Practice and experience predict coarticulation in child speech

Margaret Cychosz[1,2], Benjamin Munson[3], and Jan R. Edwards[1]

[1]Department of Hearing and Speech Sciences, University of Maryland, College Park

[2]Center for Comparative and Evolutionary Biology of Hearing, University of Maryland, College Park

[3]Department of Speech-Language-Hearing Sciences, University of Minnesota, Twin Cities

Author Note

Author to whom correspondence should be addressed: Margaret Cychosz, 0100 Samuel J. LeFrak Hall, University of Maryland, College Park, College Park, USA, 20742. Email: mcychosz@umd.edu.

Abstract

Much research in child speech development suggests that young children coarticulate more than adults. There are multiple, not mutually-exclusive, explanations for this pattern. First, children may coarticulate more because they are limited by immature motor control. Children might also coarticulate more because they initially represent phonological segments in larger, more holistic units such as syllables or feet. We tested these explanations by evaluating how four-year-old children's language experience and practice predict their coarticulation between adjacent segments in real words and paired nonwords. Children with larger vocabularies coarticulated less, especially in real words, though there were no reliable coarticulatory differences between real words and nonwords after controlling for word duration. Children who vocalized more throughout a daylong audio recording also coarticulated less. Quantity of child vocalizations was more predictive of the degree of children's coarticulation than a measure of receptive language experience, adult word count. Overall, these results suggest strong roles for children's phonological representations, as well as their immature motor control, for coarticulatory development.

Keywords: speech development, speech production, coarticulation, phonology, word repetition, naturalistic recording

1   Introduction

A long standing question in language development is whether the representation of linguistic forms differs between adults and children. Numerous experimental paradigms and techniques examining speech perception and lexical access (e.g. head-turn preference, eye-tracking), have addressed this question (Majorano, Vihman, & DePaolis, 2014; Storkel & Morrisette, 2002; Swingley & Aslin, 2002). However, children's own speech production has also been studied to shed light on the nature of their phonological representations (Ferguson & Farwell, 1975; Goffman, Smith, Heisler, & Ho, 2008; Nittrouer, Studdert-Kennedy, & McGowan, 1989; Nittrouer, Studdert-Kennedy, & Neely, 1996; Noiray, Popescu, et al., 2019; Noiray, Abakarova, Rubertus, Krüger, & Tiede, 2018; Redford, 2018; Song, Demuth, Shattuck-Hufnagel, & Ménard, 2013; Song, Demuth, Evans, & Shattuck-Hufnagel, 2013; Zharkova, Hewlett, & Hardcastle, 2011). These works have demonstrated that speech production patterns, such as the duration, variability, and especially fluency of speech, can lend insight into speakers' phonological representations and planning.

One production metric that is frequently used to measure speech fluency in development is coarticulation, or the temporal and gestural overlap that results from producing two speech sounds in a sequence (Gerosa, Lee, Giuliani, & Narayanan, 2006; Goffman et al., 2008; Zharkova et al., 2011). Coarticulation is not simply noise in the speech signal. It conveys important auditory-acoustic information for speakers and listeners alike. Appropriate coarticulatory overlap, such as the ability to anticipate forthcoming speech segments, indicates mature, adult-like speech (Barbier et al., 2020; Bradlow, 2002; Whalen, 1990). Children speak more slowly overall, and with less coordinated movement (Green, Moore, Higashikawa, & Steeve, 2000; Goffman, Gerken, & Lucchesi, 2007; Lee, Potamianos, & Narayanan, 1999). It might then be logical to conclude that children coarticulate less than adults. Some studies have indeed concluded that children show less long-distance anticipatory coarticulation than adults, for example across syllable boundaries (Barbier et al., 2020). However, a consensus is now growing that children actually coarticulate more than adults, both over longer distances (Rubertus & Noiray, 2018, 2020; cf. Goffman et al.,

2008) and between adjacent segments within syllables (Nittrouer et al., 1989, 1996; Noiray, Wieling, Abakarova, Rubertus, & Tiede, 2019; Rubertus & Noiray, 2020; Zharkova et al., 2011; Zharkova, Hewlett, Hardcastle, & Lickley, 2014).

1.1   The development of coarticulation

A number of explanations have been offered for why children coarticulate between adjacent segments more than adults. Children's coarticulation could reflect 1) low-level pressures upon speech due to the protracted development of fine motor control, 2) the phonological reorganization of speech from more holistic units, such as syllables, into phonemes, 3) domain-general cognitive pressures from children's limited working memory and planning capacities, or a combination of these explanations.

