.93, 0.96]. For the 7;4 child's list, the difference between the aligners' average consonant duration was 2ms and the average difference in vowel duration was 2ms. Pearson correlations between the aligners for the 7;4 child's list were significant for consonants: r=0.98 p<.001, 95% CI=[0.97, 0.98] and vowels: r=0.95 p<.001, 95% CI=[0.94, 0.96]. The high levels of agreement between aligners suggest high fidelity to the alignment protocol.

As mentioned in the previous literature, much of the work that uses acoustic instead of articulatory coarticulation techniques employs measures such as center of gravity for fricatives, Peak equivalent rectangular bandwidth (ERB_N) (Reidy et al. 2017), or formant-based measurements (e.g. transitions or spectral peaks) for vowels (Lehiste & Shockey 1972; Öhman 1966). However, valid estimates of child formants and even sonorants are notoriously difficult to measure (Chen et al. 2019). In this study, coarticulation is instead measured 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). Specifically, for this analysis, Mel-frequency log spectra were measured over the middle third of two adjacent phones (e.g. [a] and [p]). From these spectral vectors, an average spectrum was computed for each phone. Finally, the Euclidean distance between the averaged spectra was computed where a greater Euclidean distance equated to *less* coarticulation between the phones and a smaller Euclidean distance equated to *more* coarticulation. Using this technique, coarticulation was measured automatically using a custom Python script running, now available open-source (`https://github.com/megseekosh/Meas_Quechua_coartic`), running Librosa functions (McFee et al. 2015).

5 Results

The primary research question in this study asks how child and adult Quechua speakers coarticulate between and within morphemes. The results begin with descriptive statistics concerning the amount of coarticulation by age group (children aged 5 through adults) and morphological environment (within versus between morphemes). Additional descriptive statistics outline the duration of the VC sequences by age group and morphological environment, as the phone durations may interact with coarticulatory patterns. Then, a series of models are fit to predict coarticulation and duration by age and morphological environment. These models are complemented by an analysis highlighting how coarticulation interacts with duration differently in adults and children in the two morphological environments. Scripts to compile these results and fit the models are publicly available (`https://github.com/megseekosh/Meas_Quechua_coartic`).

All analyses were conducted in the RStudio computing environment (version: 1.2.5019; Team (2020)). Data visualizations were created with `ggplot2` (Wickham 2016). Modeling was conducted using the `lme4` (Bates et al. 2015), `lmerTest` (Kuznetsova et al. 2017), and `glmmTMB` (Brooks et al. 2017) packages and summaries were presented with `papaja` (Aust & Barth 2018) and `Stargazer` (Hlavac 2018). Tests of residual normality were conducted using the `normtest` package (Gavrilov & Pusev 2014). The significance of potential model parameters was determined using a combination of log-likelihood comparisons between models, AIC estimations, and p-values procured from model summaries. In all models, continuous predictors were mean-centered to facilitate model interpretation.

5.1 Descriptive statistics

Coarticulation. The degree of coarticulation between the VC sequences [ap] and [am] was measured using the spectral distance metric described in the Methods. For this coarticulation metric, coarticulation is quantified as the Euclidean distance between the spectral vectors of two adjacent phones, henceforth the Mel spectral distance. In this outcome measure, a larger spectral distance between phones equates to *less* coarticulation between the phones. See Methods for further details.

Table 2
Mean spectral distance between [a] and [p] by age and word position

	Across boundary		Within boundary	
Age	Spectral Distance	SD	Spectral Distance	SD
5	17.42	4.64	19.63	4.56
6	17.62	4.70	19.78	4.59
7	17.04	4.68	19.18	5.90
8	18.50	6.03	21.35	6.76
9	16.21	7.02	17.51	7.79
10	14.92	5.09	15.93	5.07
adult	14.99	3.64	14.51	3.03

Table 3
Mean spectral distance between [a] and [m] by age and word position

	Across boundary		Within boundary	
Age	Spectral Distance	SD	Spectral Distance	SD
5	7.94	3.13	8.71	3.39
6	8.44	3.08	9.54	4.76
7	8.43	3.11	8.86	3.80
8	9.89	3.99	9.73	3.78
9	7.43	2.64	8.06	3.11
10	8.21	3.43	8.30	4.25
adult	7.44	4.19	7.00	3.34

Table 2 shows the mean Mel spectral distance between the segments in [ap] (words inflected with *-pi* for the across morpheme boundary condition) and Table 3 shows this for [am] (words inflected with *-man* for the across morpheme boundary condition). Unsurprisingly, there is a larger average spectral distance between the vowel and plosive in [ap] than the vowel and nasal in [am] because the segments in [am] have increased acoustic similarity (sonority, voicing). Next, looking by age group for coarticulation between [ap], it is apparent that the amount of coarticulation between segments increases as children age. This is likely due to the increased speaking rate in the older cohorts and adults, as will become apparent when these results are crossed with sequence duration. There is less within age group variability in [ap] productions in the adult speakers as well (reflected in the smaller SD of the mean for the adults). Variability does not appear to decrease linearly as both the nine and ten-year-old cohorts exhibit larger SDs than the five and six-year-olds. However, this could also reflect the smaller sample sizes in the older cohorts (5 ten-year-olds and 5 nine-year-olds but 10 each in the five- and six-year-old cohorts).

For the coarticulation patterns between segments in [am], adults and children appear to coarticulate similarly, irrespective of age group. There is slightly greater variability in the amount of coarticulation (larger SD) in the adult group. Overall, the difference in developmental coarticulatory patterns between the VC sequence [am] and the sequence [ap] could be due to differences in the *-man* suffix use in Quechua (frequency, productivity). Alternatively, differences between these sequences could be due to the acoustic differences between [ap] and [am] as segments in [am] are more acoustically similar than segments in [ap].

Duration. Next, descriptive statistics outlining the interaction of coarticulation and sequence duration are presented. Duration could interact with coarticulation, and age, as speakers coarticulate more in fast speech (Gay 1981), and adults speak faster than children (Lee et al. 1999).

Table 4
Mean duration of [ap] sequence by age and word position

	Across boundary		Within boundary	
Age	Duration (ms)	SD	Duration (ms)	SD
5	228.7	47	339.9	52
6	242.5	45	334.4	59
7	245.1	57	319.8	60
8	226.9	56	329.3	43
9	216.6	36	312.6	56
10	212.8	50	302.3	63
adult	205.7	47	197.0	59

Table 5
Mean duration of [am] sequence by age and word position

	Across boundary		Within boundary	
Age	Duration (ms)	SD	Duration (ms)	SD
5	214.4	45	251.2	66
6	217.9	58	247.8	62
7	209.8	42	244.6	60
8	218.2	63	247.1	53
9	199.8	30	231.8	64
10	194.3	33	239.9	57
adult	175.6	36	173.7	39

Table 4 maps average sequence duration of [ap] by age and word position and Table 5 does similarly for [am]. Overall, duration of [ap] decreases with age, with adult speakers exhibiting the shortest average duration for [ap] for both word positions. Of note is that the average duration of [ap] sequences within morphemes tended to be longer than the duration of [ap] sequences across boundaries for all of the children. This pattern was reversed in the adult speakers, however, whose average [ap] duration within morpheme boundaries was actually *shorter* than across. This pattern will be revisited in the modelling portion of the Results.

Turning to the sequence [am], the average duration of the [am] sequence in adult speakers was shorter than all of the child speakers; and like [ap], [am] duration also appears to decrease with age. This pattern of [am] duration decreasing with age is consistent for both word positions, across and within morphemes.

Concerning differences by word position, all age groups showed, on average, shorter [am] durations in the across morpheme condition, contrary to the finding that [ap] sequence duration was longer in the across morpheme condition for adults. However, it is important to note the average duration of [am] by word position in adults only differed by approximately 3 ms while the average within morpheme condition in children was approximately 50 ms greater than the across morpheme condition in children. Thus, for both [am] and [ap] sequences, there appears to be a difference in sequence duration by word position for adult and child speakers. In the following section we turn to the modeling of coarticulation before illustrating how degree of coarticulation interacts with duration differently in the adults and children.

5.2 Modelling interaction of coarticulation and duration

The central research question in this study asked if child and adult Quechua speakers would coarticulate similarly between VC sequences across versus within morpheme boundaries. To answer this question, a series of generalized linear mixed effect models (GLMMs) were fit to predict degree of coarticulation (Mel spectral distance between each V and C). GLMMs were chosen instead of the more common linear mixed effect models due to the non-normality of the residual **VC Sequence duration** (henceforth simply **Sequence Duration**) which was included in all models. (Shapiro tests of kurtosis and skewness for **Sequence duration** indicated that we could reject the null hypothesis that the residual's distribution did not differ significantly from a normal distribution. Kurtosis $t=5.53$, $p<.001$ and skewness: $t=1.07$, $p<.001$ (Shapiro et al. 1968).) Specifically, the response variable **Spectral distance** and the residual **Sequence duration** are both limited to non-negative values (as all VC sequences had some distance between the V and C and all had a duration), with a resultant right skew to the data distribution. The choice to fit gamma GLMMs, as opposed to log-normalizing **Sequence duration** and fitting linear mixed models, reflects recent suggestions in cognitive psychology to avoid data transformation, even for commonly-transformed variables such as time/duration, in an effort to facilitate between-study comparison (Lo & Andrews 2015). Consequently, gamma GLMMs were fit using a log linking function to appropriately model the skewed, non-Gaussian distribution of the residual.

A GLMM was fit to predict the spectral distance between segments in the VC sequences. Baseline models included random effects of **Participant** and **Word** (models with random slopes of **Participant** by **Word** did not converge, possibly due to the number of repetitions per speaker; see Methods). Model building then began in a forward-testing manner with predictors added in the following order: **Sequence duration**, **VC sequence** ([ap] or [am]), **Age** (adult or child), **Environment** ([within morpheme or between morphemes]), and interactions. The best model fit included the four-variable interaction of **Sequence duration**, **VC sequence**, **Age**, and **Environment**. The summary for the model containing adults and children together is included in Appendix D. This four-variable interaction indicates that the coarticulation and duration patterns differ between adults and children. Consequently, given the difficulty in interpreting four-variable interactions, separate models were fit for adults and children to facilitate coefficient interpretation.

For both the adult and child groups, models were fit to predict the spectral distance between segments in the VC sequences. Best model fit for the adult and child models included the three-variable interaction of **Sequence duration**, **VC sequence**, and **Environment**: the improvement of models with this interaction over models with the three independent effects was significant for the adult model with alpha level <.05 (χ^2 = 12.69, df=7,11 p=.01) and the child model (χ^2 = 15.18, df=7,11, p=.004). Throughout the results, the children's patterns are additionally broken apart by age to view differences between age groups. However, note that in the child model, the addition of the variable **Age Group** (levels: 5, 6, 7, 8, 9, 10) did not improve upon a model with the interaction of **Sequence duration**, **VC sequence**, and **Environment**. This fact suggests that the pattern of coarticulation and duration by morphological environment did not significantly vary by the child's age group.

The final adult model summary is listed in table 6 and the child model summary is listed in table 7. Note that in the model summaries for the children and adults, the coefficients and standard error measurements were multiplied by 100 to make the otherwise small coefficients more interpretable. This step does not effect the direction or magnitude of the effect between predictors and outcome variables.

In the adult and child models, a positive coefficient for the predictor **VC sequence**, with the reference level '[ap]', shows that there was greater spectral distance between the segments in [ap] than [am], as we would anticipate given the acoustic signatures of [m] (voiced, sonorant) versus [p] (voiceless, transient) (adult model: β=61.78, z=10.78, p<.001; child model: β=72.39, z=16.32, p<.001).

Also of note in both the adult and child models is the direction of the **Sequence dura-**

Table 6
Model predicting coarticulation in adults

term	estimate	S.E.	z.statistic	p.value	95% CI
Intercept	213.56	5.50	38.84	<.001	2.24,2.03
Sequence duration	0.33	0.06	5.52	<.001	0,0
Environment:across morpheme	-7.56	6.89	-1.10	0.27	0.06,-0.21
VC sequence:ap	61.78	5.73	10.78	<.001	0.73,0.51
Sequence duration*Environment:across morpheme	-0.22	0.08	-2.88	0.003	0,0
Sequence duration*VC sequence:ap	-0.21	0.07	-2.95	0.003	0,0
Environment:across morpheme*VC sequence:ap	6.80	8.05	0.84	0.40	0.23,-0.09
Sequence duration*Environment:across morpheme*VC sequence:ap	0.29	0.09	3.20	0.001	0,0

Table 7
Model predicting coarticulation in children

term	estimate	S.E.	z.statistic	p.value	95% CI
Intercept	216.80	3.70	58.66	<.001	2.24,2.1
Sequence duration	0.06	0.02	3.15	0.001	0,0
Environment:across morpheme	-0.94	3.33	-0.28	0.78	0.06,-0.07
VC sequence:ap	72.39	4.44	16.32	<.001	0.81,0.64
Sequence duration*Environment:across morpheme	0.07	0.03	2.38	0.02	0,0
Sequence duration*VC sequence:ap	-0.01	0.03	-0.32	0.75	0,0
Environment:across morpheme*VC sequence:ap	-6.06	5.08	-1.19	0.23	0.04,-0.16
Sequence duration*Environment:across morpheme*VC sequence:ap	-0.09	0.04	-2.40	0.02	0,0

tion predictor: a positive coefficient for **Sequence duration** indicates that longer duration VC sequences tended to also be less coarticulated (greater spectral distance between phones) (child model: β=0.06, z=3.15, p=0.002; adult model: β=0.33, z=5.52, p<.001). The coefficients suggest that when speakers, both adults and children alike, speak slower, they tend to coarticulate less. There is, however, an interaction between several of these predictors, which will demonstrate that children in particular do not always coarticulate less in longer-duration sequences. The direction of the interaction between **Sequence duration**, **VC sequence**, and **Environment** differs between the adult and child speakers so this will be interpreted separately for the two groups in the following section.

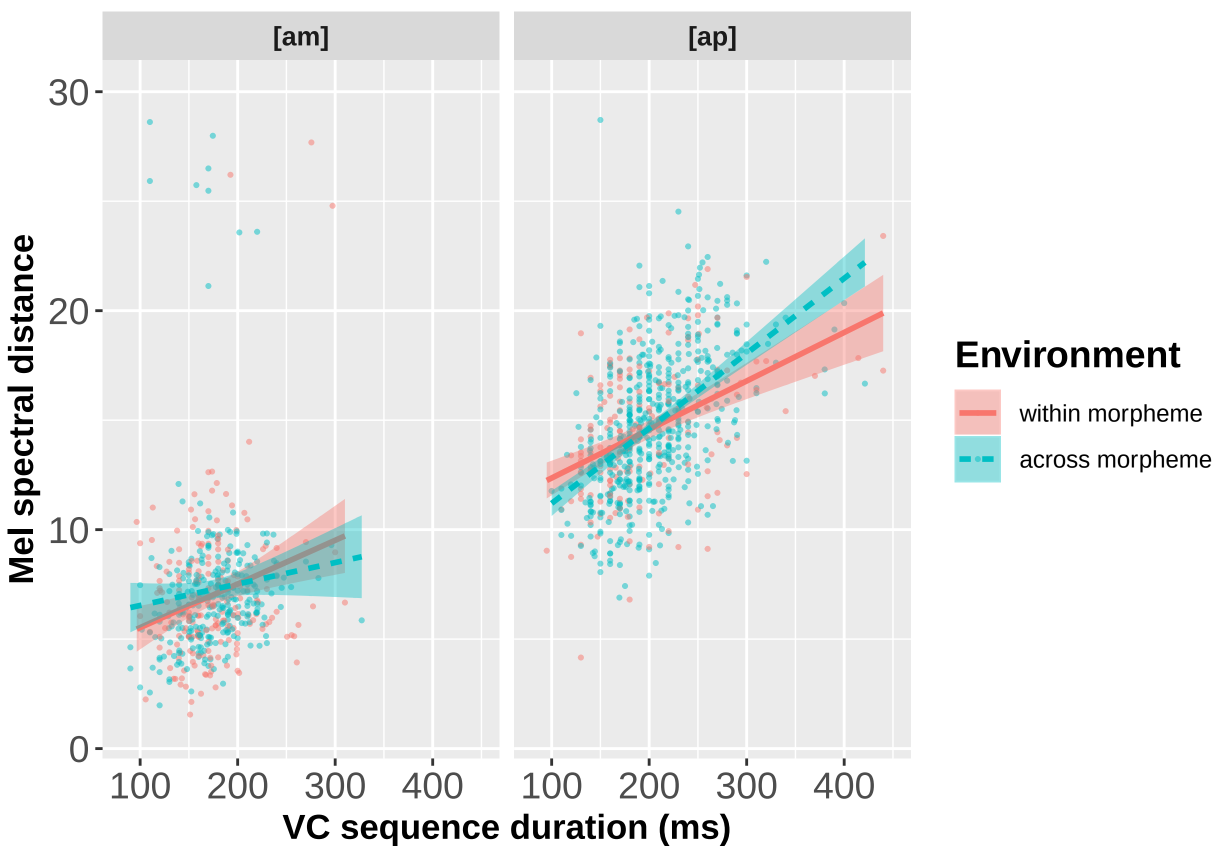

Figure 1. Coarticulation within VC sequence by sequence duration and morphological environment in adult speakers

Adults. For the adult model, the interaction between **Sequence duration**, **VC sequence**, and **Environment** suggests a difference in the relationship between the response variable - amount of coarticulation - and **Sequence duration** that differs by **Environment** and **VC sequence**. As Figure 1 demonstrates, this difference by **Environment** is apparent in the steepness of the slope for the ‘across morpheme’ and ‘within morpheme’ conditions for [am] and [ap]. To quantify this difference for the sequence [am], the slopes of the two conditions were calculated. As the [am] panel in Figure 1 suggests, the slope for the ‘within morpheme’ condition was steeper (2.14) than the slope for the ‘across morpheme’ condition (2.06),[11] suggesting a different relationship between duration and coarticulation between the two word environments in adults.

