

# UC Berkeley

## UC Berkeley PhonLab Annual Report

### Title

Holistic Lexical Storage: Coarticulatory Evidence from Child Speech

### Permalink

<https://escholarship.org/uc/item/1j22q02m>

### Journal

UC Berkeley PhonLab Annual Report, 14(1)

### Author

Cychosz, Meg

### Publication Date

2018

### DOI

10.5070/P7141042477

### Copyright Information

Copyright 2018 by the author(s). All rights reserved unless otherwise indicated. Contact the author(s) for any necessary permissions. Learn more at <https://escholarship.org/terms>

Peer reviewed

HOLISTIC LEXICAL STORAGE:  
COARTICULATORY EVIDENCE FROM CHILD SPEECH

MEG CYCHOSZ

**Abstract**

Adult speakers readily decompose morphologically-complex words into their component parts. Overgeneralizations in children's early speech (e.g. *goed*) demonstrate that they must share in this ability. However, acoustic evidence from child speech suggests that children do not always break words down, instead storing language in more holistic chunks such as syllables or even entire words.

How are morphologically-complex forms represented throughout childhood? To answer this, we measured coarticulation within and between morphemes in adult and child (age 5-10) South Bolivian Quechua speakers. Coarticulation was quantified as the difference between the averaged Mel frequency cepstral coefficient vectors of the adjacent phones. Experiment 1 replicates known coarticulatory findings from the literature, demonstrating the validity of the MFCC measurement for calculating adjacent coarticulation. Experiment 2 then measures coarticulation between a single biphone sequence, [ap], in two environments: 1) within morphemes and 2) across morpheme boundaries. Results show that adult speakers coarticulated less across morpheme boundaries than within root morphemes. This is further evidence that adults decompose complex words. Children, however, coarticulated equally across and within morphemes. This suggests that the child speakers store inflected words more holistically than adults, even in this highly agglutinating language.

**Keywords:** coarticulation, acoustics, L1 acquisition, MFCCs, field phonetics, morphology

**1. Introduction**

Adult speakers efficiently compose novel words, freely converting nouns into verbs (e.g. *monetize*) and verbs into adjectives (e.g. *twinkly*). Since the Wug Test (Berko, 1958), developmental researchers have acknowledged that young children must share this morphological productivity with adults; if not, they would not be able to extend morpho-phonological patterns to novel word environments.

Despite these assumptions, it is likewise apparent that even for adult speakers, morphological decomposition varies by factors such as word frequency: both highly frequent complex words (Baayen, 1992) and derived words that are more frequent than the corresponding base form (e.g. *disentangle* vs. *entangle*) are less likely to be decomposed (Hay, 2003).

Furthermore, phonetic production data have called into question the nature of children's early lexical representations. A consistent, though not universal finding from this literature is that children coarticulate between adjacent and near-adjacent phones more than adults (Goodell & Studdert-Kennedy, 1992; Nittrouer, Studdert-Kennedy, & McGowan, 1989; Zharkova, Hewlett, & Hardcastle, 2011; cf. Barbier et al., 2015). Children's greater coarticulation suggests that they organize speech more holistically, at the syllabic level or higher. Anticipatory coarticulation between adjacent phones then decreases as children age; this may represent the individuation and abstraction of language units.

One method of studying the nature of lexical storage is to examine the effects of morphological structure on speech production (Baker, Smith, & Hawkins, 2007; Cho, 2001; Pluymaekers, Ernestus, Baayen, & Booij, 2010). A morphophonetic relationship is apparent as, for example, English /l/ is darker in stem-final position (*coolest*) than affix-initial (*couples*) (Lee-Kim, Davidson, & Hwang, 2013). Acoustic cues provide implicit evidence of morphological decomposition in adult, and some evidence suggests, child speech (Song, Demuth, Evans, & Shattuck-Hufnagel, 2013; Song, Demuth, Shattuck-Hufnagel, & Ménard, 2013).

Here we bring these two lines of research – morphophonetic interplay and child coarticulation – together and employ Mel cepstral measurements to examine morphological (de)composition in adult and child speech. We measure this in South Bolivian Quechua (SBQ), a highly agglutinating language with over 200 nominal and verbal suffixes. Coarticulation was measured between the adjacent phones [a] and [p] carried in words from one of two contexts: within a root morpheme (e.g. *papa* 'potato') or at a morpheme boundary (e.g. *llama-pi* 'llama-LOC').

Child coarticulation results suggest that word representations initially develop more holistically. If this is applied to SBQ, we predict that child speakers do not decompose complex words. Children would then coarticulate more than adults at morpheme boundaries (e.g. *llama-pi* 'llama-LOC'), indicating holistic lexical storage. Adults would coarticulate less across boundaries (e.g. *llama-pi* 'llama-LOC') than within morphemes (e.g. *papa* 'potato'), suggesting that they have learned to parse speech segmentally and store it morpho-phonemically. We can tell that this comes from experience because adults' phonetic realization of the *exact same* [ap] sequence differs within a morpheme versus across a boundary.

## **2.2 Background**

### **2.2.1 Storing complex words**

Adult speakers' ability to compose novel, morphologically-complex words (e.g. *twinkly*) appears to be highly flexible and abstract. Children's early overgeneralization errors (e.g. *I taked*) demonstrate how even very young speakers can analyze the internal structure of words and apply morphemes in new, unheard environments.

Yet there is evidence that adult speakers do not uniformly decompose complex words. For example, Baayen (1992) demonstrated that frequent words may instead be accessed holistically. In a later extension of this finding, Hay (2003) argued for relative frequency effects: derived words that are more frequent than their corresponding base form (e.g. *disentangle* vs. *entangle*) are likely accessed holistically and not decomposed (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 analytic languages (Baker et al., 2007). And less frequent words (Kemps, Wurm, Ernestus, Schreuder, & Baayen, 2005), or words that only ever occur with a given suffix, do not necessarily manifest the same pattern of morphological decomposition.

Findings that cast doubt upon word decomposition in adults are highly relevant for the study of child language. Conflicting evidence from child speech research over the past three decades has suggested that children may store speech forms more holistically than adults, at the syllabic level or higher. Perhaps, for extremely frequent collocations, speech could even be represented in chunks that transcend word boundaries (Arnon, 2010). Production evidence for this conclusion comes primarily from studies of children's anticipatory coarticulation patterns. Despite the logic behind a lack of coarticulation in children's speech – children speak slower and with less coordinated movement (Smith & Goffman, 1998) – many studies have concluded that children coarticulate more than adults (Goodell & Studdert-Kennedy, 1992; Nittrouer et al., 1989; Nittrouer, Studdert-Kennedy, & Neely, 1996; Zharkova, Hewlett, & Hardcastle, 2008; Zharkova et al., 2011).

Nittrouer et al. (1989) studied the coarticulation of fricative-vowel sequences in 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 strongly affected child fricative production. They concluded that this was because the children did not distinguish between /s/ and /ʃ/ as reliably as adults. It was not 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, I, 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., 2008, 2011; Zharkova, Hewlett, Hardcastle, & Lickley, 2014) have incorporated articulatory ultrasound data and have come to similar conclusions concerning child coarticulation patterns. 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 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 organization by preadolescence.

Citing acoustic and articulatory data, the above studies support theories of holistic, syllable-sized representations in early child phonology. According to such an interpretation, representations progressively individuate and become more adult-like with age. However, results from child coarticulation research are mixed. Elsewhere, findings suggest that child speech develops bottom-up and is more discretely organized into individual segments than adult phonology. This approach cites evidence that adults coarticulate more than children; children learn to coordinate and appropriately overlap as part of standard phonetic/phonological development. For example, Kent (1983) found evidence of anticipatory spectral movement for the /a/ of 'box' in adults but not four-year-old children. This suggested that coarticulation is something to be acquired as children construct their phonological representations. Whiteside & Hodgson (2000) also found that for children (6;0-10;0), formant transitions between CV and VC syllables increased with age, suggesting increased coarticulatory ability. Crucially, Whiteside and Hodgson elicited more naturalistic data, via a picture description task, instead of repetition elicitation of nonce syllables employed elsewhere.

Additional evidence that children's coarticulatory ability improves with age comes from articulatory ultrasound data: in  $V_1CV_2$  sequences, adults anticipate  $V_2$  in  $V_1$  more than children (Barbier et al., 2013, 2015). Finally, (Katz, Kripke, & Tallal, 1991) found almost no difference in coarticulation patterns by age. They report that in fricative-V syllables, adults anticipated the following vowel just slightly more than the children (3;0-8;0). Several other studies likewise could not find any differences between adult and child coarticulation (Flege, 1988; Goffman,

Smith, Heisler, & Ho, 2008; Noiray, Ménard, & Iskarous, 2013; Sereno, Baum, Marean, & Lieberman, 1987; Sereno & Lieberman, 1987).

A feature of many studies that took a bottom-up interpretation is that they examined long-distance coarticulation, or coarticulatory effects that spread across multiple segments. For example, while Repp (1986) had found that the young child in the study anticipated lip rounding in C/u/ sequences more than the older child, the author only found anticipatory F2 spectral movement in the /ə/ of ə#CV in the older child. This suggested that coarticulation developed with age. Both Barbier et al. (2013; 2015) and Goffman et al. (2008) looked at anticipatory vowel coarticulation across a V<sub>1</sub>CV<sub>2</sub> sequence and entire phrase, respectively, and accordingly found that children did not anticipate the vowel as much as the adults. Long-distance versus adjacent coarticulation likely does not entirely explain the different findings in this literature, however. Rubertus, Abakarova, Ries, & Noiray (2013) also examined long-distance anticipatory coarticulation in /V<sub>1</sub>CV<sub>2</sub>ə/ sequences, similar to those used in Barbier (2013; 2015), and actually found that children's coarticulation decreased with age.