First, children could coarticulate more because their fine motor control develops gradually throughout childhood and well into early adolescence (Barbier et al., 2020; Kent, 1976; Perkell, 2013; Walsh & Smith, 2002). The spatiotemporal coordination of the lips and jaw, for example, undergoes significant development between ages two and six—younger children tend to rely more on vertical jaw movement, only learning to simultaneously coordinate both the lips and jaw with age (Green et al., 2000). Unsurprisingly, this motor control development has implications for children's phonetic production, including coarticulation. For example, Zharkova, Hardcastle, and Gibbon (2018) measured coarticulation within /əsV/ and /əʃV/ sequences in children aged 7;0-13;0 and adults and found both that the ability to differentiate between /s/ and /ʃ/ increased with age and that the youngest children were most variable. The authors attribute these findings to the children's developing control of the tongue, especially its coordination with the lips and jaw. Rubertus and Noiray(2018) likewise attribute 3;6-7;0 children's coarticulation patterns—all children engaged in more V-V anticipatory coarticulation than adults—to protracted lingual coordination.

Another reason younger children coarticulate more than adults and older children is if they initially represent language more holistically than adults, in units such as syllables, feet, or morae,

instead of phoneme-sized segments. The exact unit of representation may vary by lexical item and depend upon the child's developmental stage (Davis & Redford, 2019) as well as their familiarity with the phoneme sequence(s) or word in question. More familiar, frequent words likely have more segmental representations (Edwards, Beckman, & Munson, 2004), but representations may also be redundant: a given word may be encoded both holistically as well as segmentally, as is often argued to be the case even in adult phonology (Pierrehumbert, 2003). Overall, however, these larger speech units would be activated during children's speech planning resulting in significant anticipatory effects of adjacent segments on one another.[1] Anticipatory coarticulation between adjacent phones then decreases with age as phonological units reorganize into smaller segments in a phenomenon of child phonology often termed PHONOLOGICAL REORGANIZATION (Goodell & Studdert-Kennedy, 1992; Metsala, 1999, 1999; Nittrouer et al., 1989, 1996; Noiray, Popescu, et al., 2019; Noiray et al., 2018; Redford, 2018; Zharkova et al., 2011).[2] This phonological reorganization, children's advancement from more holistic to segmental speech representations, is driven by myriad factors including vocabulary growth (Metsala & Walley, 1998), construction of (denser) phonological neighborhoods (Storkel & Hoover, 2011), and, starting around age 5, phonological awareness/literacy development (Metsala, 1999). For example, vocabulary growth and the restructuring of the lexicon drive reorganization because children have more word types in their mental lexicons—and stronger relationships between sublexical units within and across neighborhoods—over which they can generalize and calculate statistics to construct a segmental

---

[1] The literature does, however, remain somewhat mixed on the directionality of coarticulatory development. Some studies have been unable to find a difference between adult and child coarticulation between adjacent segments (Katz, Kripke, & Tallal, 1991; Noiray, Ménard, & Iskarous, 2013; Sereno & Lieberman, 1987; Zharkova et al., 2018), while others have found that children coarticulate less (Kent, 1983).

[2] It is important to underscore how this phonological reorganization could play out differently cross-linguistically. Results from tests of phonological awareness in Mandarin-speaking children, for example, suggest that syllable and tone awareness are more important for character recognition than sub-syllabic awareness (McBride-Chang, Bialystok, Chong, & Li, 2004; Shu, Peng, & McBride-Chang, 2008). The developmental trajectories of phonological reorganization may then differ based on the language of exposure.

phonology (Storkel, 2002).