Overall, the significance of the interaction **Sequence duration**, **VC sequence**, and **Environment** in adult speakers shows two important results: first, adults distinguish by word environment, both for [ap] versus [a#p] sequences and [am] versus [a#m] sequences. Second, complicating this finding, is the fact that adults distinguish between word environments differently depending upon the VC sequence. For [ap], though adults coarticulate roughly equally across and within morphemes, the relationship between duration and coarticulation (longer duration equates to less coarticulation) is stronger in the ‘across morpheme’ condition. For [am], adults also distinguish between the two morphological environments by the relationship of VC duration and coarticulatory degree, but the effect of condition is reversed: the relationship between duration and coarticulation is stronger for the ‘within morpheme’ condition.

Thus, returning to one part of the central research question - does adult coarticulation differ by word environment - we find that adults do coarticulate differently in the two word environments. However, despite these significant differences, there was nevertheless a positive relationship between duration and amount of coarticulation for all combinations of VC sequences and word environments. Adults consistently coarticulate less in longer-duration sequences. This result suggests that adult speakers may have one overarching articulatory plan for all environments and both VC sequences measured. The following section demonstrates how this relationship between duration and coarticulation may not be uniform between adults and children.

Children. Turning to the child model, the significant interaction of **Sequence duration**, **VC sequence**, and **Environment** suggests that children do not coarticulate similarly in longer-duration sequences for all combinations of **Environment** and **VC sequence** (Figure 2). Specifically, for [ap] sequences that occur across morpheme boundaries, the negative slope indicates that children actually coarticulate *more* in longer duration sequences. The positive slope for the within morpheme boundary condition suggests that children coarticulate less in longer-duration sequences, in line with all of the adult patterns. So, children coarticulate more between segments at morpheme boundaries in words inflected with the locative marker *-pi* than between those same segments that occur within morphemes.

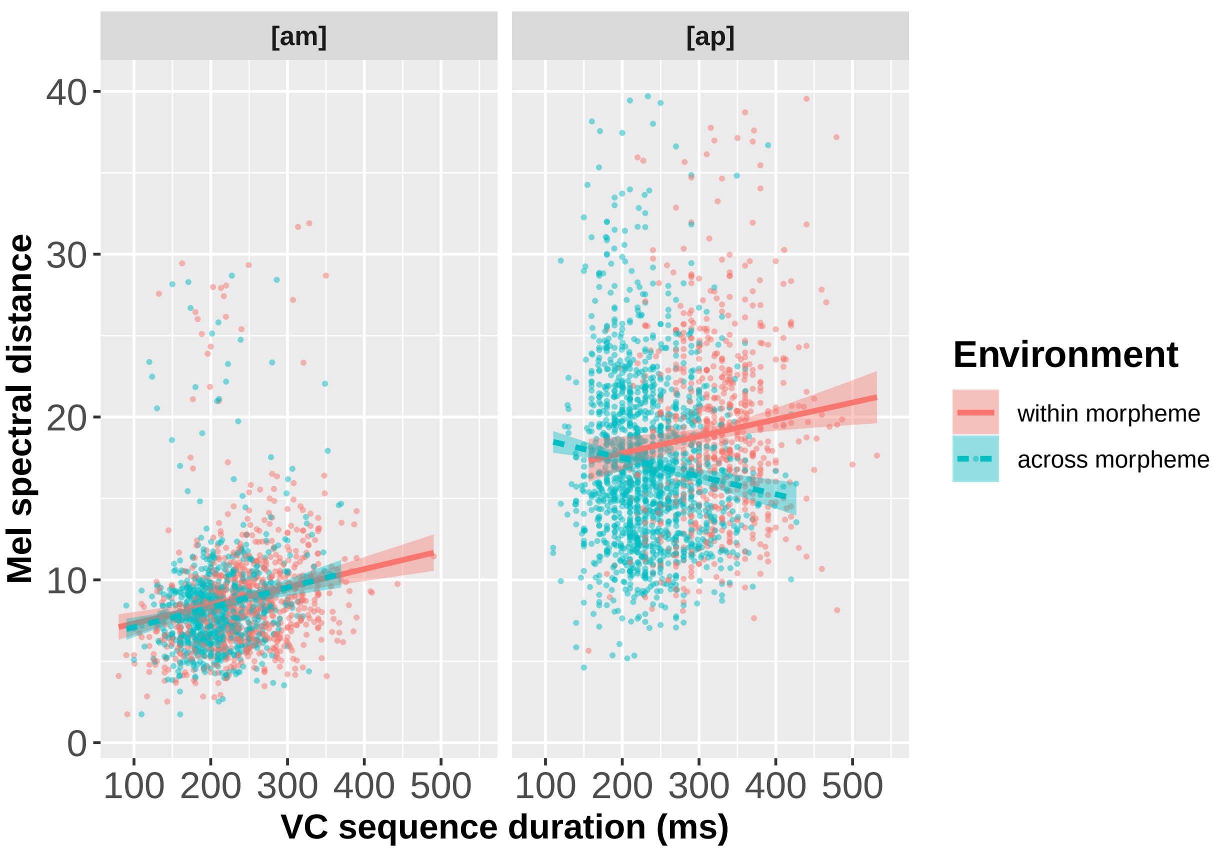

Figure 2. Coarticulation within VC sequence by sequence duration and morphological environment in all child speakers

Note that this negative relationship between duration and coarticulation is counter to the

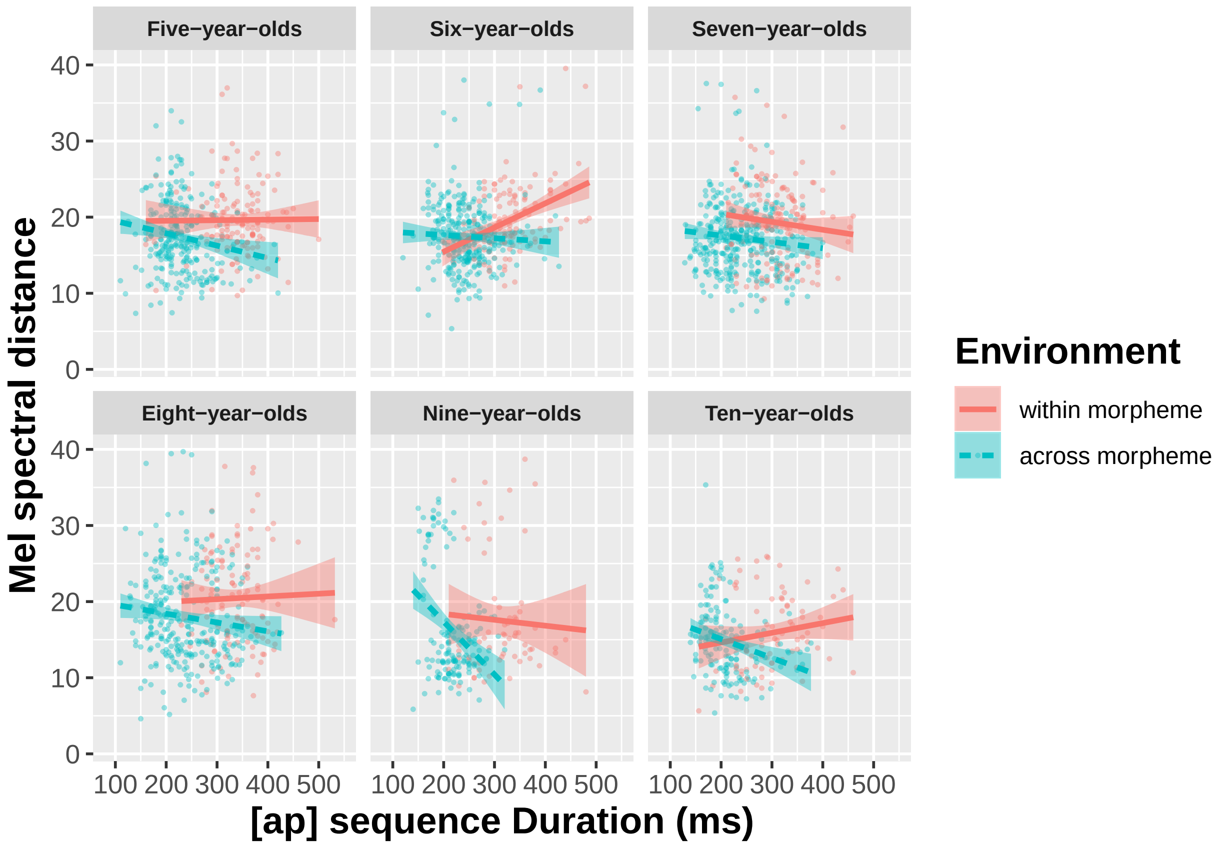

Figure 3. Coarticulation within [ap] by sequence duration, morphological environment, and age in child speakers

positive relationship for every combination of VC sequence and word environment in adult speakers. Adults consistently coarticulate less in longer-duration sequences regardless of environment or VC sequence. The facet plot in Figure 3 plots this relationship between duration and coarticulation for [ap] for each age group (5-10 years) to ensure a consistent pattern across the groups. (Again the predictor **Age Group**, with the levels 5, 6, 7, 8, 9, and 10 years, did not improve upon the child model fit.) All age groups show the same negative relationship: the longer the [ap] sequence, the more the children coarticulate between [a] and [p] in the across morpheme condition.

The results for [am] in children demonstrate broadly similar results to the adult speakers: children coarticulate less between segments in longer-duration [am] sequences. The facet plot in Figure 4 once again shows a similar effect for each age group. Given the between-subject variability that typically characterizes child speech, these patterns by environment are further broken apart by

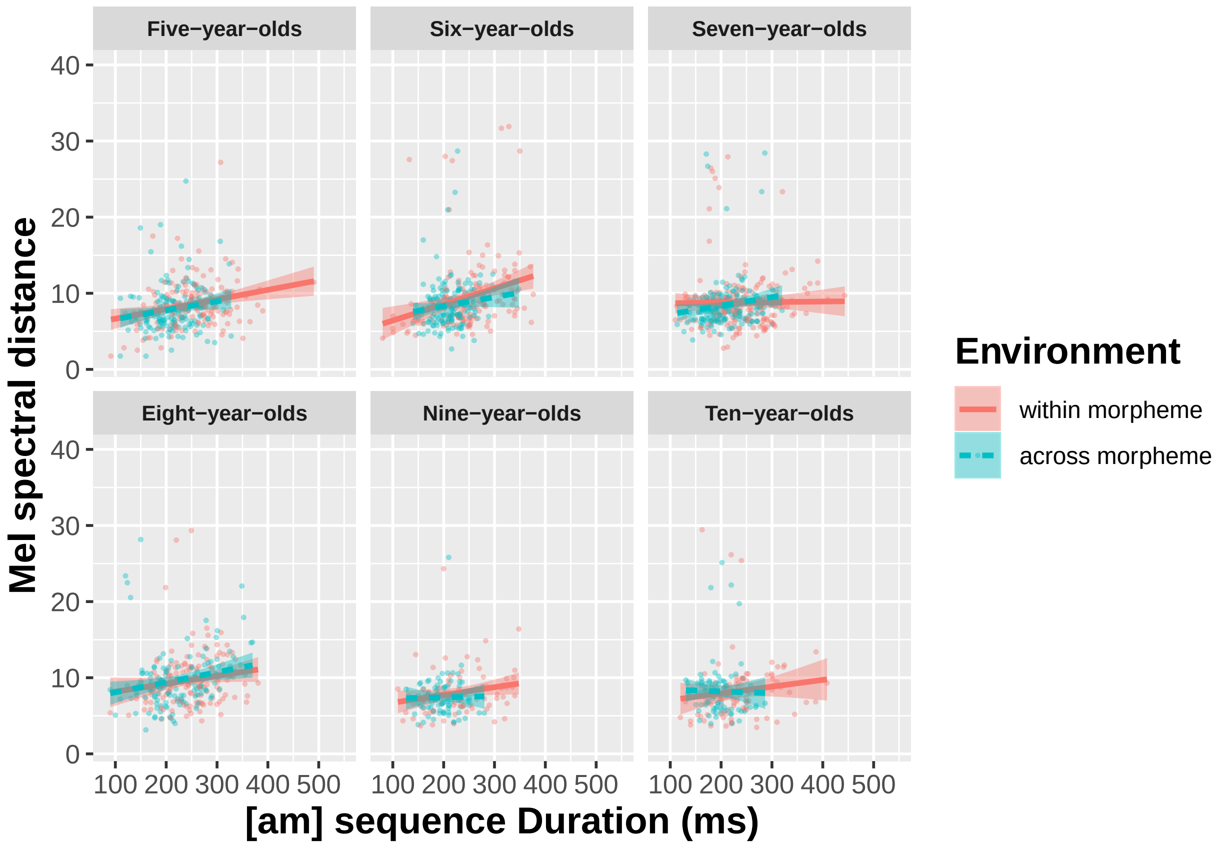

Figure 4. Coarticulation within [am] by sequence duration, morphological environment, and age in child speakers

individual child for each age group (age 5-10) (Appendix E) to ensure no large outliers with regards to the patterning by word environment. The results by are broadly similar across speakers.

5.3 Interim discussion

Comparing between the adult and child models, several preliminary conclusions can be made. First, responding to the original research question - do adults and children coarticulate differently within versus between morpheme boundaries - we find that both adults and children differentiate by morphological environment. However, they do so in different ways. Adults have a single plan for both environments, and even both VC sequences: adults coarticulate less in longer-duration sequences, and overall, they coarticulate less between [a] and [p] within morphemes than across morpheme boundaries. The stark difference between adults and children emerges in the [ap] sequence patterning. Children differentiate between morphological environments via the relationship between duration and coarticulation as they coarticulate more in longer-duration sequences across morpheme boundaries and coarticulate *less* in longer-duration sequences within morphemes.

For words inflected with *-man*, children show a similar pattern to adults, though children do not differentiate by environment coarticulatorily. Rather, across morpheme sequences are shorter in duration than within morpheme sequences for the children. On the basis of these results, two questions remain. First, why do children differentiate between environments via a combination of duration and coarticulation; specifically, why do children produce shorter duration VC sequences across morpheme boundaries and longer duration sequences within morphemes? Second, why do children coarticulate more in longer-duration [ap] sequences that cross morpheme boundaries (e.g. *llama-pi* 'llama-LOC')? All of the other combinations of morphological environment and VC sequence in the adults and children suggest that the speakers coarticulate less in longer-duration segments.

The finding that children produce shorter-duration VC sequences in the 'across morpheme' condition than the 'within morpheme' condition for [am] and [ap] could be explained by a confound in word size and morphological environment. Coarticulation for the 'across morpheme' condition was, necessarily, measured across morphemes. However, to derive an across morpheme environment in Quechua, nouns are inflected with suffixes (e.g. *llama* 'llama' -> *llama-pi* 'llama-LOC'). As a result, almost all of the stimuli in the 'across morpheme' condition are at least one syllable longer in length than the stimuli for the 'within morpheme' condition. For example, coarticulation between [ap] was frequently measured within two syllable base roots (e.g. *llapa* 'lightening' and *llama* 'llama'). However, for the 'across morpheme' condition, [ap] coarticulation was frequently measured in three-syllable inflections of these nouns (e.g. *llapa-pi* 'lightening-LOC' and *llama-pi* 'llama-LOC'). Even for prosodically longer words where within morpheme coarticulation was measured, such as the three-syllable *hampiri* 'healer' and *hatun mama* 'grandmother', there were equivalent across morpheme stimuli that were one syllable longer (e.g. *hatun mama-mang* 'grandmother-ALL').