In some respects, theoretical interest in coarticulation has advanced beyond the more versus less detail debate. With the increasing acceptance of exemplar-theoretic models of speech perception and production (Johnson, 1997; Pierrehumbert, 2003) and evidence from infant research that phonological categories are at least partially constructed bottom-up (Houston & Jusczyk, 2003; Werker & Curtin, 2005), many contemporary development theories now presume that children's phonological representations contain detailed acoustic and somatosensory information (McAllister Byun, Inkelas, & Rose, 2016; McAllister Byun & Tessier, 2016; Menn, Schmidt, & Nicholas, 2013; Seidl, Onishi, & Cristià, 2014). As a result, some contemporary studies on child coarticulation, particularly those that include articulatory ultrasound data in addition to inferential acoustics, see coarticulation as a phenomenon to illustrate the nature of children's multifaceted phonological representations (Barbier et al., 2013, 2015), rather than a distinction between more versus less detail or gestures versus acoustics as the basis of representation.

### **2.2.2 Morphological structure and speech production**

Until quite recently, research on the acquisition of morphology had been pursued independently from children's phonetic development.<sup>1</sup> Yet one window into the nature of stored words in the mental lexicon is the effect of morphological structure on speech production. Morphophonetics, or how morphological structure interacts with phonetic variability, is increasingly studied in adult speech. We now know that morphological composition dictates some adult speech variability (Baker et al., 2007; Cho, 2001; Sugahara & Turk, 2009; cf. Hanique

& Ernestus, 2012), but it is often not clear when, how much, or even why (Mousikou, Strycharczuk, Turk, Rastle, & Scobbie, 2015).

### **2.2.3 In adult speakers**

Frequency appeared to mitigate the phonetic realization of morphemes in Losiewicz (1995) where English past tense *-ed* was temporally longer on low-frequency verbs than high-frequency. Elsewhere, Cho (2001) found that intergestural timing in Korean, measured with EMA and EPG, 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). /l/ darker in stem-final position (*coolest*) than affix-initial (*couples*), independent of duration, because *coolest* is analogizing to word-final /l/ in its base form *cool*. Other notable findings about the relationship between morphological composition and phonetic variability in adult speech include that pseudo-prefixes (*mistake*) were shorter in duration than real prefixes (*misjudge*) (Baker et al., 2007) and that heteromorphemic words are longer than otherwise identical monomorphemic words (e.g. *sighed* versus *side*) (Sugahara & Turk, 2009).

However, the relationship between duration and morphology is not always straightforward. For example, (Plag, Homann, & Kunter, 2017) found that, after controlling for a host of other factors such as number of syllables, speaking rate, and surrounding context, 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/ of third person markers. The cause of this difference is not clear. Elsewhere, others have argued that it is not morphological structure but word information load that predicts these interactions between speech production and word structure (Haniue & Ernestus, 2012; Pluymaekers et al., 2010).

### **2.2.4 In child speakers**

Despite this disparate work for adult speech, we know next to nothing about how complex relationships of acoustic variability and morphological composition develop in children. Understanding this could help us better explore children's phonological faculty, as speech development seems to interact with word structure (Mealings, Cox, & Demuth, 2013; Mealings & Demuth, 2014). Beyond that, the study of child morphophonetics could also respond to a long-standing theoretical question in language development: how is children's early language, specifically morphologically-complex words, represented throughout childhood?

Song, Demuth, Shattuck-Hufnagel, et al. (2013) studied morphophonetic development in English-learning children. They employed acoustic and ultrasound analyses to study adult and 2;0 children's CC syllable (/ks/) coarticulation in the coda of bimorphemic (*rocks*) and monomorphemic (*box*). 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 leads the authors to conclude that the primary articulatory target for monomorphemic *box* is C<sub>1</sub> of the consonant cluster (/ks/) while the articulatory target for bimorphemic *rocks* is C<sub>2</sub>, 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, et al. (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 2;0, child speakers of American English reliably distinguish between otherwise identical morphemic and non-morphemic segments. This finding means that children may not rote-memorize morphologically complex forms, but may compose the words online.

### **2.3 Current study**

The primary goal of this work is to study adult and child coarticulation production patterns as a lens into the storage of morphologically-complex words. To do this, I employ an acoustic metric to gauge the coarticulatory influence of adjacent phones on one another. Coarticulation has played numerous illustrative roles for phonetic and phonological theory. It served as the theoretical motivation and foundation for Articulatory Phonology, a theory that takes individual, overlapping gestures as the basic units of speech (Browman & Goldstein, 1989, 1992). Its ubiquity and relevance for accurate perception challenge more abstractionist models (Lahiri & Marslen-Wilson, 1991). However, in this paper, these theoretical frameworks and debate surrounding the form of gestural units are less relevant. Here coarticulation is quantified as the shared acoustic information between two adjacent phones. Of course articulatory gestures and acoustics are interrelated – modifications in the former modulate the acoustic signal. And articulation is a significant factor in coarticulatory patterns and development (Iskarous et al., 2013). However, this paper focuses upon the ensuing acoustics and not underlying coarticulatory gestures.



The child speech apparatus creates multiple issues for the study of acoustic phonetics, spectral analyses in particular (Vorperian & Kent, 2007). Small vocal tracts result in higher resonant frequencies meaning harmonics in the spectral envelope can be widely spaced. This can render an undersampled spectral shape obfuscating formant frequency peaks. Consequently, I employ an acoustic measure to study coarticulation that is immune to these challenges: the difference between Mel frequency cepstral coefficient vectors averaged over adjacent phones. Because this measurement is not commonly employed in the child coarticulation literature, I first validate it (Experiment 1) as I replicate broader anticipatory coarticulatory findings (Mooshammer, Hoole, & Geumann, 2006; Recasens & Espinosa, 2009; Recasens, Pallarès, & Fontdevila, 1997). In Experiment 1, coarticulation between adjacent phones in the syllabic contexts .C/a/ and /a/.C, where the articulation manner of C varies, is measured. The degree of coarticulation should vary as a function of the articulation manner of C, with less sonorous consonants resisting the coarticulatory influence from the adjacent vowel. This pattern should apply across adults and children alike.

After validating the coarticulatory measurement, in Experiment 2 I turn to lexical storage mechanisms for adults and children. SBQ provides unique insight into interactions of morphology and phonetics in adult and child speakers. Adult SBQ speakers have a highly flexible inflectional and derivational lexicon. 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, arrange novel noun-adjective pairings.

SBQ is understudied, both in child language research and linguistics more broadly. Still, much like speakers of well-studied analytic languages such as English and French, child speakers of Cuzco Quechua appear to have a productive morphology apparent by age 5;0, if not earlier (Courtney & Saville-Troike, 2002). However, as outlined in the background research, child coarticulation results suggest that word representations initially develop more holistically. This conclusion from L1 phonetics provides acoustic evidence for recent proposals in broader child language research suggesting that children initially store language more holistically than adults (Lieven, Pine, & Baldwin, 1997; Tomasello, 2003; Vihman & Keren-Portnoy, 2013). Such approaches argue that children do not always deconstruct phrases into words or words into phones. Instead, they may access language in chunks; particularly frequently-occurring combinations or phonotactically-probable “subwords.” In English you could imagine that chunks like *Immagonna* or *lookatit* could be stored in this way, in addition to their deconstructed form.

Consequently, I predict that child speakers will not decompose complex words. I test this in a tightly-controlled experimental paradigm where the coarticulation patterns between the phones [a] and [p] are measured. When [ap] straddles a morpheme boundary (e.g. llama-pi ‘llama-LOC’), children will coarticulate between the phones more than adults, indicating that

they are not deconstructing the word, but are instead storing it holistically. Adults will coarticulate more within morphemes (e.g. **papa** ‘potato’) than across morpheme boundaries (e.g. **llama-pi** ‘llama-LOC’). This is because adults have learned to parse speech segmentally and store it morpho-phonemically. We can tell that this pattern likely comes from experience because adults’ phonetic realization of the *exact same* [ap] sequence differs within a morpheme versus across a boundary.

This approach to language development may seem radical to many 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 (supposedly) untenable demands upon psycholinguistic representation and instigated logical fallacies such as the Poverty of the Stimulus. However, a “chunking” approach to language acquisition does not need to reject abstraction, it just postpones it until later in development. This line of research is emergent but already a series of papers have documented how high-frequency collocations could be represented holistically, with redundant representation at multiple levels, in adults (Arnon & Christiansen, 2017; Arnon & Cohen Priva, 2013; Arnon & Snider, 2010) and crucially, children (Arnon, 2010; Arnon & Christiansen, 2017; Bannard & Matthews, 2008).

So the concept of more holistic storage in children is not crazy. In fact, 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 even with spoken language, might segment and store speech differently than adults. This study empirically tests this theory.

Study hypotheses and design, including the partitioning of participant age groups, were pre-registered in the Open Science Framework on September 11, 2018. Registration, raw acoustic data, and code to replicate the analyses in this paper are available in the affiliated project (Cychosz, 2018).

## **2.4 Experiment 1 Methods**

### **2.4.1 Participants**

30 children (15 girls, 15 boys) and 10 adults (10 female), all bilingual Spanish–SBQ speakers from a mid–size town in southern Bolivia, completed a picture–prompted word elicitation task. Child participants were grouped by age: five 5–6–year–olds (mean = 6;6 [years;months]), nine 7–year–olds (7;7), six 8–year–olds (8;5), five 9–year–olds (9;8), and five 10–year–olds (10;7).