Finally, it is possible that children coarticulate more than adults because domain-general constraints, such as limited short-term working memory capacities, affect children's speech planning. To the authors' knowledge, no study has explicitly tested the effects of short-term working memory, or related correlates like cognitive load, upon children's phonetic production or coarticulation. Nevertheless, several lines of research on adult phonetics suggest that children's memory constraints could explain some of their coarticulation patterns. First, Franich (2015) showed a modest effect of cognitive load, instantiated as digit recall, upon tonal coarticulation in adult speakers of Mandarin: speakers' anticipatory coarticulation increased under heightened cognitive load conditions. Second, more indirectly, increased cognitive or work load appears to affect adult speakers' vowel space structure and speech articulation rate (though the direction of the effect is not always consistent) (Huttunen, Keränen, Pääkkönen, Päivikki Eskelinen-Rönkä, & Leino, 2011; Lively, Pisoni, Summers, & Bernacki, 1993). Since short-term working memory capacities and their ability to predict outcomes like fluid intelligence change throughout childhood (Engel de Abreu, Conway, & Gathercole, 2010; Hitch & Halliday, 1983), it is plausible that the demands of speech production may be more burdensome for children than adults, which could result in children's increased coarticulation. This possibility notwithstanding, the traditional explanations for children's greater coarticulation are protracted fine motor control and phonological reorganization. The current study focuses on disentangling these two explanations, while acknowledging a possible role for additional variables such as short-term working memory.

There is no doubt that fine motor control develops with age, with clear implications for children's speech faculty and coarticulation (Barbier et al., 2020; Zharkova et al., 2018). However, there is also ample evidence suggesting that child coarticulation is also explained by phonological reorganization. Support for this idea has often come from studies with cross-sectional designs: numerous studies, measuring coarticulation between the same segments using the same methods (e.g. acoustics, articulatory), have shown that children coarticulate less between adjacent segments as they age. Since the mental lexicon reorganizes, and phonological awareness skills increase, with

age (Metsala, 1999; Storkel & Hoover, 2011), age-related changes in coarticulation are cited as evidence that phonological reorganization explains children's coarticulation. In a series of studies, Nittrouer et al. (1989) and Nittrouer et al. (1996) found that as children aged 3;0-8;0 got older, they tended to coarticulate less between segments in fricative-vowel sequences. Similar conclusions were drawn by Zharkova et al. (2011) who measured children's (aged 6;0-9;0) coarticulation using articulatory imaging methods (ultrasound) and found that the degree of coarticulation decreased with age. Finally, in a cross-sectional sample of children aged 3;6-7;0, Noiray et al. (2018) also found that children coarticulated less in CVC sequences as they aged. However, the authors warn that this may not be attributable to a single representational strategy, such as the representation of speech in syllables or chunks. Rather, there may be different organizational—and thus production—strategies for different segments. Children must learn the degree of coarticulation permitted between phones in sequences such as /bi/ as opposed to that permitted in /di/ or /gi/. (See also Zharkova [2018] for discussion of segment-specific effects on coarticulation.)

The aforementioned studies used a cross-sectional study design to support the idea that children's coarticulatory patterns reflect phonological reorganization. However, cross-sectional designs cannot rule out the exclusive role of speech motor control for coarticulatory development because motor control coincides with phonological reorganization over the course of development—both develop and change as children age. As a result, the underlying causes behind children's coarticulatory patterns have somewhat eluded researchers because many common methodological designs have not attempted to decouple the motor maturation and phonological reorganization explanations (cf. Noiray, Popescu, et al., 2019).

Additional evidence in support of phonological reorganization needs to come from predictors that, unlike age, mature at least somewhat independently of speech motor control. Noiray, Popescu, et al. (2019) correlated children's coarticulation with two such independent measures of linguistic experience—vocabulary size and phonological awareness (i.e. metalinguistic recognition of phoneme-sized units). Neither of these factors correlates directly with motor maturation, though it is certainly the case that children with larger vocabularies may have stronger motor

skills because they have more experience articulating diverse syllables and word types. The authors found that children (aged 4;0-7;0) with greater phonological awareness, and larger expressive vocabularies, coarticulated less between segments, independent of chronological age. From these findings the authors conclude that children's coarticulation is likely explained by both speech motor control maturation and phonological reorganization.

To be clear, the motor maturation and phonological reorganization explanations for children's coarticulation are not mutually exclusive. Children's fine motor planning is continuously honed and developed as children age, with implications for child coarticulation—these conclusions are clear (Barbier et al., 2020; Rubertus & Noiray, 2018; Zharkova et al., 2018). The objective of the present analysis is to determine to what extent children's speech representations *also* explain their coarticulatory patterns, over and above the role of fine motor development. To do so, the current study expands upon Noiray, Popescu, et al. (2019) by evaluating the role of several additional, independent factors on child coarticulation: speech planning, vocabulary size, and speech practice. In typically-developing children, these factors should develop somewhat independently of speech motor control and chronological age (a direct correlate of motor control). If one or more of the tested variables predicts the degree of children's coarticulation, it suggests extra-motoric influences on the development of coarticulation, shedding further light on the underlying causes of this speech phenomenon in children. Note that it may never be possible to entirely disentangle motoric and representational explanations for child coarticulation since speech planning skills, vocabulary size, and speech practice are likely positively correlated with motor maturation: children who speak more throughout the day and have larger vocabularies have likewise honed their fine motor skills. Nevertheless, the variables tested here develop independently of chronological age so any effect they have on speech patterns will advance our understanding of the causal mechanisms behind children's coarticulation.