5.4 Compensatory shortening

To explore the possibility that durational differences between the ‘across morpheme’ and ‘within morpheme’ conditions could be due to word length, an exploratory analysis was conducted. It was anticipated that sequence duration would shorten in words with more syllables, regardless of morphological context. This well-known tendency for segment durations to shorten in longer-duration/polysyllabic words is known as Compensatory Shortening (Harrington et al. 2015 Lehiste 1972; Munhall et al. 1992). To illustrate how this unfolds in the children’s speech production for the current study, Figure 5 plots VC sequence duration by the number of syllables for the children and Figure 6 plots duration as a function of number of syllables for the adults. As the children’s figure demonstrates, children’s VC sequences are consistently shorter in words with more syllables, most notably between two- and three-syllable words. The same pattern is not apparent in the adult data: adults have fairly similar sequence lengths regardless of the number of syllables in the word (see Table 8 for descriptive statistics of duration by word length in syllables for the children and Table 9 for duration by word length results for the adults).

Table 8
Mean VC sequence duration by number of syllables in word for children

	[am]		[ap]	
Syllables	Duration (ms)	SD	Duration (ms)	SD
2	272.9	52	325.3	57
3	211.6	51	235.2	51
4	207.6	46	220.7	51
5	204.1	35	230.4	51

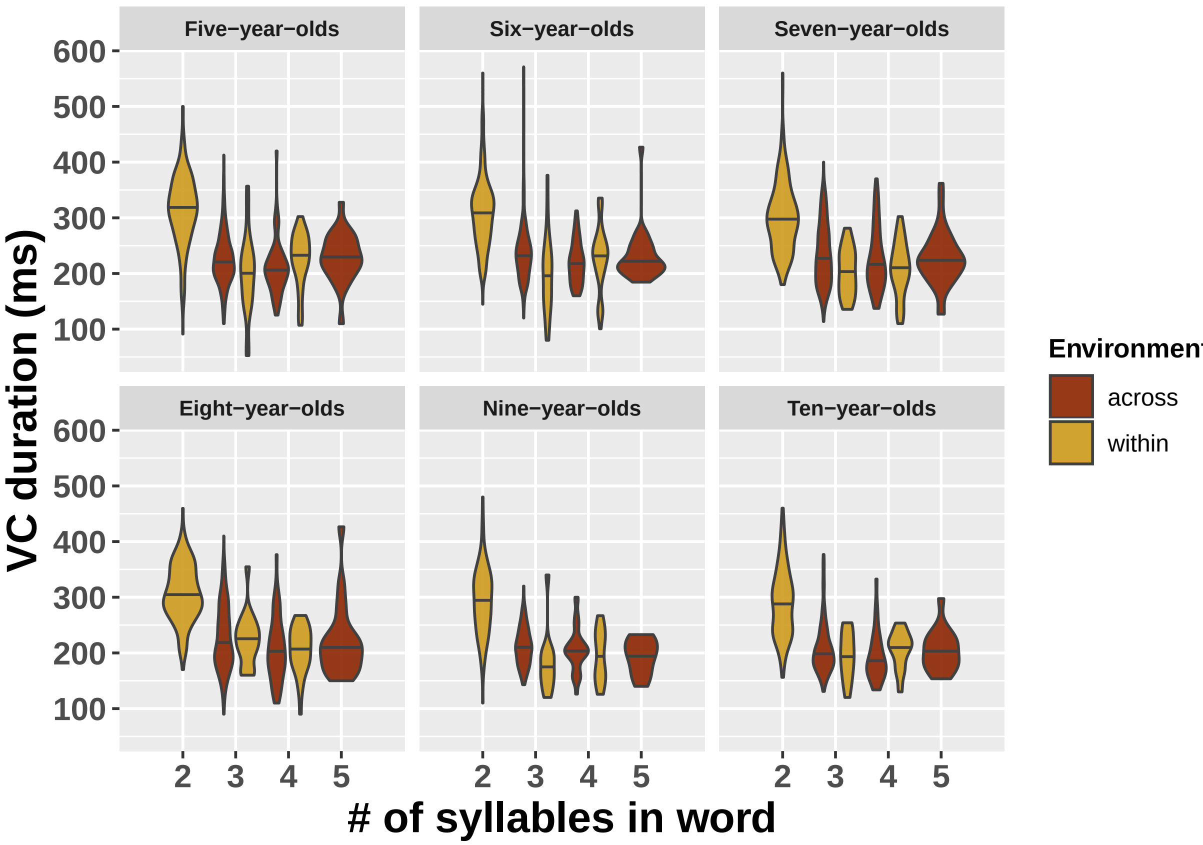

Figure 5. Sequence duration by word length and word environment: Children

Table 9
Mean VC sequence duration by number of syllables in word for adults

	[am]		[ap]	
Syllables	Duration (ms)	SD	Duration (ms)	SD
2	177.6	39	192.4	57
3	176.5	35	210.0	50
4	168.0	36	195.3	43
5	178.5	51	207.3	40

To further explore the phenomenon of Compensatory Shortening in the children's speech, a linear mixed effects model was fit to predict the VC sequence duration in the children's speech (no skewed/non-negative predictors were included in the modelling so GLMMs were not necessary). Model fitting occurred as before in a forward-testing manner: the base model contained random effects of individual **Child** and **Word** (random slopes of Speaker by Word did not converge).

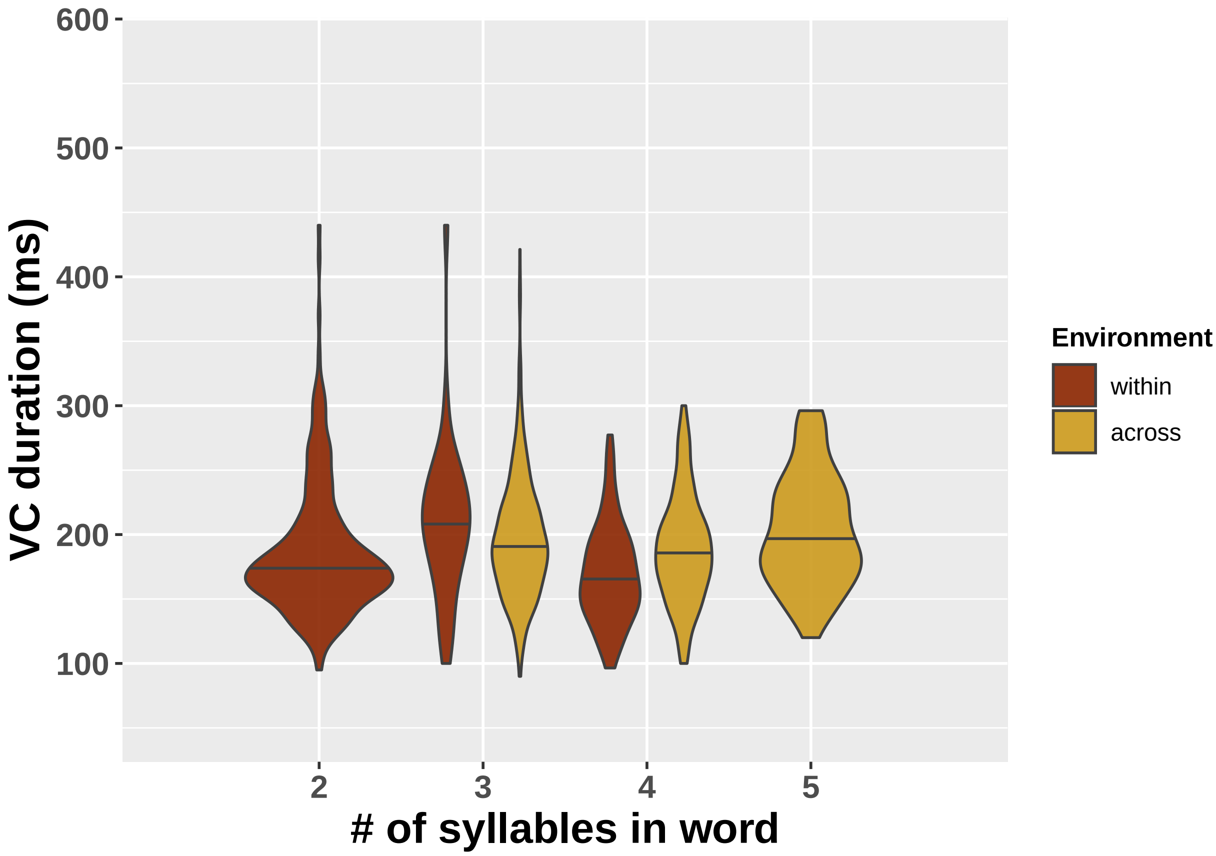

Figure 6. Sequence duration by word length and word environment: Adults

Then, parameters were added in the following order: **Age** (5-10), **Number of Syllables** (2-5), **Environment** (across versus within), the interaction of **Number of Syllables** and **Environment**, and **VC sequence**. Only the predictors **Number of Syllables** and **VC sequence** improved baseline model fit (see Table 10 for model summary).

In the model summary, the positive beta coefficient for VC sequence with a reference level [am] indicates that the [ap] sequence was significantly longer than [am] sequences (as previous models demonstrated). Next, the negative beta coefficients for **Syllable Count** with a reference level of '2 syllables' indicated that VC sequence duration was approximately 80 ms shorter in three syllable words than two syllable (β=-79.68, t=-14.61, p<.001). Similarly, VC sequences were approximately 93 ms shorter in four syllable words than two syllable (β=-93.38, t=-13.06, p<.001) and 87 ms shorter in five syllable words than two syllable words (β=-87.01, t=-7.87, p<.001). The insignificance of **Environment** and child **Age** for the modelling suggests that that this relationship

Table 10
Model predicting VC duration: children

Intercept	289.36***
	(276.14, 302.58)
Three syllables	−79.68***
	(−90.37, −68.99)
Four syllables	−93.38***
	(−107.40, −79.37)
Five syllables	−87.01***
	(−108.68, −65.34)
VC sequence:[ap]	27.66***
	(19.56, 35.77)
Observations	3,877
Log Likelihood	−20,122.98
Akaike Inf. Crit.	40,261.97
Bayesian Inf. Crit.	40,312.07
Note:	*p<0.05; **p<0.01; ***p<0.001

between duration and word length is independent of morphological environment and child age.

As these coefficients demonstrate, VC sequence decreases in larger words, with the largest differences between two- and three-syllable words. The diverse stimuli in the two- and three-syllable conditions (many different word types) suggest that this relationship by word length is relatively robust.

The only exception to the tendency to shorten sequences in larger words was that [ap] sequences are slightly longer in duration in 5-syllable words than 4-syllable words. However, in this exploratory analysis, the differences between the four and five syllable words were not tightly controlled: there were only two different five-syllable word stimuli: *hatun mama-man* 'grandmother-ALL" and *hatun mama-pi* 'grandmother-LOC'. We can only speculate that this relationship between sequence duration and number of syllables is strictly linear and would generalize to additional words with more syllables.

In conclusion, on the basis of this modeling, it is proposed that children may differentiate by morphological environment in their speech production. However, in Quechua, morphological structure is crossed with prosodic structure: complex words are always structurally longer than base forms. Regardless, children distinguish between morphological/prosodic environments primarily via the acoustic cue of *duration*: in child Quechua, the duration of sequences is shorter in words with more syllables (Compensatory Shortening). This duration pattern by word length was not present in the adult speech, a finding that is explored in the Discussion.

6 Discussion

The experiment in this study employed a spectral measure of coarticulation to measure the coarticulatory patterns of adult and child Quechua speakers in two morphological environments. Two predictions were made for this experiment: 1) adult speakers would coarticulate less between two phones at a morpheme boundary than between the same phones within a morpheme boundary, and 2) child speakers would coarticulate more than adults between phones at morpheme boundaries. The reasoning behind these hypotheses was that frequency ratios between suffixes, and between roots and suffixes, predict decompostability in adults (Hay 2003; Kemps et al. 2005; Plag & Baayen 2009). Adult speakers have more experience with language than children and have larger vocabularies (Lorge & Chall 1963). Consequently, adults may weigh the ratios of inflected to base forms in their lexicons differently. And children may structure relationships between base and suffixal forms differently as they age and their vocabularies grow. Moreover, children appear to coarticulate less as they age and gain more experience with words and segments (Noiray, Popescu, et al. 2019). A decrease in coarticulation suggests that phonological representations individuate into segment-sized units. For these reasons, it was predicted that adults might decompose words differently from children.

The first hypothesis of this study proposed that adult speakers would coarticulate less between two phones at a morpheme boundary than between the same phones within a morpheme boundary.

Study results appear to confirm this hypothesis. Adult speakers did, overall, coarticulate less across morpheme boundaries than within. This is to be expected because, if as previous studies suggest, speech production - coarticulation and duration - indexes lexical retrieval and composition (Cho 2001; Kemps et al. 2005; Lee-Kim et al. 2013; Plag et al. 2017; Pluymaekers et al. 2010; Song, Demuth, Shattuck-Hufnagel, & Ménard 2013; 2013; Sugahara & Turk 2009; Tomaschek et al. *under review*). Specifically, the adult speakers tended to coarticulate less between phones in a VC sequence at a morpheme boundary than within a morpheme because adults are more likely to compose morphologically complex words from their component parts. Though decomposition is probabilistic (e.g. Hay 2003), overall, adults may be less likely to access morphologically complex words holistically.

The second hypothesis of this study proposed that child speakers would coarticulate more than adults between phones at morpheme boundaries. The results and conclusions of this second second hypothesis proved far more complex, particularly given the interactions between degree of coarticulation and speaking rate/VC sequence duration. In slower speech, adult speakers coarticulated less, both across and within morphemes. This replicates known interactions between speaking rate (here instantiated as VC sequence duration) and coarticulation (Agwuele et al. 2008; Matthies et al. 2001). In children, the same relationship between VC sequence duration and coarticulation appeared in the within-morpheme condition: within morphemes, children coarticulated less when they spoke slower. However, one primary difference between adults and children was that, in children, coarticulation did *not* vary as a consequence of VC sequence duration in the across-morpheme condition. Of further interest was the finding that the children's VC sequence duration was reliably shorter in the across-morpheme condition than the within-morpheme condition. The following section interprets the duration finding from the children's data.

6.1 Compensatory shortening

As suggested in the results section, one possible explanation for the shorter sequence duration in the across-morpheme condition for the children is the morphological structure of Quechua words.

To construct morphologically complex words in this agglutinating language, additional suffixes must be appended to the root. As a result, almost all of the across-morpheme stimuli were approximately one syllable longer than the within-morpheme stimuli. It was suggested that the longer prosodic (and temporal) length of the stimuli in the across-morpheme condition caused children to shorten the duration of the phones in each of the stimuli items (or, variably, lengthen phone duration in the shorter, within-morpheme stimuli). This aligns with the well-known phenomenon of compensatory shortening, or segment duration shortening in longer-duration/polysyllabic lexical items (Harrington et al. 2015; Lehiste 1972; Munhall et al. 1992).[12] This finding reinforces previous work on adult speech, as described in the literature review, that has demonstrated how roots have different variants in their bound and free forms (Kemps et al. 2005). One challenge in morphophonetics has been to identify the explanatory mechanisms behind morphologically-conditioned speech variation; at least in the current data, compensatory shortening could be one of these explanatory mechanisms.

This relationship between-prosodic word size and segment duration in the children's speech is notable. For example, it is unclear if these data mean that the children have some minimal planning unit size causing them to elongate prosodically short words (or shorten prosodically long words) to fit within the unit's temporal domain. However, even more interesting perhaps than the children's patterns is the *lack* of compensatory shortening in the adults. While children demonstrated a trade-off in sequence duration and prosodic word size, adults were insensitive to word size. Yet previous findings on compensatory shortening came from adult populations. Why don't adults compensate for word size in their speech production as the children, even the eldest ten-year-olds, appear to?

There are several potential explanations for the difference in compensatory shortening between adults and children, some related to language experience and others reflecting sociolinguistics in Bolivia. First, adults speak faster than children (Lee et al. 1999), leading to more extreme phonetic reduction in their speech. Adult Quechua speakers may speak so much faster, and reduce so much more, than children that there could be insufficient freedom in their speech duration to differentiate sequence duration by prosodic word size. Children do not approximate adult-like speaking rates until early puberty (Lee et al. 1999; B. L. Smith 1992). This increase in speaking rate is replicated in the present study, even in a tightly-controlled experimental setting, as the average VC sequence duration for the adult speakers was just 192 ms (σ=47) compared to 252 ms (σ=69) for the five-year-olds, 255 ms (σ=70), for the eight-year-olds, and 231 ms (σ=63) for the ten-year-olds.

Speakers reduce significantly more in fast speech compared to slower, controlled speech. In fast speech, the vowel space is smaller and more centralized (Fourakis 1991; Tsao et al. 2006) - which may compromise contrasts (Koopmans-Van Beinum 1980), increase coarticulation between phones (Agwuele et al. 2008; Matthies et al. 2001), and cause the omission of entire segments or syllables (Johnson 2004). For the current study, the durational data summarized in the previous paragraph suggest that all adult speech, regardless of prosodic or morphological structure, is maximally fast and reduced.