### 2.4.2 Stimuli

To elicit the words, participants were presented with photos of 32 culturally-appropriate nouns that children in these communities recognize (e.g. *house, flower, cow*). Each word was carried with the locative marker *-pi*, which was used in experiment 2 (see 2.3). Only words with [a]<sub>C<sub>onsonant</sub></sub> transition or C<sub>onsonant</sub>[a] were analyzed (C[a]: *N*=13; [a]C: *N*=9) (Table 1). Because SBQ is canonically open-syllabic, nearly all VC syllables transcend syllable boundaries and nearly all CV syllables are syllable-internal. The vowel in the VC syllables is always stressed, but this is not the case for the CV syllables. I control for stress in the statistical model for CV syllables. The vowel [a] was chosen to study in VC and CV coarticulation as it is the most frequent vowel in SBQ, making it easier to find child-friendly items to elicit during the task.

| WORD            | TRANSLATION         | SYLLABLE STRUCTURE |
|-----------------|---------------------|--------------------|
| 'api            | 'corn/citrus drink' | V.C                |
| 'chaki/'chaki   | 'potato'            | V.C/.CV            |
| ka'tari/ka'tari | 'snake'             | V.C/.CV            |
| 'llama/'llama   | 'llama'             | V.C/.CV            |
| 'maki/'maki     | 'hand'              | V.C/.CV            |
| 'papa/'papa     | 'potato'            | V.C/.CV            |
| 'waka/'waka     | 'cow'               | V.C/.CV            |
| 'wasi/'wasi     | 'house'             | V.C/.CV            |
| 'wawa/'wawa     | 'baby'              | V.C/.CV            |
| 'qhari          | 'boy'               | .CV                |
| 'sunkha         | 'beard'             | .CV                |
| 't'ika          | 'flower'            | .CV                |
| 'wallpa         | 'chicken'           | .CV                |
| 'warmi          | 'woman'             | .CV                |

**Table 1:** Words elicited and syllabic structure: between or within morpheme. For words that contained both syllabic environments, both .[a]C and C.[a] transitions were measured. Stress marked with ' ; ejectives marked with '.

I first had to ensure that children knew the words that they were producing. There is no equivalent to the *Macarthur Bates Communicative Development Inventory* (Fenson et al., 2007) which reports stages of age-normed vocabulary development, for any Quechuan language or Bolivian Spanish. Nor is there a large, naturalistic child-directed speech corpus to infer stages of SBQ vocabulary development. I remedied this in two ways: first, unlike other child elicitation techniques, children here did not mimic a model speaker but instead produced the word spontaneously. If a child did not know a word, it was skipped. Second, a pre-test confirmed that children as young as 3;0 could name the items to be elicited. Two items elicited variable Spanish

borrowing/SBQ responses from some children: ‘sheep’ prompted both *uveja* and *ulla* and ‘duck’ prompted *patu* and *pili*. Additionally, *qowi* ‘guinea pig’ was not readily recognized by all children. These items were not analyzed. Because these steps were taken, I am confident that children knew all of the words they produced.

### **2.4.3 Data collection**

All participants were recorded with a portable Zoom H1 Handy Recorder at a 44.1 kHz sampling rate while completing the word elicitation task. Children in these communities have limited exposure to technology so instead of eliciting items on a screen, a photo of each item was pasted onto a single page in a binder. For this reason, the words could not be randomized and were presented in the same order for all participants.

Participants first produced all words in Spanish and then repeated the task in SBQ; only SBQ results are analyzed here. Each of the items was elicited twice, via two distinct pictures on separate trials. Ideally, this combination would result in 4 tokensX14 words. However, due to noise interference from wind and livestock, as well as children’s occasional nervousness, some children only produced the target word one or two times. I controlled for this between–subject variability by testing the parameter **Utterance number** in the statistical models; it did not improve model fits.

For each of the 32 items, participants named the item in the photo in a carrier phrase by twice repeating, “I say in the \_ two times” (*Noqa nini \_-pi iskay kutita*). Most children under 8;0 could not remember the carrier phrase and instead first identified the bare noun (e.g. *llama*). Then, the researcher placed a large plastic toy insect on top of the photo and prompted the child, “Where is the bug?” to which the child produced the word with the correct suffixal carrier e.g. *llama-pi* (*llama-LOC*, “on the llama”) two times. The task took 20–30 minutes. All participants were monetarily compensated. Children could additionally choose one small item from a toy bag.

### **2.4.4 Data analysis**

Productions were manually segmented in Praat (Boersma & Weenik, 2018). Much of the child coarticulation literature employs acoustic measures of coarticulation such as center of gravity for fricatives or formant–based measurements (e.g. transitions or spectral peaks) for vowels. However, measurements of child formants are notoriously difficult to obtain reliably (see Section 2.3). The coarticulation method tested here closely follows a methodology seen in Gerosa, Lee, Giuliani, & Narayanan (2006) who likewise studied children’s coarticulation between stops and vowels. I compute coarticulation via an automatically–extracted measure of cepstral change:

the difference between the averaged MFCC vectors from two adjacent phones. To compute this, the acoustic signal was first downsampled to 12 kHz. Then, each phone was segmented into 25.6ms frames, with a 10ms step. 11–14 MFCCs were parameterized from the ensuing slice. The MFCC vectors from each phone were then averaged with the resulting vector scaled by the duration of the carrier word to control for speaking rate. Finally, I measured the Euclidean distance between the averaged MFCC vector for each relevant biphone sequence for each word:

$$(1) \quad d_{ap} = \sqrt{\sum \left( \frac{\bar{x}_a}{D_{word}} - \frac{\bar{x}_p}{D_{word}} \right)^2}$$

where  $d_{ap}$  is the Euclidean distance between segments  $a$  and  $p$  in the biphone sequence  $ap$ ,  $\bar{x}_a$  and  $\bar{x}_p$  are the averaged MFCC vectors of each segment, and  $D_{word}$  is the duration of the carrier word. The measurement was implemented in the same manner for all biphone sequences regardless of syllable structure (V.C or .CV) or phone identity.

## 2.5 Experiment 1 Results

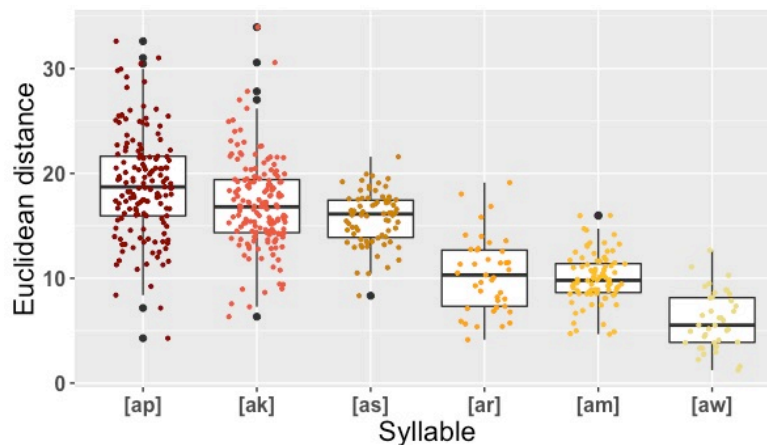
Linear mixed effects regression models were fit to predict the Euclidean distance between the phones in the [a].C or .C[a] sequences using the lme4 (Bates, Maechler, Bolker, & Walker, 2015) and lmerTest packages (Kuznetsova, Brockhoff, & Christensen, 2017) in R (R Core Team, 2018). Potential model parameters were evaluated using a combination of between-model log-likelihood comparisons, AIC estimations, and  $p$ -values. Continuous variables were mean-centered.

For both models, .C[a] and [a].C, baseline models were first fit with random intercepts for **Speaker**. Models were then built in a forwards testing procedure. **Age** [adult/child] did not significantly improve upon baseline fit for either model. However, for both .C[a] and [a].C models, the parameter **Word duration** did improve model fit. The coarticulation metric was also scaled by word duration during its measurement, as shown in example (1), but including a parameter for word duration in the model maximally controls for the prevalent speaking rate differences between adults and children. The **Word duration** coefficient demonstrates that temporally longer words showed less coarticulation for both syllable types (Table 2 & 3).

|                | $\beta$ | S.E. | $t$ -value | $p$ -value | 97.5% C.I.     |
|----------------|---------|------|------------|------------|----------------|
| INTERCEPT      | 19.21   | 0.43 | 44.36      | <.001      | 14.72, 18.74   |
| CONSONANT: [p] |         |      |            |            |                |
| [k]            | -2.17   | 0.39 | -5.51      | <.001      | -2.94, -1.40   |
| [s]            | -3.59   | 0.47 | -7.56      | <.001      | -4.52, -2.66   |
| [r]            | -9.13   | 0.63 | -14.61     | <.001      | -10.36, -7.91  |
| [m]            | -8.96   | 0.45 | -19.71     | <.001      | -9.86, -8.07   |
| [w]            | -12.63  | 0.60 | -21.11     | <.001      | -13.81, -11.46 |
| WORD DURATION  | 4.26    | 1.71 | 2.49       | 0.013      | 0.91, 7.60     |

**Table 2.** Model predicting Euclidean distance (coarticulation) between adjacent phones in [a].C syllables

The parameter for **Consonant** also improved fit for both models, indicating that the amount of coarticulation within the sequence depended upon consonant identity. For the [a].C model, I set the reference level to [p] and I found that the Euclidean distance (i.e. degree of coarticulation) differed significantly between [p] and each of the other preceding consonants (Table 2). The effect increases such that more sonorous phones have larger negative  $\beta$  values. This indicates that the distance between phones decreases as the consonant becomes more sonorous (i.e. further from reference level [p]). Figure 1 demonstrates this phenomenon of coarticulatory distance by manner of articulation. Speakers coarticulate less between segments in the sequence [ap], for example, than [as] ( $\beta = -3.59$ ,  $t = -7.56$ ,  $p < .001$ , CI = -4.52, -2.66) or between [aw] ( $\beta = -12.63$ ,  $t = -21.11$ ,  $p < .001$ , CI = -13.81, -11.46).



**Figure 1.** V.C cepstral distance (coarticulatory resistance) by syllable. Black line represents median distance. Error bars represent 1.5x the IQR in each direction.

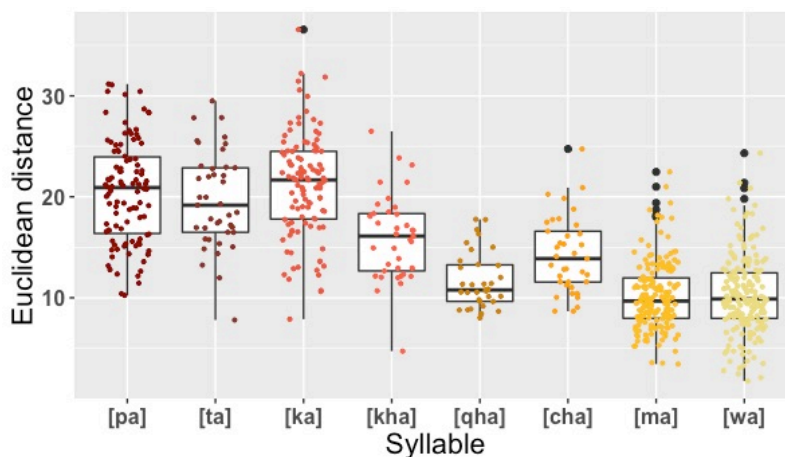
I next tested the parameter **Consonant** for the .C[a] model. However, unlike the syllables tested in the [a].C model, not all .C[a] syllables fell into the same prosodic position – some fell

in primary stress position and some did not. Consequently, a final binary parameter, [+/-] **Stress**, was added to the .C[a] model, which significantly improved model fit ([+Stress]:  $\beta = -1.79$ ,  $t = -4.24$ ,  $p < .001$ , CI = -2.62, -0.96).