First, the current study evaluates the role of speech planning on children's coarticulation by measuring coarticulation between phoneme sequences embedded in nonwords and matched real words (e.g. [su] in the nonword *sudras* versus the real word *suitcase*). Unlike real words, nonwords

must be produced without the support of lexical, semantic, or complete phonetic representations (Chiat & Roy, 2007; Cychosz, Erskine, Munson, & Edwards, 2020; Gathercole, Willis, Emslie, & Baddeley, 1991; Keren-Portnoy, Vihman, DePaolis, Whitaker, & Williams, 2010). Consequently, nonwords place additional demands upon children during speech planning which could result in less fluent production, or more coarticulation.

Second, in an attempt to replicate Noiray, Popescu, et al. (2019), this study evaluates the role of vocabulary size on coarticulation—though we expand upon Noiray, Popescu, et al. (2019) by comparing the role of receptive and expressive vocabularies. Vocabulary size could be expected to predict coarticulation because it predicts the emergence of discrete phonological units during development (Edwards et al., 2004; Sosa & Stoel-Gammon, 2012; Stoel-Gammon, 2011; Storkel & Morrisette, 2002) and, most relevant for the current work, speech accuracy and fluency (Cychosz et al., 2020; Edwards et al., 2004; Metsala, 1999; Munson, Edwards, & Beckman, 2005; Zamuner, 2009). Children with larger receptive vocabularies have more unique word types over which they can generalize to infer the importance of certain segments in their native language(s). Those with larger expressive vocabularies also have more practice producing more unique words, and thus phoneme combinations, potentially resulting in further segmental production mastery (Beckman & Edwards, 2010). For these reasons, both expressive and receptive vocabulary sizes are expected to predict degree of coarticulation in children.

The third predictive factor for coarticulation evaluated here is speech practice. The role of speech practice is measured by correlating elements from children's daily language experience—namely the frequency of child vocalizations—with coarticulation. Children's daily language experiences been shown to predict a number of developmental milestones in speech and language including lexical processing (Weisleder & Fernald, 2013), expressive and receptive vocabulary sizes (Hart & Risley, 1995; Hoff, 2003; Mahr & Edwards, 2018), and babbling complexity (Ferjan Ramírez, Lytle, Fish, & Kuhl, 2019). It is thus reasonable to predict that children's daily linguistic experiences, specifically how frequently they vocalize, could likewise predict their coarticulation patterns.

The factors of speech planning, vocabulary size, and speech practice all concern the role of speech production practice for coarticulatory development. For example, we anticipate that real words will be produced more fluently than nonwords in part because children have more practice accessing and producing the real words. A similar argument can be made for the potential roles of vocabulary size and vocalization frequency on child coarticulation. The following section outlines why speech production practice may be expected to predict children's coarticulation patterns.

1.2 The role of production practice for phonological development

A child's developing anatomy is highly transient with anatomical changes that are non-linear (Vorperian et al., 2005) and non-uniform because different speech articulators mature at distinct rates (Nittrouer, 1993). These elements of phonological development present a challenge for children who must establish accurate, replicable, articulatory-acoustic mappings in the face of rapid, uneven anatomical change. It is analogous to shooting an arrow at a bullseye with an ever-changing ratio of arm to bow length.