Though there is little work on the phonetics of highly morphologically complex languages, the author can anecdotally attest to how speaking rate interplays with word structure in Quechua. Just as the orthographic forms of English words rarely correspond to their phonetic realization in fast, spontaneous speech, spoken words in Quechua deviate from their citation form. In Quechua, this is especially true with large words that contain numerous suffixes, where the most extreme reduction can be seen. In Quechua, the further a suffix is found from the root, the *more* likely it is to be reduced (shorter in duration, omission of segments/syllables). The explanation for this is partially aerodynamic (e.g. airflow). But there are also likely perceptual and information-theoretic explanations. When suffixes are highly reduced compared to stems, speakers can more easily demarcate between suffixes and roots, and identify word boundaries (Zingler 2018). Furthermore, word meanings are likely increasingly predictable as additional suffixes are added and the available semantic space of the word narrows.[13] Variability in the phonetic realization of English morphology (e.g. plurals) is dependent upon the predictability of plurality given the sentence frame (Cohen 2014), so it is possible that speakers likewise make probability calculations over multiple suffixes. Much more work is needed to understand probabilistic reduction based on word structure. However, overall, one explanation for adults' lack of compensatory lengthening is their speaking rate and phonetic reduction.

A second explanation for the lack of compensatory shortening in adults may be that the adults are more dominant in Quechua than the children. This means that the adults may speak faster and, again, may be unable to differentiate prosodic structure via duration. Recent changes in Bolivia's educational policy, as well as the country's general sociolinguistic situation, may have led to adults' increased fluency. Bilingual education in Bolivia became mandatory in 1994 and has, in theory, been relatively widespread since the early 2000s (Benson 2004). This means more children, especially indigenous students and young girls, are attending and completing more schooling than ever before (Hornberger 2009).

In practice, however, bilingual education often takes the form of Spanish-only classrooms. In Quechua-speaking areas, many trained teachers do not speak Quechua fluently or are not provided with teaching materials and textbooks written in Quechua. The result - students completing more schooling but instructed in Spanish - has been rapid language shift (Hornberger & King 1996). This has been apparent even in the last decade, as the first generation of women educated in this system are now raising their children using both Quechua and Spanish, instead of monolingual Quechua, in an increasingly Spanish-dominant environment.

These sociolinguistic dynamics could manifest in the present sample as the adult female participants (many of whom are mothers) may be more Quechua-dominant than some of the children in the sample. Though the adult females who participated in the study were only, on average, 13 years older than the eldest children (adult μ_{age}=23 years; σ=5.46 years), and all adults and children identified as bilingual Quechua-Spanish speakers, their language practices may reflect the recent changes in educational policy. It is also important to note that all of the children in the sample attended school, in Spanish, for 3-4 hours per day. However, only one of the adult females was still attending school (taught in Spanish). This difference could also explain usage patterns between the age groups.

If the adult females were more Quechua-dominant, or used Quechua more frequently, they would speak faster, more fluently, and could reduce more, as described above. Thus, although the adult speakers in this study undoubtedly did speak faster than the children for speech maturation reasons - it takes time and practice to master the articulatory speed of an adult (Lee et al. 1999) - the adults may also have spoken Quechua faster, and thus failed to compensate for prosodic structure to the degree that the children did, because they use Quechua more frequently. This proposal may not entirely explain the differences in compensatory shortening between the adult and child Quechua speakers - previous studies on compensatory shortening (Lehiste 1972; Munhall et al. 1992) reported on highly-fluent, monolingual adult speakers who *did* compensate for prosodic structure - but it is one explanation for the differences observed between age groups in the present work.

While this study did not report participants' bilingual language dominance, all speakers identified as Quechua-Spanish bilinguals and reported using both languages.[14] The decision not to conduct a traditional language usage survey was made for several reasons. Traditional measures of self-reported bilingual dominance, such as the bilingual language profile (Birdson et al. 2012), often rely on participant literacy or familiarity and comfort with written behavioral research surveys. Also, the traditional stigmatization of indigenous languages in Bolivia may render self-reports of language dominance unreliable. Nevertheless, one potential difference between the adults and children could be their Quechua language usage or dominance.

6.2 Do children distinguish between morphological environments in their speech production?

The original research question in this study asked if adults and children coarticulated differently across versus within morpheme boundaries. As the discussion of compensatory shortening outlined, children appear to distinguish by morphological environment, suggesting that they decompose words. However, this could be a consequence of prosodic structure, which is correlated with morphological structure in Quechua. The different temporal and coarticulatory patterns seen in the across- and within-morpheme conditions could reflect morphological structure and lexical access, or the patterns could reflect prosodic structure. Children may modulate their speech temporally to demarcate between morphological environments - faster segments across morphemes and slower segments within. Or, as compensatory shortening suggests, children may shorten the duration of VC sequences in prosodically longer words.

One way to evaluate these two explanations is to control for prosodic structure by measuring duration and coarticulation across morphemes versus within morphemes, but always in words of the same length (in number of syllables). This is made more difficult in Quechua because the language has canonical penultimate stress: even if duration/coarticulation were measured between [am] in a three-syllable within-morpheme stimulus (e.g. ll**a**.'**m**a.-man 'llama-ALL'), and a corresponding three-syllable across-morpheme stimulus (e.g. pa.'p**a**.-**m**an 'potato-ALL'), the stress would not be controlled. In other words, any differences in duration/coarticulation patterns between the across-morpheme and within-morpheme stimuli could be attributable to known acoustic correlates of stress, and might not reflect morphological *or* prosodic structure. This issue reflects some recent acknowledgements in morphophonetics that it may be nearly impossible to design studies that perfectly isolate morphological effects from the plethora of other correlated variables (Strycharczuk 2019, cf. Seyfarth et al. 2018).

In the current dataset there are three lexical items that control for prosody (and thus stress) and allow us to tease apart the prosodic versus morphological explanations for the children's patterns. Specifically, there are three four-syllable stimuli - two across morpheme (*imi'll**a-m**an* 'girl-ALL' and *juk'u'ch**a-m**an* 'mouse-ALL') and one within morpheme (*hatun'm**am**a*) - where the [a] of [am] falls in stressed position. In the following exploratory analysis, the duration of [am] and the coarticulation between [a] and [m] were measured for all of the children and adults. Results are listed in tables 11 and 12.

Table 11
Mean duration of [am] across and within morphemes in prosodically-controlled stimuli

	Across boundary		Within boundary	
Age	Duration	SD	Duration	SD
5	193.3(ms)	32	245.7	30
6	199.0	25	242.9	50
7	188.8	34	229.8	36
8	205.3	49	225.2	30
9	189.4	37	226.4	32
10	170.9	26	216.6	17
adult	165.4	30	230.6	33

Table 12
Mean spectral distance between [a] and [m] across and within morphemes in prosodically-controlled stimuli

	Across boundary		Within boundary	
Age	Spectral Distance	SD	Spectral Distance	SD
5	7.34	2.52	8.49	2.45
6	9.78	5.84	9.51	5.39
7	7.85	1.30	8.82	5.05
8	9.61	4.00	8.89	3.37
9	7.62	1.88	6.63	2.15
10	8.34	1.57	8.29	4.66
adult	8.90	4.17	5.48	1.20

As the results demonstrate, both the children and adults have shorter [am] sequences in the across-morpheme condition (stimuli *imilla-man* and *juk'ucha-man*). The duration of [am] also decreases with age, as the adults speak fastest. Thus, when controlling for prosody, it appears that children do pattern like the adults and may distinguish by morphological environment using durational cues. However, for coarticulation, the results differ between adults and children. While the children do not appear to distinguish greatly between morphological environments, coarticulating roughly equally in the two environments regardless of age, the adults distinguish between the two areas. As anticipated from previous work (e.g. Cho 2001), the spectral distance between [a] and [m] is greater in the across-morpheme stimuli (*imill**a-m**ang* and *juk'uch**a-m**an* than the within-morpheme stimulus (*hatunm**am**a*). Unlike the durational results, the exploratory coarticulation results suggest that adults, but not children, distinguish between the the morphological environments.

These data are too sparse to make definitive conclusions; only three distinct lexical items were tested. Concerning the central research question - do children distinguish between morphological environments - the exploratory analysis gave conflicting results: children distinguish between environments like adults durationally, but not coarticulatorily. However, the exploratory analysis does suggest that children's compensatory lengthening is likely morphological in nature, not prosodic, since the [am] was shorter in duration in the across-morpheme condition in the four-syllable stimuli.

Thus, using a combination of coarticulatory and temporal cues in their production, children appear to distinguish between across-morpheme and within-morpheme stimuli. But, in most of the current stimuli, morphological environment is correlated with prosodic environment (to derive morpheme boundaries, extra syllables must be added to roots). It therefore remains unclear if children's different speech patterns across the two morphological environments reflect prosodic planning, lexical planning, or both.

Future work, ideally on languages with different and complex morphological structures, is needed to further explore the relationships between children's speech production and word structure. Going forward, researchers may also benefit from the use of articulatory measures, in addition to acoustic measures, to explore the relationships between coarticulation and morphological structure in children. Given the logistics of ultrasound imaging, and the limitations of fieldwork, the current study was not able to collect articulatory data, which would otherwise be a valuable addition.

This study was unable to control for, or examine the effect of, lexical frequency of the base or inflected forms. This is unfortunate because much variability in morphological parsing, and morphophonetics, is attributable to lexical frequency (Seyfarth et al. 2018) or frequency ratios between stems and suffixes (Hay 2003; Plag & Baayen 2009). At this time, however, it is not possible to reliably calculate lexical frequency statistics for Quechua, and indeed for most of the world's languages. A large, naturalistic corpus of child and adult Bolivian Quechua has recently been collected (Cychosz 2018), but it will be years before it is sufficiently transcribed to calculate lexical statistics. In the mean time, researchers interested in the morphophonetics of underdocumented languages could possibly use age of acquisition as a proxy for word frequency (Morrison et al. 1992) and this may a promising methodological approach should researchers hope to include additional under-documented languages in the study of morphophonetics.[15]

7 Conclusion

The primary goal of this study was to compute coarticulation in adult and child Quechua speakers in two morphological environments: within morphemes and across morpheme boundaries. Results showed that, using a combination of coarticulatory and temporal cues, adults distinguished between the two morphological environments in their speech production. This replicated known speech production patterns by word environment but in an understudied, morphologically complex language. The children showed increased prosodic sensitivity where the adults did not: children shortened the duration of sequences in prosodically longer words, which also happened to be morphologically complex. It was suggested that the difference between adults and children could be attributable to adults' faster speaking rate and increased practice with Quechua. An exploratory analysis isolating prosodic effects from morphological was conducted on a subset of the original data. This analysis suggested that children's speech patterns may be reflecting morphological structure, though future work is needed to determine the level of children's speech planning: prosodic, morphological, or both. Overall, these analyses have demonstrated some of the complexities that arise in morphophonetic patterning and speech development, and the importance of extending recent studies in this sub-domain of phonetics to languages with vastly different word structures and speakers with different language experiences.

Declaration of interest

The author has no competing interests to report.

Acknowledgements

Footnotes

[1]Recent statistical advances have called into question the results of Losiewicz (1995) (see Seyfarth et al. (2018)).

[2]Cuzco Quechua and South Bolivia Quechua are distinct varieties, but they are mutually-intelligible. There should be no reason to assume that children's morphological productivity would vary greatly between them.

[3]Concerning Quechua orthography, it is important to emphasize two things. First, there *is* a standardized, written form of South Bolivian Quechua. For example, students who study Quechua at the university level learn to read and write in the language and *some* textbooks for school-aged children contain some words written in Quechua. Second, in my fieldwork, I have found that speakers certainly *can* write in Quechua, because the orthography is relatively transparent and Quechua shares many phonological characteristics with Spanish (e.g. five vowel system). But I have found this to be uncommon in the communities where I work.

[4]Cuzco Quechua and South Bolivia Quechua are distinct varieties, but they are mutually-intelligible. There should be no reason to assume that children's morphological productivity would vary greatly between them.

[5]Late talker status was not collected from the participants recruited from the school.

[6]This is reported because the presence of front teeth could have notable consequences for speech acoustics (e.g. anterior fricatives). This information is not typically reported in speech development research, but arguably should be.

[7]Quechua is canonically open-syllabic, so all VC syllables transcend syllable boundaries.

[8]The only exception to this was the item *ham'piri* 'healer', and its inflected form *hampi'ri-pi* 'healer-LOC', where the [am] sequence does not coincide with primary stress. This item was included because it was difficult to find sufficient items for the within morpheme condition that adhered to the criterion of frequency, recognition by children, etc.

[9]Some lexical items contained both within and across stimuli (e.g. [am] and [ap] in ll**ama-p**i 'llama-LOC').

[10]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.

[11]To reflect the data visualizations, these slopes were calculated on the beta coefficients before the coefficients were scaled by 100.

[12]The term compensatory shortening is also variably used in the literature to refer to shortening of stressed vowels compared to unstressed vowels in the context of polysyllabicity (Harrington et al. 2015), or the shortening of stressed vowels in the context of unstressed vowels and consonants (Fowler 1981).

[13]For English, however, Plag and Baayen (2009) found that those suffixes farthest from the stem are also the most productive and available for use in novel environments.

[14]Computation of the children's language dominance is underway using naturalistic recordings of the children's daily language usage. This method skirts the issue of self-reported usage.

[15]The author found that obtaining age of acquisition norms was not possible for the current study because the Quechua-speaking adults were relatively unfamiliar with behavioral research and the methods that would have been required to solicit age of acquisition information.

8 References

Agwuele, A., Sussman, H. M., & Lindblom, B. (2008). The Effect of Speaking Rate onConsonant Vowel Coarticulation. *Phonetica*, *65*(4), 194–209. doi: 10.1159/000192792

Ambridge, B., Kidd, E., Rowland, C. F., & Theakston, A. L. (2015, March). The ubiquity of frequency effects in first language acquisition. *Journal of Child Language*, *42*(02), 239–273. doi: 10.1017/S030500091400049X

Arnon, I. (2010). *Starting big: The role of multiword phrases* (Dissertation). Stanford University.

Arnon, I., & Christiansen, M. H. (2017). The role of multiword building blocks in explaining L1-L2 differences. *Topics in Cognitive Science*, *9*(3), 621–636. doi: 10.1111/tops.12271

Arnon, I., & Cohen Priva, U. (2013). More than words: The effect of multi-word frequency and constituency on phonetic duration. *Language and Speech*, *56*(3), 349–371. doi: 10.1177/0023830913484891

Aust, F., & Barth, M. (2018). *Papaja:Create APA manuscripts with R Markdown.*

Baayen, R. H. (1992). Quantitative aspects of morphological productivity. In G. Booij & J. van Marle (Eds.), *Yearbook of Morphology 1991* (pp. 109–149). Kluwer Academic Publishers.

Baayen, R. H., McQueen, J. M., Dijkstra, T., & Schreuder, R. (2003). Frequency effects in regular inflectional morphology: Revisiting Dutch plurals. In R. H. Baayen & R. Schreuder (Eds.), *Morphological Structure in Language Processing* (pp. 355–390). Berlin, New York: De Gruyter Mouton. doi: 10.1515/9783110910186.355

Bannard, C., & Matthews, D. (2008). Stored word sequences in language learning: The effect of familiarity on children's repetition of four-word combinations. *Psychological Science*, *19*(3), 241–248. doi: 10.1111/j.1467-9280.2008.02075.x

Barbier, G., Perrier, P., Ménard, L., Payan, Y., Tiede, M. K., & Perkell, J. S. (2013). Speech planning as an index of speech motor control maturity. In *Proceedings of Interspeech 2013.* Lyon, France.

Barbier, G., Perrier, P., Ménard, L., Payan, Y., Tiede, M. K., & Perkell, J. S. (2015). Speech planning in 4-year-old children versus adults: Acoustic and articulatory analyses. In *Proceedings of Interspeech 2015.* Dresden, Germany.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48.

Benson, C. (2004). Bilingual schooling in Mozambique and Bolivia: From experimentation to implementation. *Language Policy*, *3*(1), 47–66. doi: 10.1023/B:LPOL.0000017725.62093.66

Berko, J. (1958). The child's learning of English morphology. *Word*, *14*(2-3), 150–177. doi: 10.1080/00437956.1958.11659661

Birdson, D., Gertken, L. M., & Amengual, M. (2012). *Bilingual Language Profile: An Easy-to-Use Instrument to Assess Bilingualism.* COERLL, University of Texas at Austin.

Boersma, P., & Weenink, D. (2018). *Praat: Doing phonetics by computer.*

Bradlow, A. R. (2002). Confluent talker- and listener-oriented forces in clear speech production. In C. Gussenhoven & N. Warner (Eds.), *Laboratory Phonology* (Vol. 7, pp. 241–273). Berlin and New York: Mouton de Gruyter.