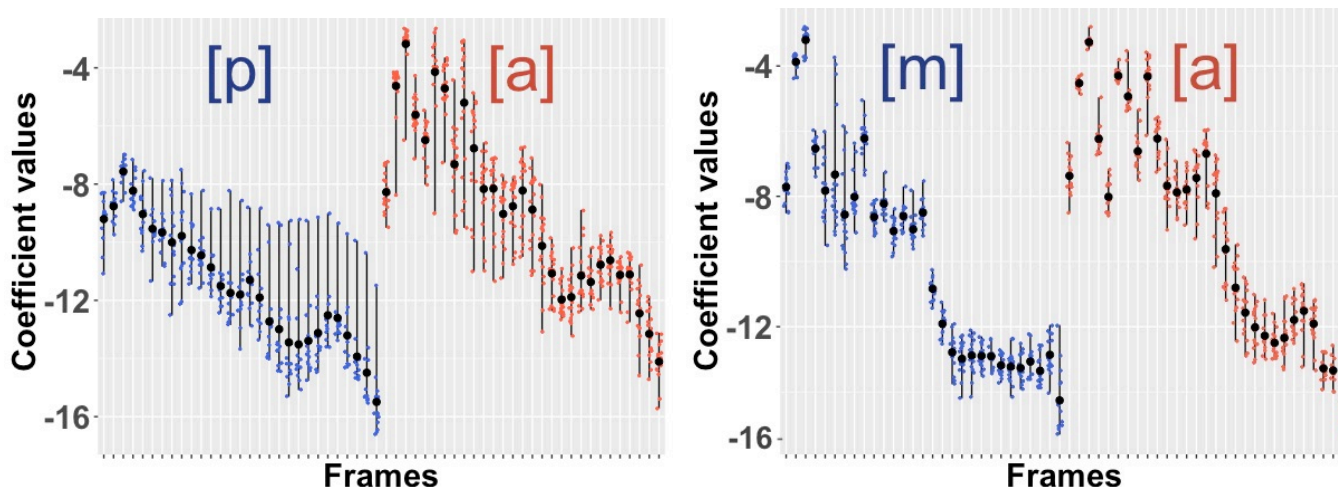
Setting the reference level for **Consonant** to [p] once again, I found significant differences between [p] and all the remaining consonants, with the distance again increasing as the phones became more sonorous (Figure 2). For example, [w] had a larger negative  $\beta$  coefficient value ( $\beta = -11.61$ ,  $t = -21.01$ ,  $p < .001$ , CI = -12.69, -10.52) than [kh] ( $\beta = -5.10$ ,  $t = -7.17$ ,  $p < .001$ , CI = -6.49, -3.71) or [t] ( $\beta = -2.94$ ,  $t = -3.72$ ,  $p < .001$ , CI = -4.50, -1.39).

|                | $\beta$ | S.E. | $t$ -value | $p$ -value | 97.5% C.I.     |
|----------------|---------|------|------------|------------|----------------|
| INTERCEPT      | 22.32   | 0.63 | 35.34      | <.001      | 21.08, 23.56   |
| CONSONANT: [p] |         |      |            |            |                |
| [t]            | -2.94   | 0.79 | -3.72      | <.001      | -4.50, -1.39   |
| [k]            | 0.99    | 0.50 | 2.0        | 0.046      | 0.02, 1.96     |
| [kh]           | -5.10   | 0.71 | -7.17      | <.001      | -6.49, -3.71   |
| [qh]           | -10.33  | 0.83 | -12.50     | <.001      | -11.95, -8.71  |
| [ch]           | -8.02   | 0.81 | -9.93      | <.001      | -9.60, -6.44   |
| [m]            | -11.14  | 0.50 | -22.40     | <.001      | -12.12, -10.17 |
| [w]            | -11.61  | 0.55 | -21.01     | <.001      | -12.69, -10.52 |
| STRESS: [+]    | -1.79   | 0.42 | -4.24      | <.001      | -2.62, -0.96   |
| WORD DURATION  | 5.63    | 1.60 | 3.52       | <.001      | 2.49, 8.76     |

**Table 3.** Model predicting Euclidean distance (coarticulation) between adjacent phones in .C[a] syllables



**Figure 2.** .CV cepstral distance (coarticulatory resistance) by syllable. Black line represents median distance. Error bars represent 1.5x the IQR in each direction.



**Figure 3.** Cepstral distance between adjacent phones in .CV syllables from a single adult speaker saying “wallpa-pi” *chicken*-LOC (L) and “llama-pi” *llama*-LOC (R). Small points indicate individual MFCCs. Large points indicate mean MFCC for a given frame. Black bars for each frame indicate S.E. of the coefficients.

To envision how this coarticulatory pattern plays out by consonant in a different way, Figure 3 displays coarticulatory differences measured with this cepstral metric from a single speaker. The coefficient values for [p] in the syllable [pa] differ from those for the adjacent [a]. The mean coefficient value for each of the [p] frames – represented with the black points – is consistently smaller across the 29 frames. However, note that for the syllable [ma], both [m] and [a] have similar patterns. This similarity indicates that as the speaker transitioned from [m] to [a], the acoustic signal changed less than for the transition between [p] and [a].

## 2.6 Interim Discussion

In Experiment 1, I validated a measure of coarticulation that did not rely on spectral analysis. To demonstrate this, I hypothesized that a cepstral metric, the difference between averaged MFCC vectors from adjacent phones, would vary by consonant manner of articulation. Specifically, more sonorous phones such as [w] and [m] should show less resistance to coarticulation with the adjacent [a] than obstruent sounds such as [p] and [k<sup>h</sup>]. This would replicate previous reports that the degree of coarticulation should vary by the manner of consonant articulation (Mooshammer et al., 2006; Recasens & Espinosa, 2009; Recasens et al., 1997). Results from Experiment 1 corroborated this hypothesis: the Euclidean distance between the MFCC vectors was greater for less sonorous sounds and the adjacent [a] than more sonorous sounds. This relationship trended in the expected relationship as well, with an expected gradual



decrease in Euclidean distance between more sonorous sounds and [a]. Furthermore, the relationship applied both across and within syllables as I tested coarticulation in .CV and V.C environments.

Having established the validity of the cepstral measure for coarticulation, I next turned to the primary investigation: evaluating coarticulatory patterns across and within morphemes in adult and child SBQ speakers.

## **2.7 Experiment 2 Methods**

### **2.7.1 Participants**

The same 40 participants (30 children, 10 female adults) in experiment 1 contributed data for experiment 2.

### **2.7.2 Stimuli**

The overall stimuli for Experiment 2 were identical to those from Experiment 1 (they were collected during the same task). However, a different subset of words, only those containing the sequence [ap], were analyzed. Each elicited word was still carried with the locative marker *-pi* (see section 2.4.2).

The locative marker was chosen instead of any number of other SBQ suffixes for a couple of reasons. First, I used a nominal suffix instead of verbal as noun declensions are easier to elicit via pictures than derived word forms (e.g. *puñu-y* ‘to sleep’ → *puñu-chi-y* ‘to make (one) sleep’) or verb conjugations. Nouns are also grammatical in SBQ with just one suffix. Some conjugated verbs require multiple suffixes (see ‘sleep’ example above) which would make elicitation and tight control of the experimental stimuli more difficult. In using nominal suffixes, I can isolate elicitation to a single stem + suffix combination for this initial study. Finally, the locative marker was also used because, absent a large corpus of child-directed SBQ speech, we can hypothesize that the locative *-pi* on high-frequency nouns, such as those I elicited, is going to be relatively frequent in a child’s input.

In addition, most research on child morphophonetics has been conducted on the English plural suffix. It would be ideal to replicate this in SBQ because any difference between studies would not be due to semantic differences between morphemes. However, SBQ speakers generally replace the SBQ plural suffix *-kuna* with the Spanish plural suffix *-s* and the plural tends to be optional in SBQ anyways.<sup>ii</sup>

Only words with adjacent [a] and [p] phones in stressed position were analyzed ( $N=11$ ). The combination of [a] and [p] was chosen as [a] is the most frequent vowel in SBQ, occurring within and across morpheme boundaries in several common nouns.

| WORD         | TRANSLATION         | MORPHOLOGICAL ENVIRONMENT |
|--------------|---------------------|---------------------------|
| 'api         | 'corn/citrus drink' | within                    |
| 'papa        | 'potato'            | within                    |
| imi'lla-pi   | 'girl-LOC'          | between                   |
| juk'u'cha-pi | 'mouse-LOC'         | between                   |
| lla'ma-pi    | 'llama-LOC'         | between                   |
| sun'kha-pi   | 'beard-LOC'         | between                   |
| t'i'ka-pi    | 'flower-LOC'        | between                   |
| uhu't'a-pi   | 'sandal-LOC'        | between                   |
| wa'ka-pi     | 'cow'-LOC'          | between                   |
| wall'pa-pi   | 'chicken-LOC'       | between                   |
| wa'wa-pi     | 'baby-LOC'          | between                   |

**Table 4:** Words elicited and morphological environment: between or within morpheme.

Stress marked with ' ; ejectives marked with '.

Most work on child coarticulation has elicited .CV sequences. However for the current experiment, V.C sequences were elicited because the default syllable structure in SBQ is open – CVC syllables, particularly word-finally, are infrequent. This makes it difficult to find high-frequency, child-friendly word stems that end in a consonant. Furthermore, there are not any vowel-initial nominal suffixes in SBQ. So it is impossible to derive a context like C + V.

### 2.7.3 Data collection and analysis

Elicited words were collected during the same task as experiment 1. Again due to interference from wind and livestock, as well as children’s occasional nervousness, some children only produced the target word one or two times. I again controlled for this between–subject variability by testing the parameter **Utterance number** in the statistical models for experiment 2; it did not improve model fits.

Productions were manually segmented in Praat (Boersma & Weenik, 2018) and the coarticulation metric is the same as described in the methods for Experiment 1.

## 2.8 Experiment 2 Results

A linear mixed effects regression model was fit to predict the Euclidean distance between [a] and [p] using the lme4 (Bates et al., 2015) and lmerTest (Kuznetsova et al., 2017) packages in R (R Core Team, 2018). Potential model parameters were evaluated using a combination of

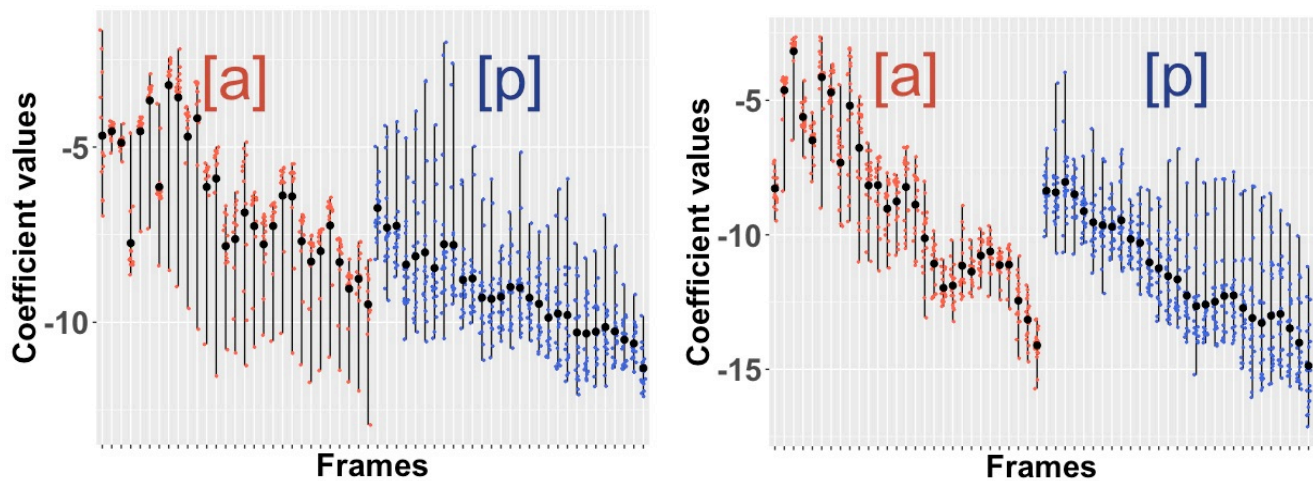
between-model log-likelihood comparisons, AIC estimations, and  $p$ -values. Continuous variables were mean-centered.

The baseline model included the random intercepts of **Word** and **Participant**. Parameters were evaluated in a forwards-testing procedure. Recall that a significant effect for **Morphological environment** [Across morpheme vs. Within morpheme] would indicate that the coarticulation pattern differed between the two morphological environments. An interaction of the parameters **Age** [Child vs. Adult] and **Morphological environment** would demonstrate that one age group, the children or the adults, distinguished coarticulatorily between the morphological environments and one age group did not distinguish between the environments.

In the first model, neither **Morphological environment**, **Age**, nor their interaction improved model fit (Table 5). Only **Word duration** was significant ( $\beta = 3.65$ ,  $t = 2.49$ ,  $p = .013$ , CI = 0.34, 6.58) where a positive beta coefficient indicates less coarticulation between phones in temporally longer words. As in Experiment 1, **Word duration** was included in the model to maximally control for the prevalent speaking rate differences between adults and children. I did this even though the coarticulation metric was scaled by word duration during measurement, as shown in example (1).

|  | $\beta$ | S.E. | $t$ -value | $p$ -value | 97.5% C.I.   |
|--|---------|------|------------|------------|--------------|
| INTERCEPT                                | 21.70   | 0.89 | 24.52      | <.001      | 20.00, 23.43 |
| AGE: [ADULT]                             | -1.96   | 1.30 | -1.51      | 0.14       | -4.56, 0.56  |
| MORPH. ENVIRONMENT:<br>[ACROSS MORPHEME] | -0.66   | 0.84 | -0.79      | 0.45       | -2.39, 0.96  |
| AGE*MORPH.<br>ENVIRONMENT                | 0.84    | 0.72 | 1.17       | 0.24       | -0.55, 2.32  |
| WORD DURATION                            | 3.65    | 1.47 | 2.49       | 0.013      | 0.34, 6.58   |

**Table 5.** Model predicting the Euclidean distance between MFCC vectors of [a] and [p]: whole phone model



**Figure 4.** Cepstral distance between adjacent phones [ap] from a single adult speaker (L) and child speaker age 6 (R) saying “wallpa-pi” *chicken-LOC* (L). Small points indicate individual MFCCs. Large points indicate mean MFCC for a given frame. Black bars for each frame indicate S.E. of the coefficients.

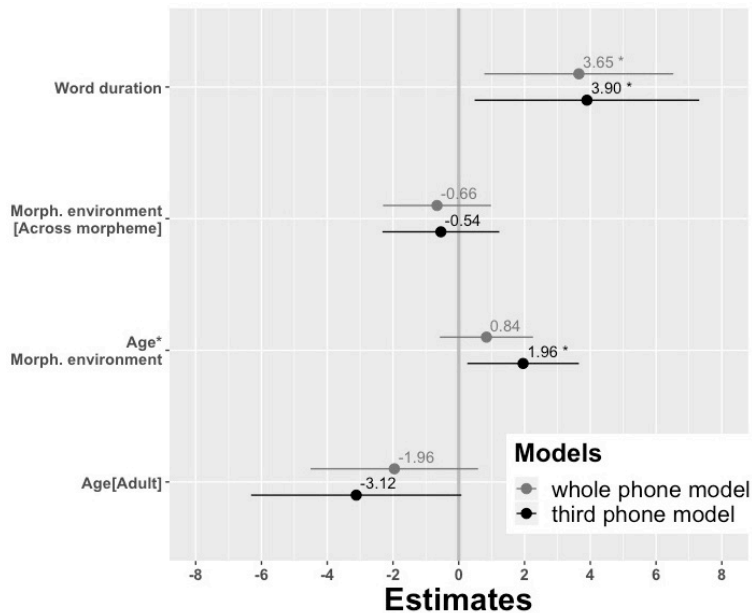
Figure 4 illustrates the similarity between adult and child coarticulation patterns in another manner. There, the mean MFCCs for the child, displayed on the right, are similar across [a] and [p] indicating that the child did not distinguish between these sounds. For the adult speaker, on the left, the mean MFCCs are once again highly similar between [a] and [p]. The result is that, as the model predicted, the Euclidean distance between the MFCC vectors of adults and children did not significantly differ.

The lack of effect for **Age** or **Morphological environment** in the model indicated that neither children nor adults distinguished between the two morphological environments. This result suggests that even *adult* speakers were storing these high-frequency nouns holistically. Alternatively, it could mean that the coarticulation measure was not sensitive enough to capture differences between adults and children.

To further explore these possibilities, I conducted an exploratory analysis hypothesizing that some coarticulatory differences between adjacent phones may be washed out when averaging MFCCs over an entire phone. To test this, I next compared the Euclidean distance between the averaged MFCC vectors taken from the *middle third* of each phone.

I fit an additional model to predict the measurements from the middle third MFCC vectors of each phone. The third-phone model fitting procedure closely followed the whole-phone procedure. Again random effects for **Participant** and **Word** were included in the baseline model. Best model fit again included the fixed effect of **Word duration** ( $\beta = 3.90$ ,  $t = 2.24$ ,  $p = .026$ ,

CI=0.15, 7.35), indicating less coarticulation in longer duration words (Table 6). However, now the interaction of **Morphological environment** and **Age** was significant ( $\beta=1.96$ ,  $t=2.26$ ,  $p=0.024$ , CI=0.28, 3.73). See Figure 5 for visual comparison of the parameters and estimates between the whole–phone and third–phone models.



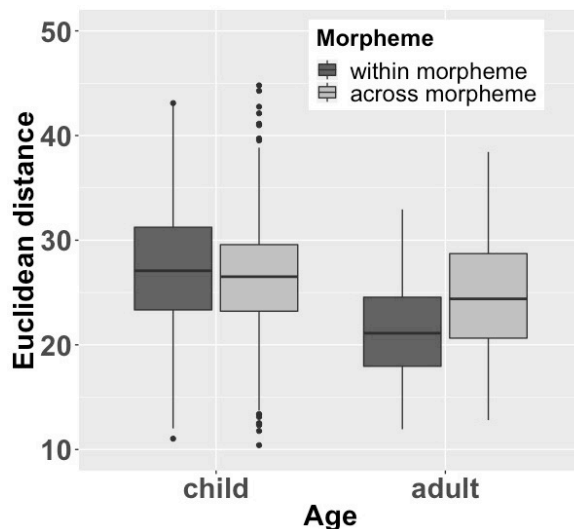
**Figure 5:** Predicting the Euclidean distance between MFCC vectors of [a] and [p].

|   | $\beta$ | S.E. | $t$ -value | $p$ -value | 97.5% C.I.   |
|---|---------|------|------------|------------|--------------|
| INTERCEPT                                       | 27.16   | 1.03 | 26.43      | <.001      | 25.18, 29.15 |
| AGE: [ADULT]                                    | -3.12   | 1.63 | -1.91      | 0.062      | -6.38, 0.04  |
| MORPH.ENVIRONMENT [ACROSS MORPHEME]             | -0.54   | 0.91 | -0.60      | 0.560      | -2.41, 1.21  |
| AGE[ADULT]* MORPH.ENVIRONMENT [ACROSS MORPHEME] | 1.96    | 0.87 | 2.26       | 0.024      | 0.28, 3.73   |
| WORD DURATION                                   | 3.90    | 1.74 | 2.24       | 0.026      | 0.15, 7.35   |

**Table 6.** Model predicting the Euclidean distance between MFCC vectors of [a] and [p]:  
third phone model

This significant interaction of **Age** with **Morphological environment** in the third–phone model suggests that adults have different coarticulatory processes between [a] and [p] when the segments co–occur within a morpheme and across morphemes (Figure 6). However, children do not have different coarticulation patterns between these environments. This could be because

children do not decompose these high-frequency noun-suffix pairings, or at least not to the extent of adults.