Brooks, M. E., Kristensen, K., van Benthem, K. J., Magnusson, A., Berg, C. W., Nielsen, A., ... Bolker, B. M. (2017). glmmTMB balances speed and flexibility among packages for zero-inflated generalized linear mixed modeling. *The R Journal*, *9*(2), 378–400.

Chen, W.-R., Whalen, D. H., & Shadle, C. H. (2019, May). F0-induced formant measurement errors result in biased variabilities. *The Journal of the Acoustical Society of America*, *145*(5), EL360-EL366. doi: 10.1121/1.5103195

Cho, T. (2001). Effects of Morpheme Boundaries on Intergestural Timing: Evidence from Korean. *Phonetica*, *58*(3), 129–162. doi: 10.1159/000056196

Chomsky, N. (1959). A review of BF Skinner's Verbal Behavior. *Language*, *35*(1), 26–58.

Cohen, C. (2014, November). Probabilistic reduction and probabilistic enhancement: Contextual and paradigmatic effects on morpheme pronunciation. *Morphology*, *24*(4), 291–323. doi: 10.1007/s11525-014-9243-y

Colé, P., Segui, J., & Taft, M. (1997, October). Words and Morphemes as Units for Lexical Access. *Journal of Memory and Language*, *37*(3), 312–330. doi: 10.1006/jmla.1997.2523

Courtney, E. H., & Saville-Troike, M. (2002). Learning to construct verbs in Navajo and Quechua. *Journal of Child Language*, *29*(03), 623–654. doi: 10.1017/S0305000902005160

Cychosz, M. (2018). *Cychosz HomeBank Corpus.*

Cychosz, M. (2020). *Word structure in early Quechua speech: Morphology and acoustic phonetics* (Open Science Framework Project).

Cychosz, M., Edwards, J. R., Munson, B., & Johnson, K. (2019, December). Spectral and temporal measures of coarticulation in child speech. *The Journal of the Acoustical Society of America-Express Letters*, *146*(6), EL516-EL522. doi: 10.1121/1.5139201

Davis, M., & Redford, M. A. (2019). The Emergence of Discrete Perceptual-Motor Units in a Production Model That Assumes Holistic Phonological Representations. *Frontiers in Psychology*, *10*(2121), 1–19. doi: 10.3389/fpsyg.2019.02121

Fenson, L., Marchman, V., Thal, D. J., Dale, P., Reznick, J., & Bates, E. (2007). *MacArthur-Bates Communicative Development Inventories User's Guide and Technical Manual* (2nd Edition ed.). San Diego, CA: Singular.

Flege, J. E. (1988, December). Anticipatory and carry-over nasal coarticulation in the speech of children and adults. *Journal of Speech Language and Hearing Research*, *31*, 525–536.

Fourakis, M. (1991, October). Tempo, stress, and vowel reduction in American English. *The Journal of the Acoustical Society of America*, *90*(4), 1816–1827. doi: 10.1121/1.401662

Fowler, C. A. (1981). A Relationship between Coarticulation and Compensatory Shortening. *Phonetica*, *38*(1-3), 35–50. doi: 10.1159/000260013

Gallagher, G. (2016). Vowel height allophony and dorsal place contrasts in Cochabamba Quechua. *Phonetica*, *73*(2), 101–119.

Gavrilov, I., & Pusev, R. (2014). *Normtest: Tests for Normality.*

Gay, T. (1981). Mechanisms in the control of speech rate. *Phonetica*, *38*, 148–158.

Gerosa, M., Lee, S., Giuliani, D., & Narayanan, S. (2006). Analyzing children's speech: An acoustic study of consonants and consonant-vowel transition. In *2006 IEEE International Conference on Acoustics Speed and Signal Processing Proceedings* (Vol. 1, pp. 393–396). Toulouse, France: IEEE. doi: 10.1109/ICASSP.2006.1660040

Goffman, L., Smith, A., Heisler, L., & Ho, M. (2008). The Breadth of Coarticulatory Units in Children and Adults. *Journal of Speech Language and Hearing Research*, *51*(6), 1424–1437. doi: 10.1044/1092-4388(2008/07-0020)

Goodell, E. W., & Studdert-Kennedy, M. (1992). Acoustic Evidence for the Development of Gestural Coordination in the Speech of 2-Year-Olds: A Longitudinal Study. *Haskins Laboratories Status Report on Speech Research*, *SR-111/1123*, 63–88.

Guy, G. R. (1991, March). Explanation in variable phonology: An exponential model of morphological constraints. *Language Variation and Change*, *3*(1), 1–22. doi: 10.1017/S0954394500000429

Hanique, I., & Ernestus, M. (2012). The role of morphology in acoustic reduction. *Lingue e linguaggio*, *11*(2), 147–164.

Harrington, J., Kleber, F., Reubold, U., & Siddins, J. (2015, January). The relationship between prosodic weakening and sound change: Evidence from the German tense/lax vowel contrast. *Laboratory Phonology*, *6*(1). doi: 10.1515/lp-2015-0002

Hay, J. (2003). *Causes and Consequences of Word Structure.* New York and London: Routledge.

Hlavac, M. (2018). *Stargazer: Well-Formatted Regression and Summary Statistics Tables.* Bratislava, Slovakia: Central European Labour Studies Institute (CELSI).

Hoff, E. (2003, October). The Specificity of Environmental Influence: Socioeconomic Status Affects Early Vocabulary Development Via Maternal Speech. *Child Development*, *74*(5), 1368–1378. doi: 10.1111/1467-8624.00612

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

Hornberger, N. H., & King, K. A. (1996, December). Language Revitalisation in the Andes: Can the Schools Reverse Language Shift? *Journal of Multilingual and Multicultural Development*, *17*(6), 427–441. doi: 10.1080/01434639608666294

Johnson, K. (2004). Massive reduction in conversational American English. In K. Yoneyama & K. Maekawa (Eds.), *Proceedings of the 1st session of the 10th International Symposium* (pp. 29–54). The National International Institute for Japanese Language.

Katz, W. F., Kripke, C., & Tallal, P. (1991, December). Anticipatory coarticulation in the speech of adults and young children: Acoustic, perceptual, and video data. *Journal of Speech Language and Hearing Research*, *34*, 1222–1232.

Kemps, R. J. J. K., Wurm, L. H., Ernestus, M., Schreuder, R., & Baayen, R. H. (2005, February). Prosodic cues for morphological complexity in Dutch and English. *Language and Cognitive Processes*, *20*(1-2), 43–73. doi: 10.1080/01690960444000223

Kent, R. D. (1983). The Segmental Organization of Speech. In P. F. MacNeilage (Ed.), *The Production of Speech* (pp. 57–89). New York, NY: Springer New York. doi: 10.1007/978-1-4613-8202-7_4

Koenig, J.-P., & Jurafsky, D. (1995). Type Underspecification and On-line Type Construction in the Lexicon. In R. Aranovich, W. Byrne, S. Preuss, & M. Senturia (Eds.), *Proceedings of the 13th West Coast Conference on Formal Linguistics* (pp. 270–285). Stanford, CA: CSLI Publications.

Koopmans-Van Beinum, F. J. (1980). *Vowel contrast reduction, an acoustic and perceptual study of Dutch vowels in various speech conditions.* Amsterdam: University of Amsterdam, Academische Per B.V.

Kuznetsova, A., Brockhoff, P., & Christensen, R. (2017). lmerTest Package: Tests in linear mixed-effects models. *Journal of Statistical Software*, *82*(13), 1–26.

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

Lee-Kim, S.-I., Davidson, L., & Hwang, S. (2013, January). Morphological effects on the darkness of English intervocalic /l/. *Laboratory Phonology*, *4*(2), 475–511. doi: 10.1515/lp-2013-0015

Lehiste, I. (1972, June). The Timing of Utterances and Linguistic Boundaries. *The Journal of the Acoustical Society of America*, *51*(6B), 2018–2024. doi: 10.1121/1.1913062

Lehiste, I., & Shockey, L. (1972). On the perception of coarticulation effects in English VCV syllables. *Journal of Speech and Hearing Research*, *15*(3), 500–506.

Lieven, E. V. M., Pine, J. M., & Baldwin, G. (1997, February). Lexically-based learning and early grammatical development. *Journal of Child Language*, *24*(1), 187–219. doi: 10.1017/S0305000996002930

Lo, S., & Andrews, S. (2015, August). To transform or not to transform: Using generalized linear mixed models to analyse reaction time data. *Frontiers in Psychology*, *6*(1171), 1–16. doi: 10.3389/fpsyg.2015.01171

Lorge, I., & Chall, J. (1963, December). Estimating the Size of Vocabularies of Children and Adults: An Analysis of Methodological Issues. *The Journal of Experimental Education*, *32*(2), 147–157. doi: 10.1080/00220973.1963.11010819

Losiewicz, B. L. (1995). Word frequency effects on the acoustic duration of morphemes. *Journal of the Acoustical Society of America*, *97*(5), 3243–3243. doi: 10.1121/1.411745

Matthies, M., Perrier, P., Perkell, J., & Zandipour, M. (2001). Variation in Anticipatory Coarticulation With Changes in Clarity and Rate. *Journal of Speech Language and Hearing Research*, *44*(2), 340–353.

McAuliffe, M., Socolof, M., Mihuc, A., Wagner, M., & Sonderegger, M. (2017). *Montreal Forced Aligner.*

McFee, B., Raffel, C., Liang, D., Ellis, D., McVicar, M., Battenberg, E., & Nieto, O. (2015). Librosa: Audio and Music Signal Analysis in Python. In *Proceedings of the 14th Python in science conference* (pp. 18–24). Austin, Texas. doi: 10.25080/Majora-7b98e3ed-003

Morrison, C. M., Ellis, A. W., & Quinlan, P. T. (1992, November). Age of acquisition, not word frequency, affects object naming, not object recognition. *Memory & Cognition*, *20*(6), 705–714. doi: 10.3758/BF03202720

Mousikou, P., Strycharczuk, P., Turk, A., Rastle, K., & Scobbie, J. M. (2015). Morphological effects on pronunciation. In *Proceedings of the 18th ICPhS.* Glasgow, Scotland.

Munhall, K., Fowler, C., Hawkins, S., & Saltzman, E. (1992, April). "Compensatory shortening" in monosyllables of spoken English. *Journal of Phonetics*, *20*(2), 225–239. doi: 10.1016/S0095-4470(19)30624-2

Muysken, P. (2012a). Contacts between indigenous languages in South America. In L. Campbell & V. Grondona (Eds.), *The indigenous languages of South America: A comprehensive guide* (pp. 235–258). Berlin, Germany: Walter de Gruyter.

Muysken, P. (2012b). Spanish affixes in the Quechua languages: A multidimensional perspective. *Lingua*, *122*(5), 481–493.

Nijland, L., Maassen, B., der Meulen, S. V., Gabreëls, F., Kraaimaat, F. W., & Schreuder, R. (2002, January). Coarticulation patterns in children with developmental apraxia of speech. *Clinical Linguistics & Phonetics*, *16*(6), 461–483. doi: 10.1080/02699200210159103

Nittrouer, S., Studdert-Kennedy, M., & McGowan, R. S. (1989). The emergence of phonetic segments: Evidence from the spectral structure of fricative-vowel syllables spoken by children and adults. *Journal of Speech Language and Hearing Research*, *32*, 120–132.

Nittrouer, S., Studdert-Kennedy, M., & Neely, S. T. (1996). How Children Learn to Organize Their Speech Gestures: Further Evidence From Fricative-Vowel Syllables. *Journal of Speech Language and Hearing Research*, *39*, 379–389. doi: 10.1044/jshr.3902.379

Noiray, A., Abakarova, D., Rubertus, E., Krüger, S., & Tiede, M. (2018, June). How Do Children Organize Their Speech in the First Years of Life? Insight From Ultrasound Imaging. *Journal of Speech, Language, and Hearing Research*, *61*(6), 1355–1368. doi: 10.1044/2018_JSLHR-S-17-0148

Noiray, A., Ménard, L., & Iskarous, K. (2013, January). The development of motor synergies in children: Ultrasound and acoustic measurements. *Journal of the Acoustical Society of America*, *133*(1), 444–452. doi: 10.1121/1.4763983

Noiray, A., Popescu, A., Killmer, H., Rubertus, E., Krüger, S., & Hintermeier, L. (2019). Spoken Language Development and the Challenge of Skill Integration. *Frontiers in Psychology*, *10*(2777), 1–17. doi: 10.3389/fpsyg.2019.02777

Noiray, A., Wieling, M., Abakarova, D., Rubertus, E., & Tiede, M. (2019, August). Back From the Future: Nonlinear Anticipation in Adults' and Children's Speech. *Journal of Speech, Language, and Hearing Research*, *62*(8S), 3033–3054. doi: 10.1044/2019_JSLHR-S-CSMC7-18-0208

O'Donnell, T. J. (2015). *Productivity and Reuse In Language: A Theory of Linguistic Computation and Storage*. The MIT Press. doi: 10.7551/mitpress/9780262028844.001.0001

Öhman, S. E. G. (1966, January). Coarticulation in VCV Utterances: Spectrographic Measurements. *The Journal of the Acoustical Society of America*, *39*(1), 151–168. doi: 10.1121/1.1909864

Pace, A., Luo, R., Hirsh-Pasek, K., & Golinkoff, R. M. (2017, January). Identifying Pathways Between Socioeconomic Status and Language Development. *Annual Review of Linguistics*, *3*(1), 285–308. doi: 10.1146/annurev-linguistics-011516-034226

Pierrehumbert, J. B. (2016). Phonological representation: Beyond abstract versus episodic. *Annual Review of Linguistics*, *2*, 33–52.

Pinker, S., & Prince, A. (1994). Regular and irregular morphology and the psychological status of rules of grammar. In S. D. Lima, R. L. Corrigan, & G. K. Iverson (Eds.), *The Reality of Linguistic Rules* (pp. 321–351). Amsterdam/Philadelphia: John Benjamins Publishing Company.

Pinker, S., & Ullman, M. T. (2002, November). The past and future of the past tense. *Trends in Cognitive Sciences*, *6*(11), 456–463. doi: 10.1016/S1364-6613(02)01990-3

Plag, I. (2014). Phonological and phonetic variability in complex words: An uncharted territory. *Rivista di Linguistica*, *26*(2), 209–228.

Plag, I., & Baayen, R. H. (2009). Suffix Ordering and Morphological Processing. *Language*, *85*(1), 109–152. doi: 10.1353/lan.0.0087

Plag, I., Homann, J., & Kunter, G. (2017, February). Homophony and morphology: The acoustics of word-final S in English. *Journal of Linguistics*, *53*(01), 181–216. doi: 10.1017/S0022226715000183

Pluymaekers, M., Ernestus, M., Baayen, R. H., & Booij, G. (2010, January). Morphological effects on fine phonetic detail: The case of Dutch -igheid. In *Laboratory Phonology* (Vol. 10, pp. 511–531). doi: 10.1515/9783110224917.5.511

Redford, M. A. (2018, June). Grammatical Word Production Across Metrical Contexts in School-Aged Children's and Adults' Speech. *Journal of Speech, Language, and Hearing Research*, *61*(6), 1339–1354. doi: 10.1044/2018_JSLHR-S-17-0126

Redford, M. A. (2019, August). Speech Production From a Developmental Perspective. *Journal of Speech, Language, and Hearing Research*, *62*(8S), 2946–2962. doi: 10.1044/2019_JSLHR-S-CSMC7-18-0130

Reidy, P. F., Kristensen, K., Winn, M. B., Litovsky, R. Y., & Edwards, J. R. (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

Repp, B. H. (1986, May). Some observations on the development of anticipatory coarticulation. *Journal of the Acoustical Society of America*, *79*(5), 1616–1619. doi: 10.1121/1.393298

Rubertus, E., & Noiray, A. (2018). On the development of gestural organization: A cross-sectional study of vowel-to-vowel anticipatory coarticulation. *PLOS ONE*, *13*(9), 1–21. doi: 10.1371/journal.pone.0203562

Sereno, J. A., Baum, S. R., Marean, G. C., & Lieberman, P. (1987). Acoustic analyses and perceptual data on anticipatory labial. *Journal of the Acoustical Society of America*, *81*(2), 512–519.

Sereno, J. A., & Lieberman, P. (1987). Developmental aspects of lingual coarticulation. *Journal of Phonetics*, *15*, 247–257.

Seyfarth, S., Garellek, M., Gillingham, G., Ackerman, F., & Malouf, R. (2018, January). Acoustic differences in morphologically-distinct homophones. *Language, Cognition and Neuroscience*, *33*(1), 32–49. doi: 10.1080/23273798.2017.1359634

Shapiro, S. S., Wilk, M. B., & Chen, H. J. (1968, December). A Comparative Study of Various Tests for Normality. *Journal of the American Statistical Association*, *63*(324), 1343–1372. doi: 10.1080/01621459.1968.10480932

Skinner, B. F. (1957). *Verbal Behavior.* Acton, MA: Copley Publishing Group.

Smith, B. L. (1992, April). Relationships between duration and temporal variability in children's speech. *The Journal of the Acoustical Society of America*, *91*(4), 2165–2174. doi: 10.1121/1.403675

Smith, R., Baker, R., & Hawkins, S. (2012, September). Phonetic detail that distinguishes prefixed from pseudo-prefixed words. *Journal of Phonetics*, *40*(5), 689–705. doi: 10.1016/j.wocn.2012.04.002

Song, J. Y., Demuth, K., Evans, K., & Shattuck-Hufnagel, S. (2013, May). Durational cues to fricative codas in 2-year-olds' American English: Voicing and morphemic factors. *Journal of the Acoustical Society of America*, *133*(5), 2931–2946. doi: 10.1121/1.4795772

Song, J. Y., Demuth, K., Shattuck-Hufnagel, S., & Ménard, L. (2013, May). 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*, 281–295. doi: 10.1016/j.wocn.2013.03.004

Strycharczuk, P. (2019). Phonetic Detail and Phonetic Gradience in Morphological Processes. In *Oxford Research Encyclopedia of Linguistics* (pp. 1–25). Oxford University Press. doi: 10.1093/acrefore/9780199384655.013.616

Strycharczuk, P., & Scobbie, J. M. (2016, November). Gradual or abrupt? The phonetic path to morphologisation. *Journal of Phonetics*, *59*, 76–91. doi: 10.1016/j.wocn.2016.09.003

Sugahara, M., & Turk, A. (2009). Durational correlates of English sublexical constituent structure. *Phonology*, *26*(03), 477–524. doi: 10.1017/S0952675709990248

Swingley, D., & Aslin, R. N. (2002, September). Lexical Neighborhoods and the Word-Form Representations of 14-Month-Olds. *Psychological Science*, *13*(5), 480–484. doi: 10.1111/1467-9280.00485

Taft, M., & Forster, K. I. (1976). Lexical storage and retrieval of polymorphemic and polysyllabic words. *Journal of Verbal Learning and Verbal Behavior*, *15*(6), 607–620.

Team, R. (2020). *RStudio: Integrated Development for R.* Boston, MA: RStudio, Inc.

Tomaschek, F., Tucker, B. V., Ramscar, M., & Baayen, R. H. (2019, July). *How is anticipatory coarticulation of suffixes affected by lexical proficiency?* (Preprint). PsyArXiv. doi: 10.31234/osf.io/gv89j

Torero, O. (1964). Los dialectos Quechuas. In *Anales Científicos de la Universidad Agraria* (Vol. 2, pp. 446–478). Lima, Peru.

Tsao, Y.-C., Weismer, G., & Iqbal, K. (2006). The effect of intertalker speech rate variation on acoustic vowel space. *Journal of the Acoustical Society of America*, *119*(2), 1074–1082.

Warner, N., Good, E., Jongman, A., & Sereno, J. (2006, April). Orthographic vs. morphological incomplete neutralization effects. *Journal of Phonetics*, *34*(2), 285–293. doi: 10.1016/j.wocn.2004.11.003

Wexler, K. (1998, December). Very early parameter setting and the unique checking constraint: A new explanation of the optional infinitive stage. *Lingua*, *106*, 23–79. doi: 10.1016/S0024-3841(98)00029-1

Whalen, D. H. (1990). Coarticulation is largely planned. *Journal of Phonetics*, *18*, 3–35.

Whiteside, S., & Hodgson, C. (2000, November). Speech patterns of children and adults elicited via a picture-naming task: An acoustic study. *Speech Communication*, *32*(4), 267–285. doi: 10.1016/S0167-6393(00)00013-3

Wickham, H. (2016). *Ggplot2: Elegant Graphics for Data Analysis.* New York: Springer-Verlag New York.

Zharkova, N., Hardcastle, W. J., & Gibbon, F. E. (2018, September). The dynamics of voiceless sibilant fricative production in children between 7 and 13 years old: An ultrasound and acoustic study. *The Journal of the Acoustical Society of America*, *144*(3), 1454–1466. doi: 10.1121/1.5053585

Zharkova, N., Hewlett, N., & Hardcastle, W. J. (2011, January). Coarticulation as an Indicator of Speech Motor Control Development in Children: An Ultrasound Study. *Motor Control*, *15*(1), 118–140. doi: 10.1123/mcj.15.1.118

Zharkova, N., Hewlett, N., Hardcastle, W. J., & Lickley, R. J. (2014, April). Spatial and Temporal Lingual Coarticulation and Motor Control in Preadolescents. *Journal of Speech Language and Hearing Research*, *57*(2), 374–388. doi: 10.1044/2014_JSLHR-S-11-0350

Zingler, T. (2018, October). Reduction without fusion: Grammaticalization and wordhood in Turkish. *Folia Linguistica*, *52*(2), 415–447. doi: 10.1515/flin-2018-0011

Appendix A

South Bolivian Quechua Consonant Inventory

	Bilabial	Dental	Postalveolar	Palatal	Velar	Uvular	Glottal
Plosive	p	t	ʧ		k	q	
Aspirated	p^h	t^h	$ʧ^h$		k^h	q^h	
Ejective	p’	t’	ʧ’		k’	q’	
Nasal	m	n		ɲ			
Fricative		s					h
Tap		ɾ					
Approximant	w		ʎ	j			
Lateral		l					

Appendix B

Stimuli used in real word repetition tasks

Real word*	Translation
'warmi	‘woman’ *training trial*
'wasi	‘house’ *training trial*
'qhari	‘man’ *training trial*
'chita	‘sheep’
'p’esqo	‘bird’
ju'k’ucha	‘mouse’
'waka	‘cow’
'wallpa	‘chicken’
'mama	‘mom’
'papa	‘potato’
't’ika	‘flower’
'llama	‘llama’
'cuca	‘coca (leaves)’
u'hut’a	‘sandal’
ham'piri	‘healer’
i'milla	‘girl’
'llapa	‘lightening’
'api	‘corn/citrus drink’
'ch’ulu	‘hat’
'punku	‘door’
'thapa	‘nest’
'punchu	‘poncho’
'pampa	‘prairie’
'sunkha	‘beard’
hatun'mama	‘grandma’
'wawa	‘baby/child’
'runtu	‘egg’
'qolqe	‘money’
'q’apa	‘palm of hand’
'alqo	‘dog’
'q’epi	‘bundle’

* For the real words, ' indicates stress, ’ indicates ejective

Appendix C

Real word repetition stimuli to elicit [am]

Real word*	Translation	Morpheme environment†
chi'ta-man	'sheep-ALL'	across
cu'ca-man	'coca (leaves)-ALL'	across
hatunma'ma-man	'grandma-ALL'	across
imi'lla-man	'girl-ALL'	across
juk'u'cha-man	'mouse-ALL'	across
lla'ma-man	'llama-ALL'	across
lla'pa-man	'lightening-ALL'	across
ma'ma-man	'mom-ALL'	across
pam'pa-man	'prairie-ALL'	across
pa'pa-man	'potato-ALL'	across
q'a'pa-man	'palm of hand-ALL'	across
sun'kha-man	'beard-ALL'	across
t'i'ka-man	'flower-ALL'	across
tha'pa-man	'nest-ALL'	across
wa'ka-man	'cow-ALL'	across
wall'pa-man	'chicken-ALL'	across
wa'wa-man	'baby/child-ALL'	across
'mama	'mom'	within
'llama	'llama'	within
ham'piri	'healer'	within
hampi'ri-pi	'healer-LOC'	within
'pampa	'prairie'	within
hatun'mama	'grandma'	within

* ' indicates stress, ' indicates ejective

Appendix D

Model predicting coarticulation in children and adults

tabular

	estimate	S.E.	z.statistic	p.value	95% CI
Intercept	2.14	0.04	59.61	0.00	2.21,2.07
Sequence duration scaled	0.00	0.00	3.65	0.00	0,0
VC sequenceap	0.73	0.05	16.20	0.00	0.82,0.64
Environment:across morpheme	-0.05	0.07	-0.67	0.50	0.09,-0.19
Sequence duration*VC sequence:ap	0.02	0.03	0.57	0.57	0.09,-0.05
Sequence duration* VC sequence:ap	0.00	0.00	-0.63	0.53	0,0
Sequence duration* Age:adult	0.00	0.00	2.66	0.01	0,0
VC sequence:ap*Age:adult	-0.10	0.06	-1.71	0.09	0.01,-0.21
Sequence duration* Environment:across morpheme	0.00	0.00	2.37	0.02	0,0
VC sequence:ap* Environment:across morpheme	-0.06	0.05	-1.17	0.24	0.04,-0.17
Age:adult* Environment:across morpheme	-0.09	0.05	-1.74	0.08	0.01,-0.2
Sequence duration* VC sequence:ap*Age:adult	0.00	0.00	-1.25	0.21	0,0
Sequence duration scaled* VC sequence:ap:Environmentacross morpheme	0.00	0.00	-2.25	0.02	0,0
Sequence duration*Age:adult* Environment:across morpheme	0.00	0.00	-3.36	0.00	0,0
VC sequence:ap*A ge:adult*Environment:across morpheme	0.16	0.07	2.31	0.02	0.29,0.02
Sequence duration* VC sequence:ap*Age:adult*Environment:across morpheme	0.00	0.00	3.44	0.00	0,0

Appendix E

Coarticulation by sequence duration, word, and morphological environment in children

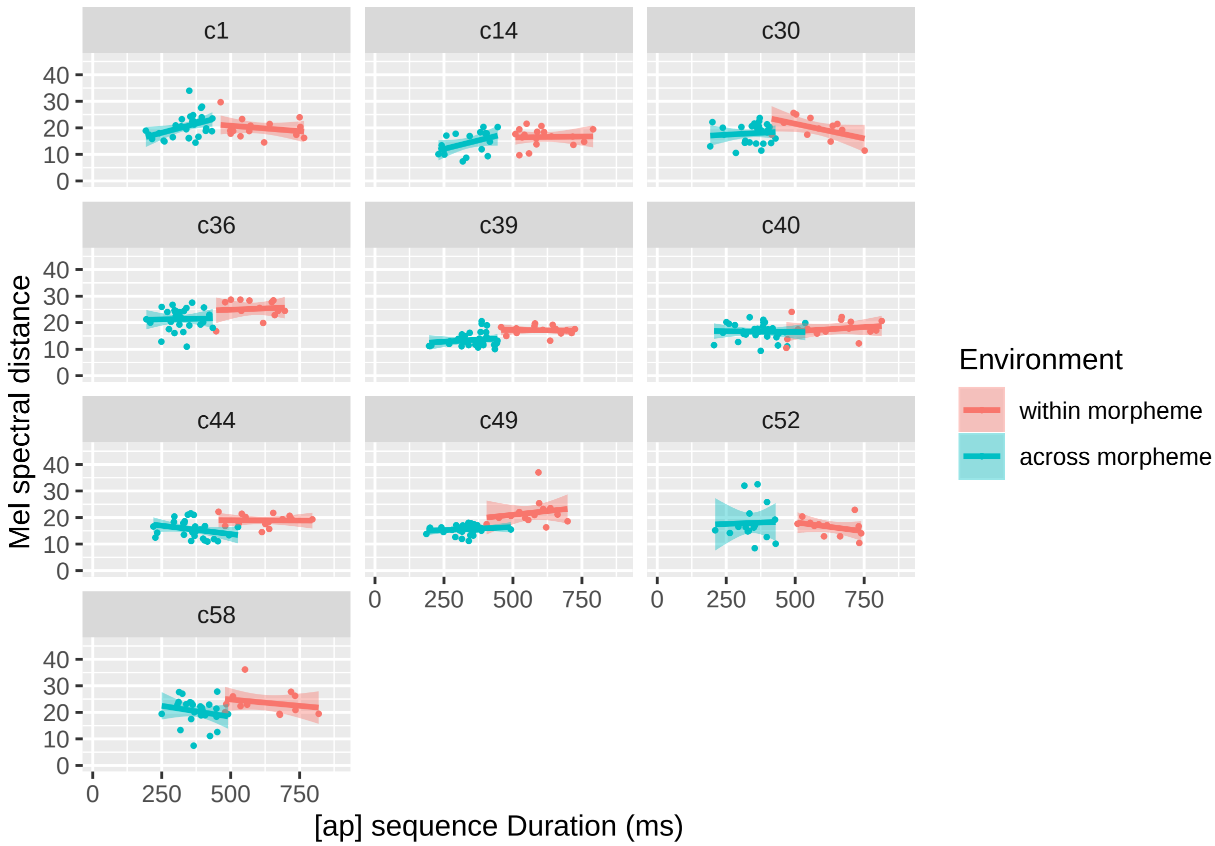

Figure E1. Coarticulation by [ap] duration, word, and morphological environment in five-year-old children

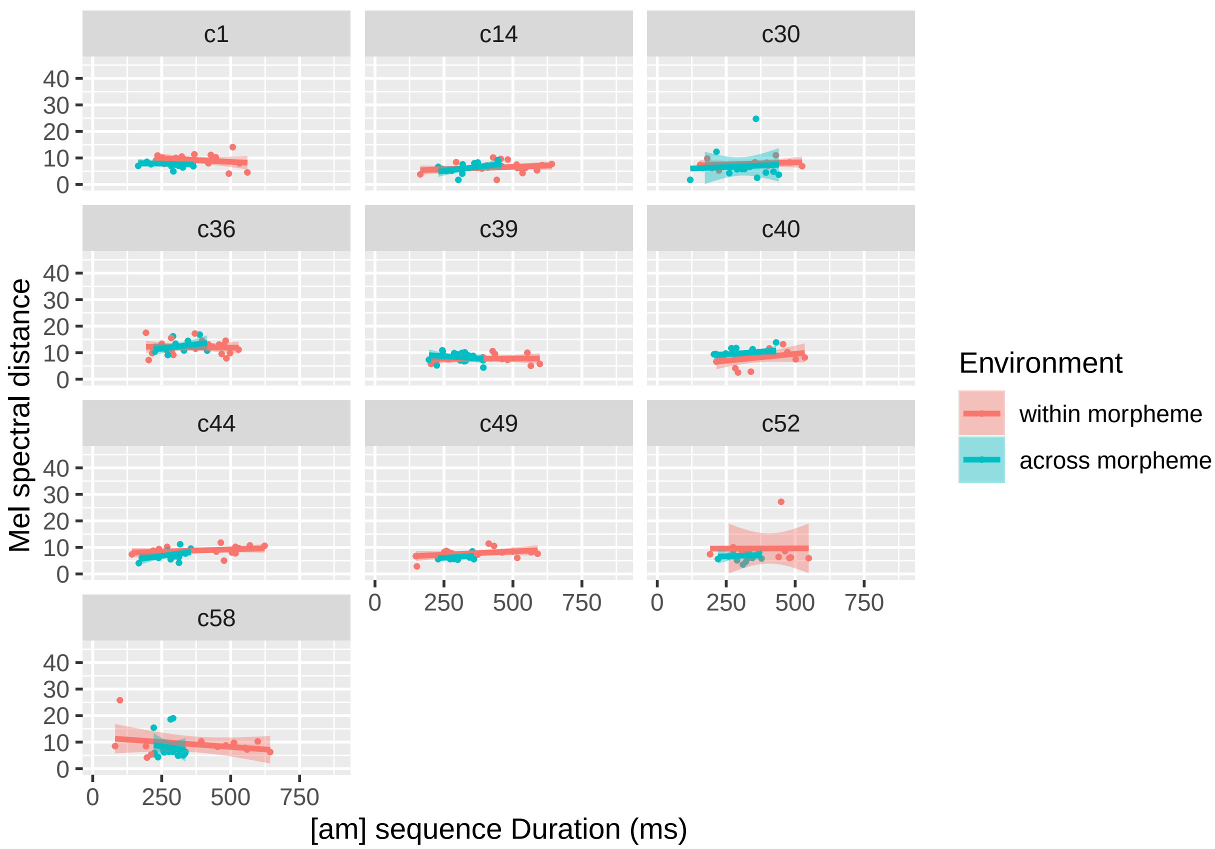

Figure E2. Coarticulation by [am] duration, word, and morphological environment in five-year-old children