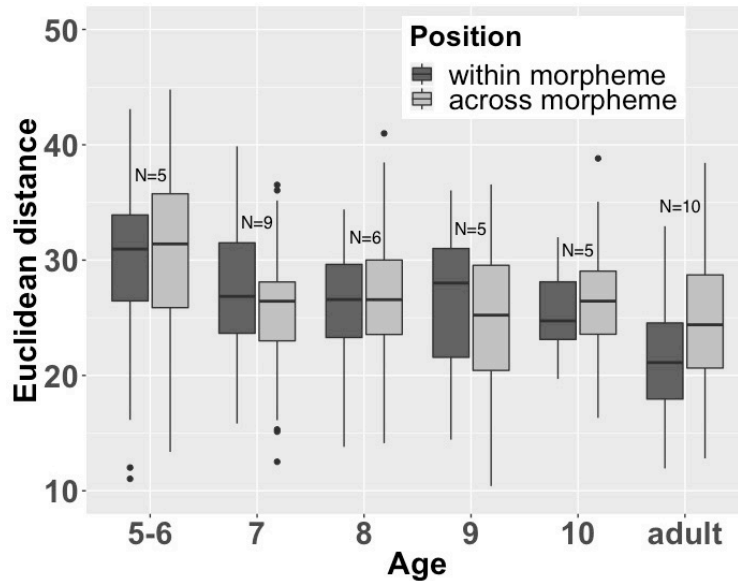


**Figure 6:** Interquartile range of cepstral distance between middle third of [a] and middle third of [p]. Black line represents median distance. Error bars represent 1.5x the IQR in each direction.

Beyond differences between adults and children, I did not have a specific hypothesis about the age trajectory for this development. In other words, I did not hypothesize at what age children would distinguish between transitions across morphemes versus within.

To test the developmental trajectory, a factored variable of **Age group** – 5 child age bins and adults – was tested but it did not improve model fit. Nevertheless, Figure 7 displays a trend suggesting that 1) children do not distinguish between the two morphological environments until age 10 and 2) the overall distance between phones decreases with age.

The trend for child participants to coarticulate more as they age is not consistent – note that the 10-year-olds pattern like adults but nine-year-olds appear to coarticulate *more* between morphemes than within. Furthermore, in a post-hoc analysis with ‘adult’ as the **Age group** reference level, only 7-year-olds reliably differed from adults ( $\beta = -3.54$ ,  $t = -3.27$ ,  $p = .001$ , 97.5% CI =  $-5.77$ ,  $-1.45$ ). There were no additional statistically significant differences between adults and any other child age group. This unclear pattern by age group may be due to the unbalanced/small sample sizes within each group. I hope to complement this analysis with more data collection on subsequent fieldtrips. This may help to disentangle the developmental trajectory.



**Figure 7:** Cepstral distance between middle third of [a] and [p] across age groups. Black line represents median distance. Error bars represent 1.5x the IQR in each direction. Number of participants per age group listed above each box.

## 2.9 General discussion

Speech does not unfold like pearls on a string. Coarticulation between segments during production is rampant, with frequent overlap of articulatory gestures (Lindblom, 1963). Phonetic realization also varies by lexical and morphological context (Cho, 2001; Lee-Kim et al., 2013; Sugahara & Turk, 2009). In this study, I attempted to examine how adult and child (aged 5–10) SBQ speakers coarticulated in two environments within morphologically–complex words: across morphemes and within a morpheme. Given the challenges that child voices bring to many standard spectral measurements, Experiment 1 first validated a relatively new coarticulatory measure. By employing MFCCs to measure coarticulation, I avoided issues inherent to child acoustics like the disparate harmonics in the spectral envelope and missing formant peaks. This first experiment demonstrated the validity of the cepstral measure for adult and child speech by replicating common findings regarding coarticulatory differences by manner of articulation. Specifically, I demonstrated that adult and child speakers coarticulated significantly more between [a] and more sonorous sounds such as [w] or [m] than between [a] and [p]. This was consistent across both .CV and V.C syllables.

In Experiment 2, I turned to the central objective of the study: measuring coarticulation across two morphological environments in SBQ. Using the validated coarticulation measure from Experiment 1, I demonstrated that adult speakers have different coarticulatory patterns between

morphemes versus within a morpheme. This is anticipated. Adult speakers have highly practiced motor routines, particularly within high-frequency words such as those I elicited. This results in well-rehearsed spectral transitions between the frequently co-occurring phones of a root morpheme. However, in adult speakers, this transitional routine is less practiced at the boundary of root morpheme and suffix. This is likely because nominal suffixes are frequently rearranged from the stem morpheme (noun) to mark different grammatical relations.

The finding that adult speakers distinguished between morphological environments does not mean that adult SBQ speakers *never* represent inflected or derived words holistically. Indeed, as data from English and Dutch has demonstrated, adults' morphological decomposition of complex words into morphemes depends heavily upon the overall frequency of the inflected word (Baayen, 1992) and the relative frequency of the component morphemes (Hay, 2003). This caveat is especially pertinent as I studied just 13 distinct words (to limit the analysis to the transition between a single biplane sequence [ap]). As a result, I can only speculate as to whether decomposition in adult SBQ speakers would play out similarly in other lexical items or in combination with other nominal suffixes. Nevertheless, I take the different patterning across these two environments as evidence that adult SBQ speakers have abstracted grammatical suffixes away from the original lexical contexts in which they heard them.

Crucial to the argument that coarticulation indicates morphological (de)composition in adult speakers is that child SBQ speakers did not distinguish between the morphological environments. Children coarticulated equally within a root morpheme as between two morphemes. This is evidence that children store these high-frequency nouns, with the coinciding locative marker, more holistically relative to adults.

The finding that children do not differentiate between morphological environments is novel evidence for an argument that has been made repeatedly, albeit to different levels of extremity, in child language acquisition: children initially represent language more holistically than adults (Arnon, 2010; Lieven et al., 1997; Tomasello, 2003; Vihman & Keren-Portnoy, 2013). Many authors who study child coarticulation have come to similar conclusions and believe that children do not discriminate between speech sounds, representing words more holistically (Goodell & Studdert-Kennedy, 1992; Nittrouer et al., 1989, 1996; Zharkova et al., 2008, 2011). However, after carefully controlling for speaking rate, I did not actually replicate reports that children coarticulate more than adults. Instead, I show that coarticulation tends to *increase* with age, independent of morphological environment (though I stress that the evidence by age group is a trend; there are too few participants within each age group to definitively conclude). The child coarticulation literature has mixed conclusions so the finding that coarticulation increases



with age has been reported elsewhere (Barbier et al., 2013, 2015; Goffman et al., 2008; Repp, 1986). However, those studies often examined coarticulation across longer-distances.

### **2.9.1 Accounting for differences**

There could be many sources for the differences between the conclusion in this study on child coarticulation and the conclusions of previous works. A difference could arise due to the MFCC measurement that I used. Only one other research group that I am aware of has employed a cepstral distance metric for child coarticulation (Gerosa et al., 2006). Most acoustic studies on child coarticulation that concluded that coarticulation decreases with age relied on formant tracking in the vowel within the syllable (Goodell & Studdert-Kennedy, 1992; Nijland et al., 2002; Nittrouer et al., 1989, 1996; Robb & Wolk, 1997). However, I again stress that formant tracking can be highly unreliable for children's voices, even when measurements are made by hand (Lee, Potamianos, & Narayanan, 1999).

This study and previous work may also differ because I employed transitions between vowels and stops. Vowels and fricatives were analyzed in much previous work. This could be relevant since, as I demonstrated in Experiment 1, certain articulation manners, such as stops and affricates, resist the coarticulatory influence of adjacent sounds more than other articulation manners. However, I do not think that the differences here could entirely be due to the segment manner analyzed. Goodell & Studdert-Kennedy (1992) and Nijland et al. (2002) likewise measured children's coarticulation between some stop-vowel sequences and they still found that coarticulation decreased with age.

In any case, coarticulation between vowel-fricative sequences would be challenging to implement in the current study. For one thing, the most frequent nominal suffixes in SBQ begin with stops or nasals (e.g. accusative: *-ta*, ablative: *-manta*). Furthermore, while the phonemic inventory of fricatives in SBQ contains just /s/ and /h/, there may be an ongoing sound change of *s* > *ʃ* diffusing within the language. This change is highly dependent upon region and degree of bilingualism with Spanish. This could likely render any vowel-fricative patterns in the language unstable and prone to item-specific effects.

I believe that the results differ, in part, because I closely controlled for speaking rate, both in measurement and statistical modeling. Children consistently speak slower, and with longer-duration phones, but previous work on child coarticulation has rarely controlled for rate differences (Goodell & Studdert-Kennedy, 1992; Nijland et al., 2002; Nittrouer et al., 1996; Robb & Wolk, 1997; Zharkova et al., 2011). It is true that several of these studies elicited nonce syllables, and not real words, so rate differences may be less relevant. Indeed, Nittrouer et al.

(1989) did not find that the mean duration of nonce /sa/ and /ʃa/ syllables differed significantly between a random sub-sample of adult and child participants. Furthermore, Zharkova et al. (2011) stress that for their static articulatory measurement, a simple correction for phone duration probably wouldn't be valid as the trajectory of lingual movement, in addition to speaking rate, differs between adults and children. Still, it is unclear if the effect of syllable duration could mitigate the coarticulation differences between adults and children that these previous studies found.