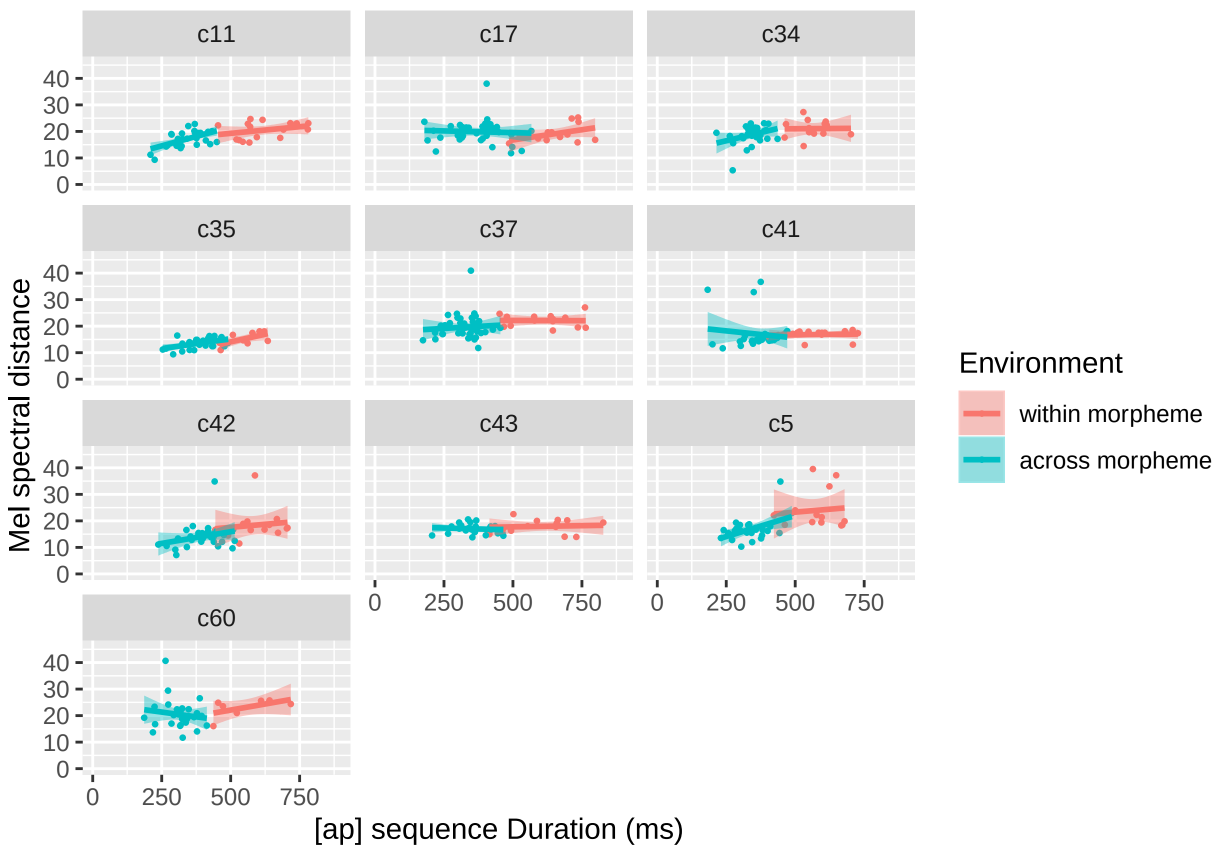

Figure E3. Coarticulation by [ap] duration, word, and morphological environment in six-year-old children

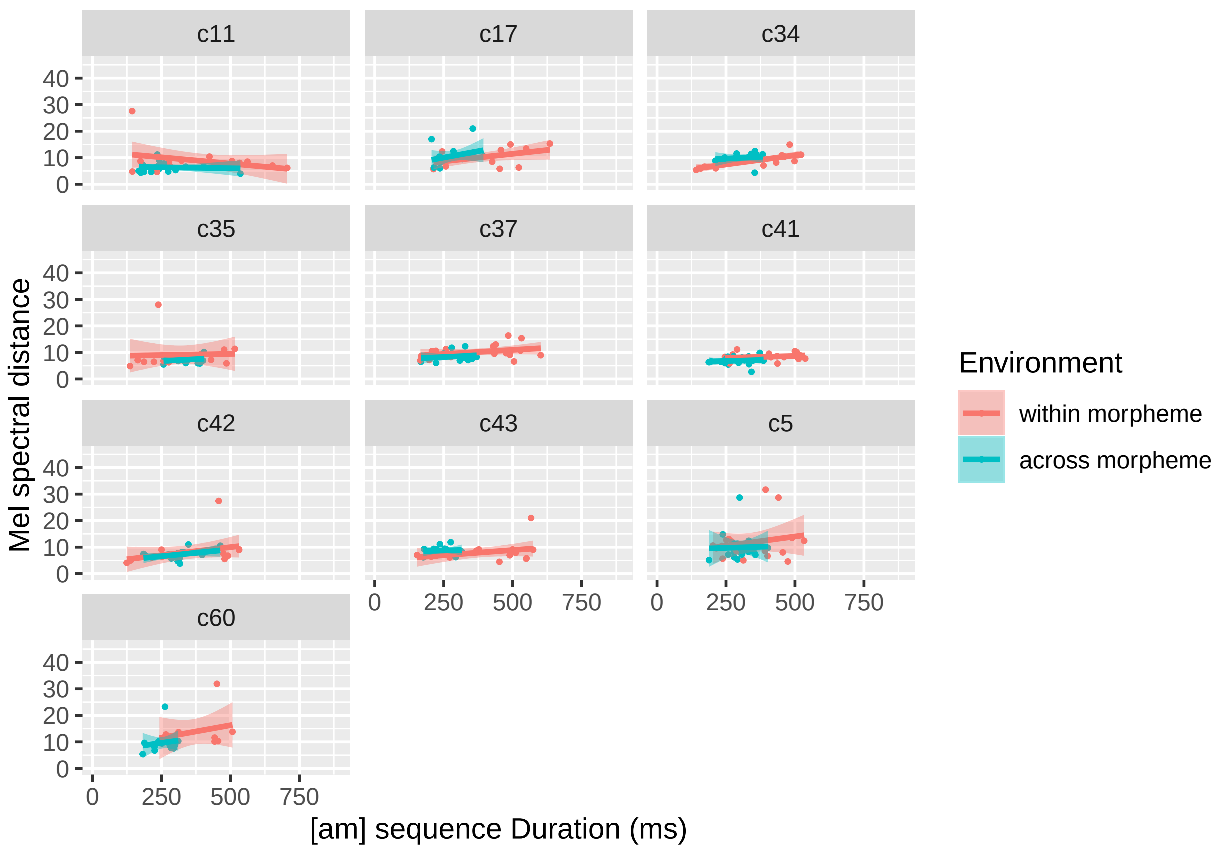

Figure E4. Coarticulation by [am] duration, word, and morphological environment in six-year-old children

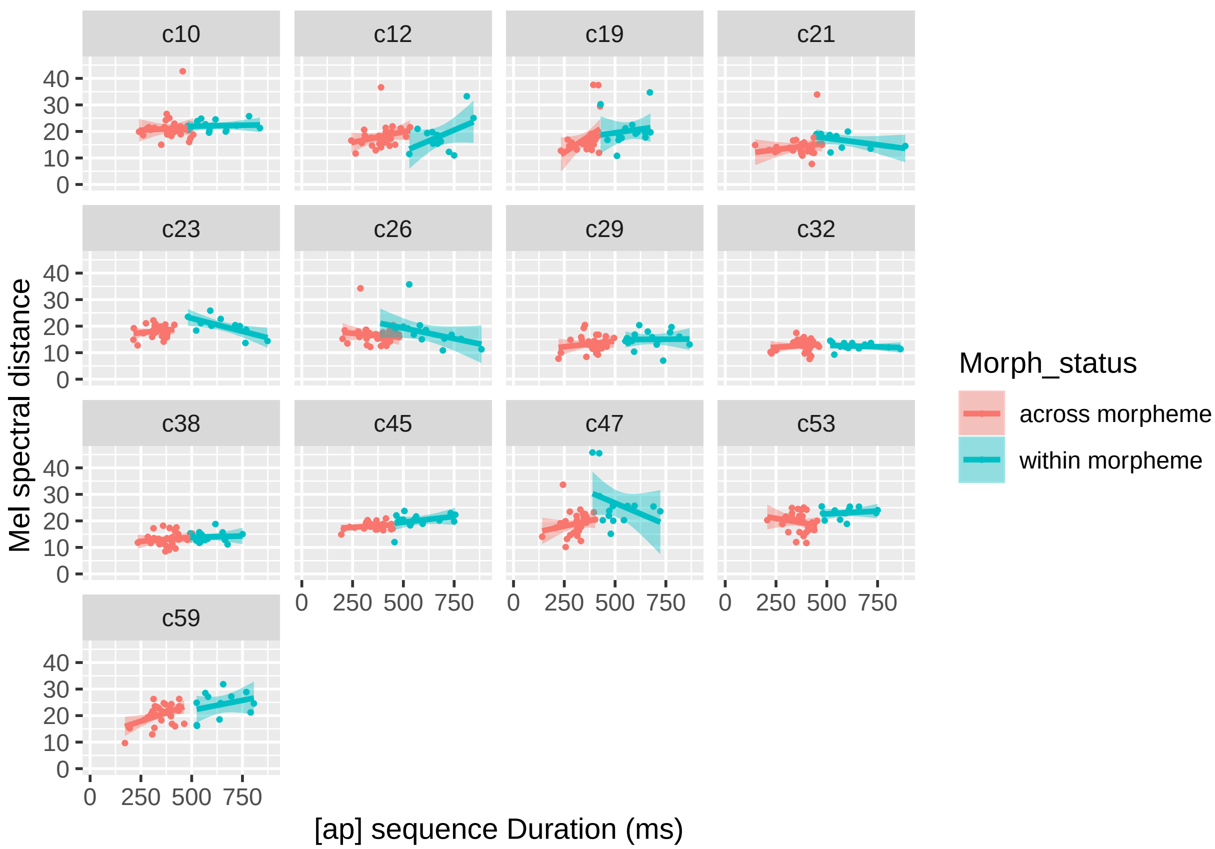

Figure E5. Coarticulation by [ap] duration, word, and morphological environment in seven-year-old children

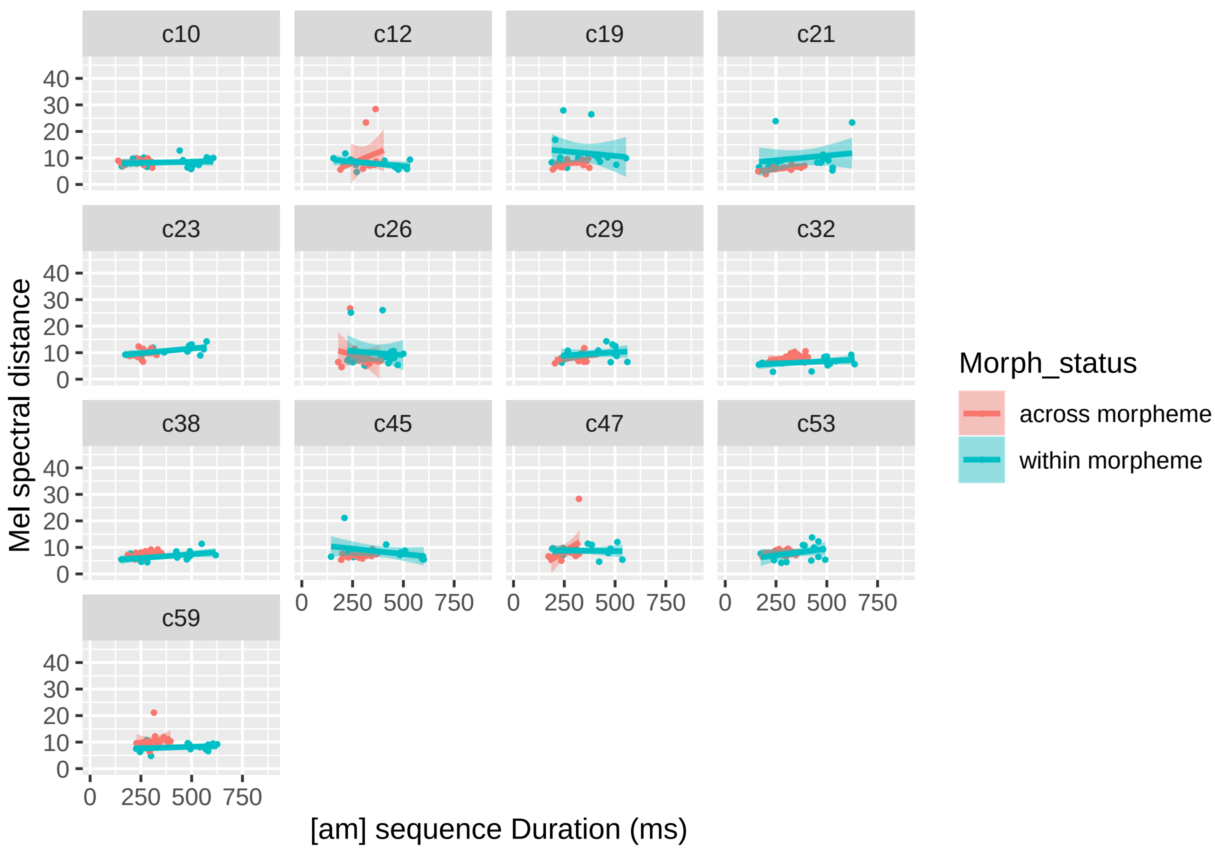

Figure E6. Coarticulation by [am] duration, word, and morphological environment in seven-year-old children

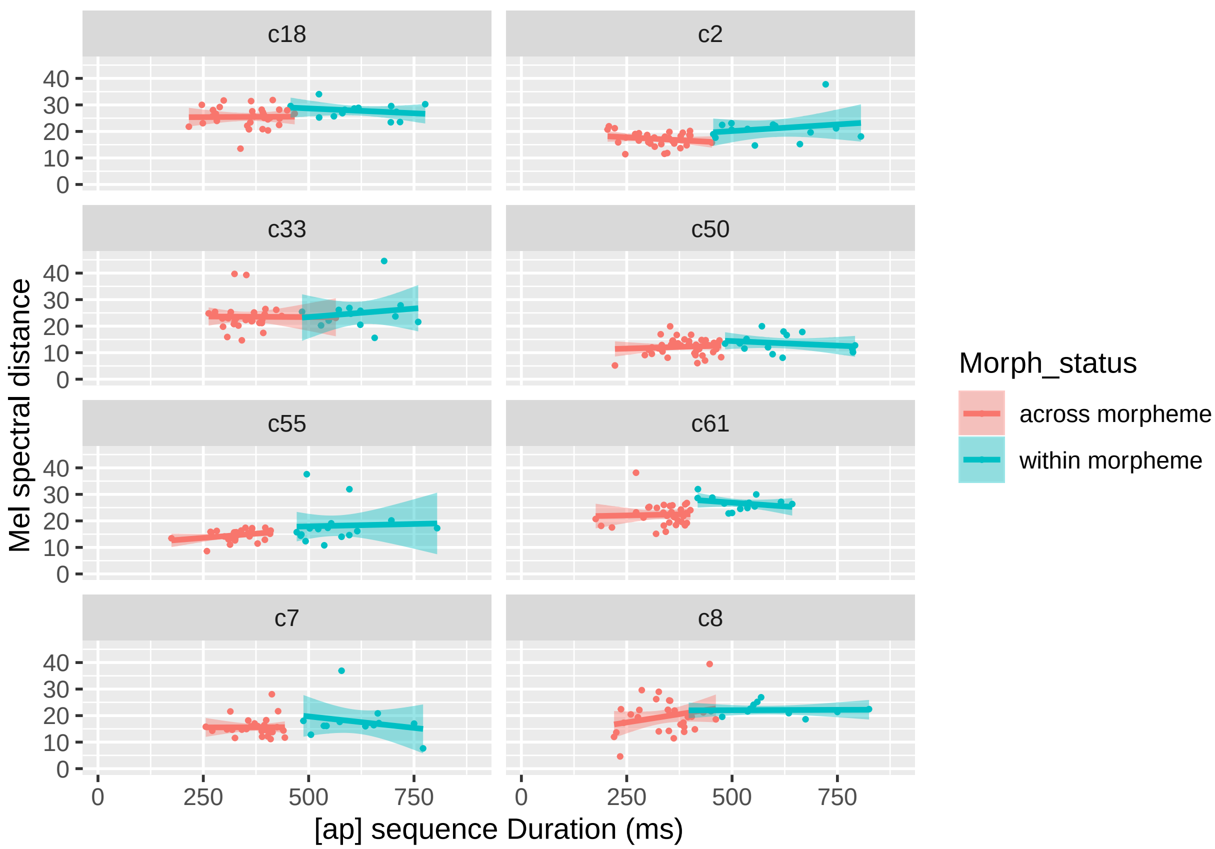

Figure E7. Coarticulation by [ap] duration, word, and morphological environment in eight-year-old children

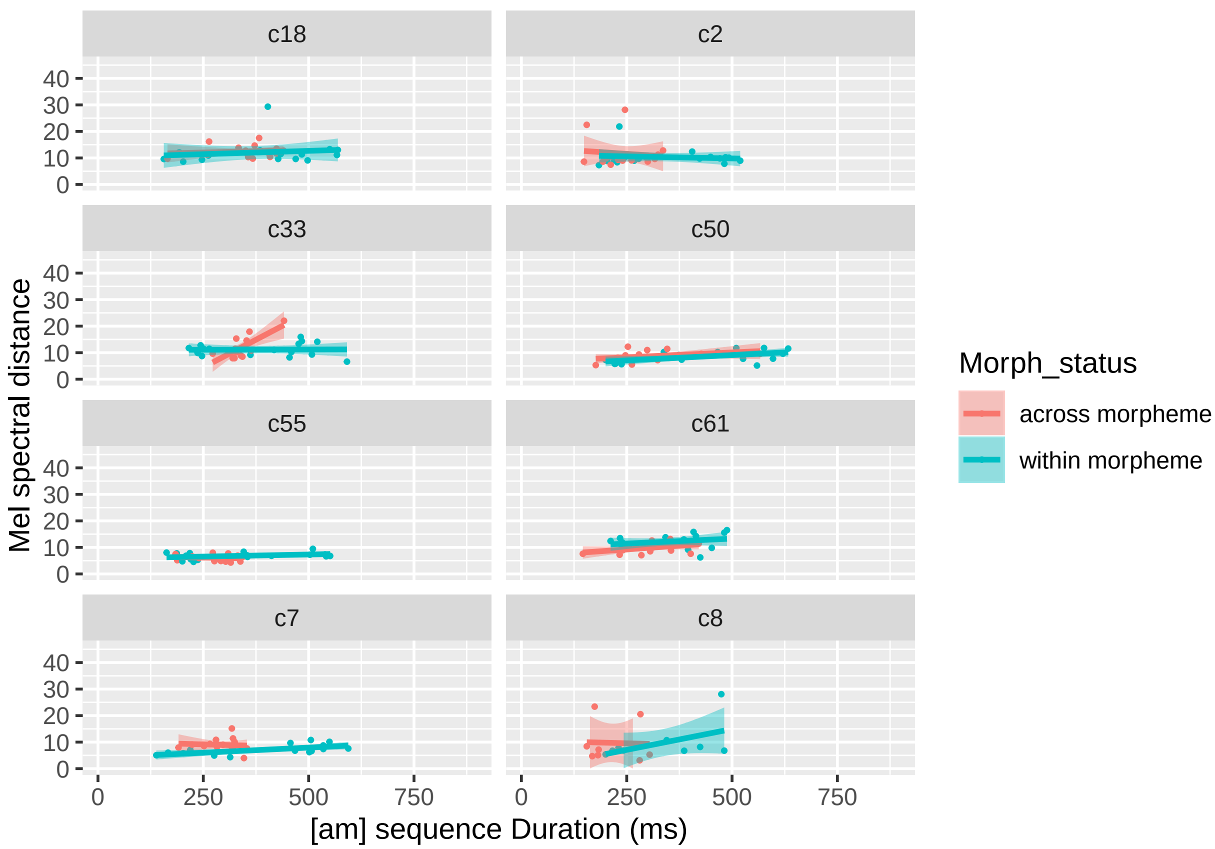

Figure E8. Coarticulation by [am] duration, word, and morphological environment in eight-year-old children

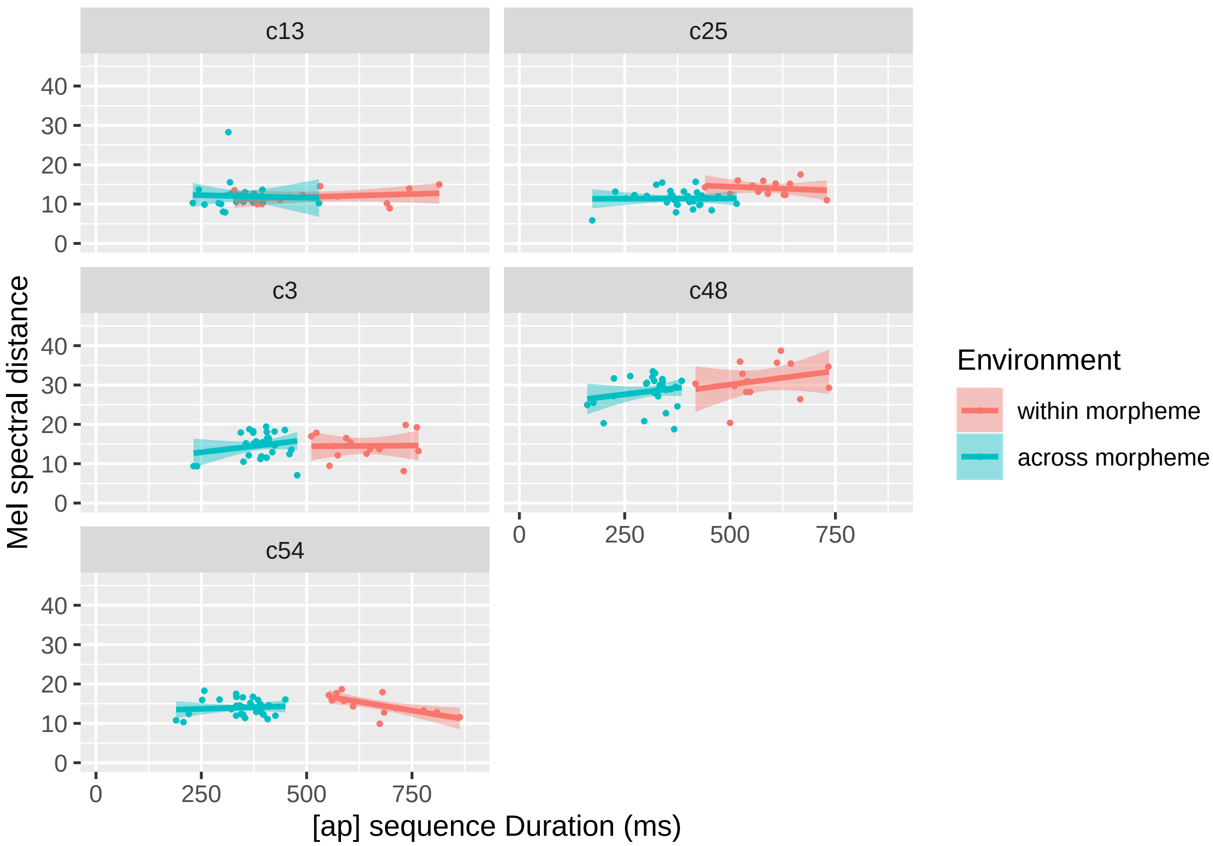

Figure E9. Coarticulation by [ap] duration, word, and morphological environment in nine-year-old children

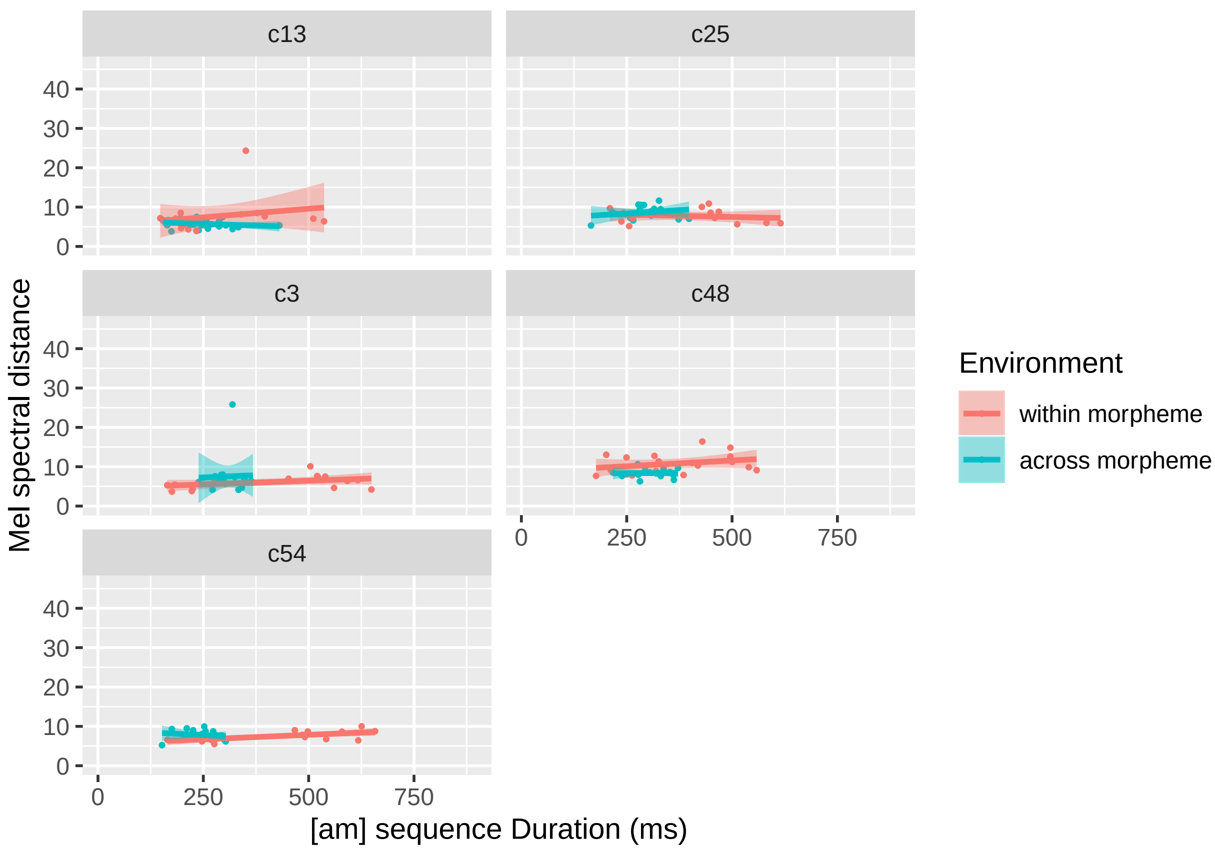

Figure E10. Coarticulation by [am] duration, word, and morphological environment in nine-year-old children

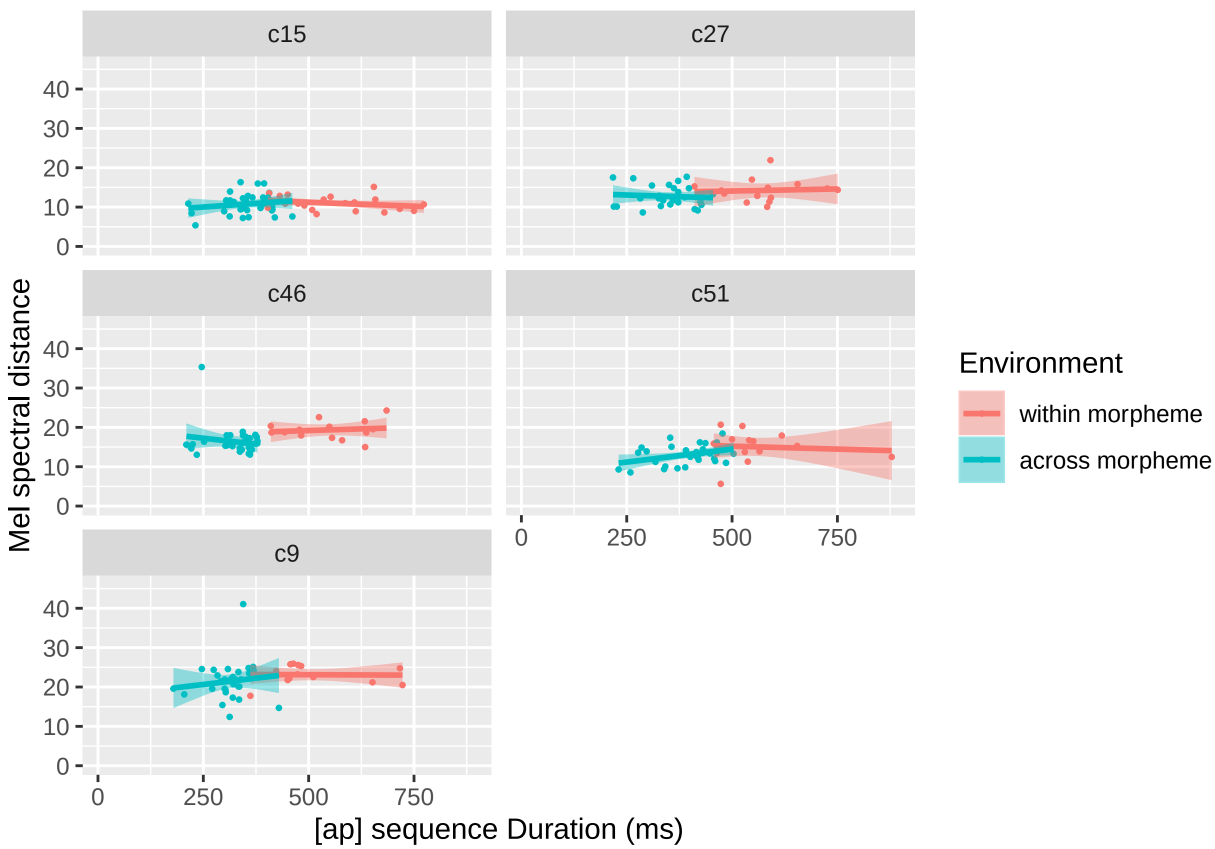

Figure E11. Coarticulation by [ap] duration, word, and morphological environment in ten-year-old children

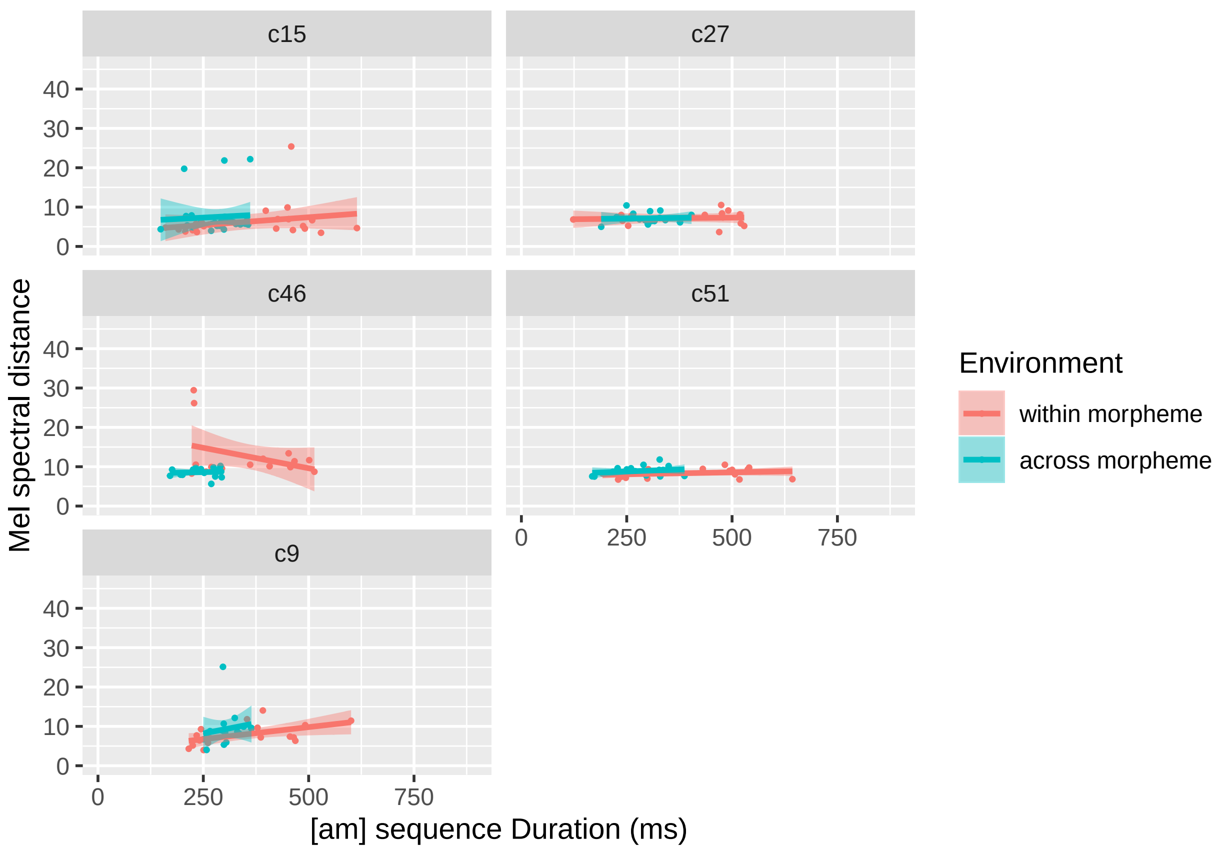

Figure E12. Coarticulation by [am] duration, word, and morphological environment in ten-year-old children