### **2.9.2 Child “coarticulation”**

The level of abstraction in children's language is a contentious issue. Even within the more limited subfield of L1 phonetics and phonology, there exist distinct theories and interpretations. Perhaps children's speech representations are highly abstracted, feature-based categories (Bernhardt & Stemberger, 1998; Fikkert, 1994; Hale & Reiss, 1998), or perhaps they more closely approximate episodic words and word-like traces (Ferguson & Farwell, 1975; Vihman & Keren-Portnoy, 2013), or perhaps some combination of these two approaches (Fikkert & Levelt, 2008; Swingley & Aslin, 2007)?

Relevant to this discussion is the fact that children have highly unreliable articulatory schema. In a model of child phonology incorporating articulatory feedback to representations (McAllister Byun & Tessier, 2016), articulatory feedback in speech development results in, messy, unreliable phonological categories. This means lots of scratches upon a metaphorical landscape into which a child may carve linguistic representations, but few deep, memorable crevices. Now imagine that instead of adult-like abstract segments, children have a word-level representation. It is somewhat abstract (i.e. entrenched), particularly if the word is frequent in the child's ambient or spoken language. However, the word is not so devoid of context so as to entirely mimic the segment-level phenomena that characterize adult phonology. Consequently, when a child produces a word, they grasp for their phonological representation which *is an entire word*. Adults have abstracted away from individual words enough that they can string individual speech sounds together; children cannot. The result is that children may appear to “coarticulate” more. It is the result of the messiness in their perceptual and articulatory representations.

Attributing child coarticulation practices to messy or underspecified phonological representations is not novel. However, if children's underdeveloped phonological representations are causing them to coarticulate more between speech segments during speech production, then researchers documenting this phenomenon have not been referring to the planned, efficiency-driven coarticulation of adult speech (Bradlow, 2002; Whalen, 1990). Instead, when previous research has suggested that children coarticulate more than adults, they have likely been

referring to something more aptly named CHILD INDISCRIMINATION. Children are not masters of coarticulation. They have less stable articulatory trajectories (Goffman et al., 2008), so it would be surprising if they could demonstrate motor control at a finer level than adults. Rather, children are just unable to discern the internal segmental structure of words and this may manifest as greater coarticulation during production.

This crucial distinction between planned, adult coarticulation and unskilled, child indiscrimination is one that that few make. Of the many studies on child coarticulation, only two remark upon the difference: “phenomena commonly lumped together under the heading of ‘coarticulation’ may have diverse origins...some forms...are an indication of advanced speech production skills whereas others may be a sign of articulatory immaturity” (Repp, 1986:1618; see also Whiteside & Hodgson, 2000). However, the findings presented here – that children coarticulate similarly across morphological contexts while simultaneously coarticulating less than adults – supports a distinction between planned coarticulation and child indiscrimination. This finding applies even though I found that children coarticulated less than adults overall. Planned coarticulation is a complex speech task that takes years of mastery and practice while child indiscrimination, I argue, is the consequence of children’s linguistic inexperience and holistic lexical storage.

## **2.10 Conclusion**

This study examined how morphological environment dictated coarticulatory patterns in adult and child (age 5-10) speakers of South Bolivian Quechua. To do so, I validated a Mel cepstrum measure that avoids the pitfalls inherent to child acoustics (dispersed harmonics, undersampling of spectral envelope) by replicating known relationships between manner of articulation and coarticulatory restraint. Next, I showed that adult SBQ speakers coarticulate more between morphemes than within them, further evidence that adults decompose complex words. Child SBQ speakers did not differentiate between morphological environments in this way. This suggested that they may instead store these inflected word forms holistically much later – up to age nine – than we would assume given children’s early morphological productivity.

## **References**

- 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. <https://doi.org/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.  
<https://doi.org/10.1177/0023830913484891>
- Arnon, I., & Snider, N. (2010). More than words: Frequency effects for multi-word phrases. *Journal of Memory and Language*, 62(1), 67–82. <https://doi.org/10.1016/j.jml.2009.09.005>
- Baayen, H. (1992). Quantitative aspects of morphological productivity. In G. Booij & J. van Marle (Eds.), *Yearbook of Morphology 1991* (pp. 109–149). Kluwer Academic Publishers. Retrieved from <http://www.sfs.uni-tuebingen.de/~hbaayen/publications/baayenYoM1992.pdf>
- Baker, R., Smith, R., & Hawkins, S. (2007). Phonetic differences between mis- and dis- in English prefixed and pseudo-prefixed words. In *Proceedings of the 16th ICPhS*. Saarbrücken.
- 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. <https://doi.org/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.
- Berko, J. (1958). The Child’s Learning of English Morphology. *Word*, 14(2–3), 150–177.  
<https://doi.org/10.1080/00437956.1958.11659661>
- Bernhardt, B. H., & Stemberger, J. P. (Eds.). (1998). *Handbook of phonological development: From the perspective of constraint-based nonlinear phonology*. New York, NY: Academic Press.
- Boersma, P., & Weenik, D. (2018). Praat: doing phonetics by computer (Version 6.0.42). Retrieved from [www.praat.org](http://www.praat.org)
- Bradlow, A. R. (2002). Confluent talker- and listener-oriented forces in clear speech production. *Laboratory Phonology*, 7.
- Browman, C. P., & Goldstein, L. (1989). Articulatory gestures as phonological units. *Phonology*, 6(02), 201–251. <https://doi.org/10.1017/S0952675700001019>
- Browman, C. P., & Goldstein, L. (1992). Articulatory Phonology: An Overview. *Phonetica*, 49(3–4), 155–180. <https://doi.org/10.1159/000261913>
- Cho, T. (2001). Effects of Morpheme Boundaries on Intergestural Timing: Evidence from Korean. *Phonetica*, 58(3), 129–162. <https://doi.org/10.1159/000056196>
- Chomsky, N. (1959). A review of BF Skinner’s Verbal Behavior. *Language*, 35(1), 26–58.
- Courtney, E. H., & Saville-Troike, M. (2002). Learning to construct verbs in Navajo and Quechua. *Journal of Child Language*, 29(03), 623–654. <https://doi.org/10.1017/S0305000902005160>
- Cychosz, M. (2018). *Coarticulatory practices as indicators of morphological decomposition* (Open Science Framework project). Retrieved from <https://osf.io/bjces/>

- Fenson, L., Marchman, V. A., Thal, D. J., Dale, P. S., Reznick, J. S., & Bates, E. (2007). *MacArthur-Bates Communicative Development Inventories User's Guide and Technical Manual* (2nd Edition). San Diego, CA: Singular.
- Ferguson, C. A., & Farwell, C. B. (1975). Words and sounds in early language acquisition. *Language*, 51(2), 419–439.
- Fikkert, P. (1994). *On the acquisition of prosodic structure*. Leiden & Amsterdam: Holland Institute of Generative Linguistics.
- Fikkert, P., & Levelt, C. (2008). How does Place fall into Place? The lexicon and emergent constraints in children's developing phonological grammar. In B. E. Dresher & K. Rice (Eds.), *Contrast in phonology: Theory, perception, and acquisition* (pp. 231–270). Walter de Gruyter.
- Flège, J. E. (1988). Anticipatory and carry-over nasal coarticulation in the speech of children and adults. *Journal of Speech Language and Hearing Research*, 31, 525–536.
- 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 Speech and Signal Processing Proceedings* (Vol. 1, pp. 393–396). Toulouse, France: IEEE. <https://doi.org/10.1109/ICASSP.2006.1660040>
- Gleitman, L. R., Gleitman, H., Landau, B., Wanner, E., & Newmeyer, F. (1988). Where learning begins: Initial representations for language learning. In F. J. Newmeyer (Ed.), *Linguistics: The Cambridge Survey: Volume 3, Language: Psychological and Biological Aspects* (pp. 150–193). Cambridge, MA: Cambridge University Press.
- 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. [https://doi.org/10.1044/1092-4388\(2008/07-0020\)](https://doi.org/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.
- Hale, M., & Reiss, C. (1998). Formal and Empirical Arguments concerning Phonological Acquisition. *Linguistic Inquiry*, 29(4), 656–683.
- Hanique, I., & Ernestus, M. (2012). The role of morphology in acoustic reduction. *Lingue e Linguaggio*, 11(2), 147–164.
- Hay, J. (2003). *Causes and Consequences of Word Structure*. New York and London: Routledge.
- Houston, D. M., & Jusczyk, P. W. (2003). Infants' Long-Term Memory for the Sound Patterns of Words and Voices. *Journal of Experimental Psychology: Human Perception and Performance*, 29(6), 1143–1154. <https://doi.org/10.1037/0096-1523.29.6.1143>
- Hsieh, L., Leonard, L. B., & Swanson, L. (1999). Some differences between English plural noun inflections and third singular verb inflections in the input: The contributions of frequency, sentence position, and duration. *Journal of Child Language*, 26(3), 531–543.
- Iskarous, K., Mooshammer, C., Hoole, P., Recasens, D., Shadle, C. H., Saltzman, E., & Whalen, D. H. (2013). The coarticulation/invariance scale: Mutual information as a measure of coarticulation

- resistance, motor synergy, and articulatory invariance. *The Journal of the Acoustical Society of America*, 134(2), 1271–1282. <https://doi.org/10.1121/1.4812855>
- Johnson, K. (1997). Speech perception without speaker normalization: An exemplar model. In *Talker variability in speech processing* (pp. 145–166). San Diego: Academic Press.
- Katz, W. F., Kripke, C., & Tallal, P. (1991). 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, H. (2005). Prosodic cues for morphological complexity in Dutch and English. *Language and Cognitive Processes*, 20(1–2), 43–73. <https://doi.org/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. [https://doi.org/10.1007/978-1-4613-8202-7\\_4](https://doi.org/10.1007/978-1-4613-8202-7_4)
- Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest Package: Tests in linear mixed-effects models. *Journal of Statistical Software*, 82(13), 1–26.
- Lahiri, A., & Marslen-Wilson, W. (1991). The mental representation of lexical form: A phonological approach to the recognition lexicon. *Cognition*, 38, 245–294.
- Lee, S., Potamianos, A., & Narayanan, S. (1999). Acoustics of children’s speech: Developmental changes of temporal and spectral parameters. *The Journal of the Acoustical Society of America*, 105(3), 1455–1468. <https://doi.org/10.1121/1.426686>
- Lee-Kim, S.-I., Davidson, L., & Hwang, S. (2013). Morphological effects on the darkness of English intervocalic /l/. *Laboratory Phonology*, 4(2). <https://doi.org/10.1515/lp-2013-0015>
- Lieven, E. V. M., Pine, J. M., & Baldwin, G. (1997). Lexically-based learning and early grammatical development. *Journal of Child Language*, 24(1), 187–219. <https://doi.org/10.1017/S0305000996002930>
- Lindblom, B. (1963). Spectroraphic Study of Vowel Reduction. *Journal of the Acoustical Society of America*, 35(11), 1773–1781.
- Losiewicz, B. L. (1995). Word frequency effects on the acoustic duration of morphemes. *Journal of the Acoustical Society of America*, 97(5), 3243–3243. <https://doi.org/10.1121/1.411745>
- McAllister Byun, T., Inkelas, S., & Rose, Y. (2016). The A-map model: Articulatory reliability in child-specific phonology. *Language*, 92(1), 141–178. <https://doi.org/10.1353/lan.2016.0000>
- McAllister Byun, T., & Tessier, A.-M. (2016). Motor Influences on Grammar in an Emergentist Model of Phonology: Motor and Grammar in Emergentist Phonology. *Language and Linguistics Compass*, 10(9), 431–452. <https://doi.org/10.1111/lnc3.12205>
- Mealings, K. T., Cox, F., & Demuth, K. (2013). Acoustic investigations into the later acquisition of syllabic -es plurals. *Journal of Speech Language and Hearing Research*, 56(4), 1260–1271. [https://doi.org/10.1044/1092-4388\(2012/12-0163\)](https://doi.org/10.1044/1092-4388(2012/12-0163))

- Mealings, K. T., & Demuth, K. (2014). The role of utterance length and position in 3-year-olds' production of third person singular. *Journal of Speech Language and Hearing Research*, 57(2), 484–494. [https://doi.org/10.1044/2013\\_JSLHR-L-12-0354](https://doi.org/10.1044/2013_JSLHR-L-12-0354)
- Menn, L., Schmidt, E., & Nicholas, B. (2013). Challenges to theories, charges to a model: the Linked-Attractor model of phonological development. In M. M. Vihman & T. Keren-Portnoy (Eds.), *The emergence of phonology: Whole-word approaches and cross-linguistic evidence* (pp. 460–502).
- Mooshammer, C., Hoole, P., & Geumann, A. (2006). Interarticulator cohesion within coronal consonant production. *Journal of the Acoustical Society of America*, 120(2), 1028–1039. <https://doi.org/10.1121/1.2208430>
- Mousikou, P., Strycharczuk, P., Turk, A., Rastle, K., & Scobbie, J. M. (2015). Morphological effects on pronunciation. In *Proceedings of the 18th ICPHS*. Glasgow, Scotland.
- Nijland, L., Maassen, B., Meulen, S. V. der, Gabreëls, F., Kraaimaat, F. W., & Schreuder, R. (2002). Coarticulation patterns in children with developmental apraxia of speech. *Clinical Linguistics & Phonetics*, 16(6), 461–483. <https://doi.org/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(2), 379. <https://doi.org/10.1044/jshr.3902.379>
- Noiray, A., Ménard, L., & Iskarous, K. (2013). The development of motor synergies in children: Ultrasound and acoustic measurements. *Journal of the Acoustical Society of America*, 133(1), 444–452. <https://doi.org/10.1121/1.4763983>
- Pierrehumbert, J. B. (2003). Phonetic Diversity, Statistical Learning, and Acquisition of Phonology. *Language and Speech*, 46(2–3), 115–154. <https://doi.org/10.1177/00238309030460020501>
- Plag, I., Homann, J., & Kunter, G. (2017). Homophony and morphology: The acoustics of word-final S in English. *Journal of Linguistics*, 53(01), 181–216. <https://doi.org/10.1017/S0022226715000183>
- Pluymaekers, M., Ernestus, M., Baayen, R. H., & Booij, G. (2010). Morphological effects on fine phonetic detail: The case of Dutch -igheid. *Laboratory Phonology*, 10, 511–531. <https://doi.org/10.1515/9783110224917.5.511>
- R Core Team. (2018). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.
- Recasens, D., & Espinosa, A. (2009). An articulatory investigation of lingual coarticulatory resistance and aggressiveness for consonants and vowels in Catalan. *Journal of the Acoustical Society of America*, 125(4), 2288–2298. <https://doi.org/10.1121/1.3089222>
- Recasens, D., Pallarès, M. D., & Fontdevila, J. (1997). A model of lingual coarticulation based on articulatory constraints. *Journal of the Acoustical Society of America*, 102(1), 544–561. <https://doi.org/10.1121/1.419727>

- Repp, B. H. (1986). Some observations on the development of anticipatory coarticulation. *Journal of the Acoustical Society of America*, 79(5), 1616–1619. <https://doi.org/10.1121/1.393298>
- Robb, M., & Wolk, L. (1997). A note on prespeech and early speech coarticulation. *Logopedics Phoniatrics Vocology*, 22(3), 99–104. <https://doi.org/10.3109/14015439709075321>
- Rubertus, E., Abakarova, D., Ries, J., & Noiray, A. (2013). Anticipatory V-to-V Coarticulation in German Preschoolers. In *Phonetik und Phonologie im deutschsprachigen Raum* (Vol. 12, p. 5). Munch, Germany.
- Seidl, A., Onishi, K. H., & Cristià, A. (2014). Talker Variation Aids Young Infants' Phonotactic Learning. *Language Learning and Development*, 10(4), 297–307. <https://doi.org/10.1080/15475441.2013.858575>
- Sereno, J. A., Baum, S. R., Mearan, G. C., & Lieberman, P. (1987). Acoustic analyses and perceptual data on anticipatory labial. *Journal of the Acoustical Society of America*, 81, 512–519.
- Sereno, J. A., & Lieberman, P. (1987). Developmental aspects of lingual coarticulation. *Journal of Phonetics*, 15, 247–257.
- Skinner, B. F. (1957). *Verbal Behavior*. Acton, MA: Copley Publishing Group.
- Smith, A., & Goffman, L. (1998). Stability and Patterning of Speech Movement Sequences in Children and Adults. *Journal of Speech Language and Hearing Research*, 41(1), 18. <https://doi.org/10.1044/jslhr.4101.18>
- Song, J. Y., Demuth, K., Evans, K., & Shattuck-Hufnagel, S. (2013). Durational cues to fricative codas in 2-year-olds' American English: Voicing and morphemic factors. *Journal of the Acoustical Society of America*, 133(5), 2931–2946. <https://doi.org/10.1121/1.4795772>
- Song, J. Y., Demuth, K., Shattuck-Hufnagel, S., & Ménard, L. (2013). The effects of coarticulation and morphological complexity on the production of English coda clusters: Acoustic and articulatory evidence from 2-year-olds and adults using ultrasound. *Journal of Phonetics*, 41(3–4), 281–295. <https://doi.org/10.1016/j.wocn.2013.03.004>
- Sugahara, M., & Turk, A. (2009). Durational correlates of English sublexical constituent structure. *Phonology*, 26(03), 477. <https://doi.org/10.1017/S0952675709990248>
- Swingle, D., & Aslin, R. N. (2007). Lexical competition in young children's word learning. *Cognitive Psychology*, 54(2), 99–132. <https://doi.org/10.1016/j.cogpsych.2006.05.001>
- Tomasello, M. (2003). *Constructing a language: A usage-based approach to child language acquisition*. Cambridge, MA: Harvard University Press.
- Vihman, M. M., & Keren-Portnoy, T. (2013). *The emergence of phonology: Whole-word approaches and cross-linguistic evidence*. Cambridge, MA: Cambridge University Press.
- Vorperian, H. K., & Kent, R. D. (2007). Vowel Acoustic Space Development in Children: A Synthesis of Acoustic and Anatomic Data. *Journal of Speech Language and Hearing Research*, 50(6), 1510–1545. [https://doi.org/10.1044/1092-4388\(2007/104\)](https://doi.org/10.1044/1092-4388(2007/104))
- Werker, J. F., & Curtin, S. (2005). PRIMIR: A Developmental Framework of Infant Speech Processing. *Language Learning and Development*, 1(2), 197–234. <https://doi.org/10.1080/15475441.2005.9684216>



- Whalen, D. H. (1990). Coarticulation is largely planned. *Journal of Phonetics*, 18, 3–35.
- Whiteside, S. P., & Hodgson, C. (2000). Speech patterns of children and adults elicited via a picture-naming task: An acoustic study. *Speech Communication*, 32(4), 267–285.  
[https://doi.org/10.1016/S0167-6393\(00\)00013-3](https://doi.org/10.1016/S0167-6393(00)00013-3)
- Zharkova, N., Hewlett, N., & Hardcastle, W. J. (2008). An Ultrasound Study of Lingual Coarticulation in Children and Adults, 4.
- Zharkova, N., Hewlett, N., & Hardcastle, W. J. (2011). Coarticulation as an Indicator of Speech Motor Control Development in Children: An Ultrasound Study. *Motor Control*, 15(1), 118–140.  
<https://doi.org/10.1123/mcj.15.1.118>
- Zharkova, N., Hewlett, N., Hardcastle, W. J., & Lickley, R. J. (2014). Spatial and Temporal Lingual Coarticulation and Motor Control in Preadolescents. *Journal of Speech Language and Hearing Research*, 57(2), 374. [https://doi.org/10.1044/2014\\_JSLHR-S-11-0350](https://doi.org/10.1044/2014_JSLHR-S-11-0350)

---

<sup>i</sup> Though acoustic salience has been argued to predict order of morpheme acquisition (Gleitman, Gleitman, Landau, Wanner, & Newmeyer, 1988: 155-160; Hsieh, Leonard, & Swanson, 1999)

<sup>ii</sup> Thank you to Gillian Gallagher for reminding me of this.