Yet articulatory and production practice is a vital component of phonological development (Bruderer, Danielson, Kandhadai, & Werker, 2015; Davis & Redford, 2019; Keren-Portnoy et al., 2010; McAllister Byun & Tessier, 2016; Menn, Schmidt, & Nicholas, 2013; Vihman, 2017; Zamuner, Strahm, Morin-Lessard, & Page, 2018). The effect that production practice appears to have upon children's speech perception, in particular, suggests that early representations consist of abstract phonological categories *and* accompanying articulatory routines. Production practice can also impact how infants process speech thorough an "articulatory filter" in early phonological development (DePaolis, Vihman, & Keren-Portnoy, 2011; DePaolis, Vihman, & Nakai, 2013; Laing & Bergelson, 2020; Vihman, 1993, 2017). According to this theory, infants pay particular attention to phonemes in their ambient environment that they can already produce, or are just starting to produce, during early babble. The effect becomes cyclical because the salient sounds from the environment, those that pass through the child's articulatory filter, in turn encourage the child to attempt to produce them more. Consequently, articulatory practice shapes the way that

infants process speech, biasing them towards those sounds that they can already produce, and shaping their phonological representations through increased practice.

The articulatory routines associated with phonological representations continue to be updated into the preschool- and school-aged years where children's speech productions may still encode articulatory traces of particular vocal tract configurations that change over development (Davis & Redford, 2019; McAllister Byun & Tessier, 2016). Both these articulatory traces and the accompanying acoustics are encoded into phonological representations, updating them to reflect the child's increased experience with language and changed anatomy (e.g. lengthened vocal tract). In this way, articulatory schemata comprise a central part of phonological representations throughout childhood.

There is ample evidence for a production effect, or an effect of speech practice upon phonological representations, in development. DePaolis et al. (2011) tested infants aged 0;10-1;4 on preference for "own production" consonants (consonants that the infant was already producing consistently) versus "other production" consonants (consonants that the infant was not yet producing consistently). In a head-turn preference task, the authors found that those infants who were already producing multiple consonants paid more attention to the "other production" consonants. At the same time, there was no effect for infants who were producing one or no consonants at the time of testing. The authors suggest that, in conjunction with early speech perception, the infants' preference for novel consonants—the consonants that they are just starting to produce and will soon produce regularly—may strengthen early speech acquisition. For one thing, these novel sounds may become more salient for an infant in the input as "production [may] initiate shifts in the way that the infant processes input speech" (2011:599). A resultant "auditory feedback loop" (Majorano et al., 2014) could facilitate phonological development as children's articulatory practice coincides with acoustic-auditory speech representations (see DePaolis et al. [2013] & Keren-Portnoy et al. [2010] for similar conclusions).

In older children, production practice in the form of repeating whole words could also be advantageous for phonological development (Icht & Mama, 2015; McAllister Byun, Inkelas, &

Rose, 2016; McAllister Byun & Tessier, 2016; cf. Zamuner et al., 2018. The argument behind this production effect in older children is that by producing a word, children are better able to encode the word's auditory-acoustic signature in semantic and phonological representations. The time required to process a new word may decrease if the word contains a well-practiced motor scheme, for example a phonotactically-probable sequence, learned elsewhere (Storkel & Maekawa, 2005; Storkel & Hoover, 2011). Icht and Mama (2015) found that 5-year-old children remembered novel words better when they looked at a picture of a novel object and produced the new word out loud than when they looked at the object and only heard the experimenter repeat the item name. Most recently, however, Zamuner et al. (2018) found that 4;5-6;0 children remembered novel words better during a "hear only" condition than a "hear and produce" condition, suggesting that the production practice effect may vary by task demand and the child's developmental stage.

A final way that the production effect could affect phonological development is through parental engagement. Infants and children, of various ages, who produce more speech tend to elicit more parental responses (Franklin et al., 2014; Goldstein & Schwade, 2008; McGillion et al., 2017; Pretzer, Lopez, Walle, & Warlaumont, 2019; Warlaumont, Richards, Gilkerson, & Oller, 2014). Furthermore, positive correlations have been found between the number of words spoken by adults in the environment and the frequency of children's own vocal productions for children aged 0;2-4;0 (Gilkerson & Richards, 2009; Orena, Byers-Heinlein, & Polka, 2019; Weisleder & Fernald, 2013).

Franklin et al. (2014) and Goldstein and Schwade (2008) illustrate how this parent-child engagement effect plays out for infant vocalizations. Franklin et al. (2014) six-month-old infants who were presented with a still face after interacting with parents vocalized more to re-engage parents suggesting that children engage in vocal feedback loops from a very young age. Elsewhere, Goldstein and Schwade (2008) found that nine-month-old infants adapted the structure of their babble shapes (i.e. proportion of CV syllables) to match caregiver productions, but only when the caregivers spoke contingently to the infants by, for example, moving closer to, smiling at, or touching the infant. Together these results from studies of infant-caregiver behavior suggest that, from the first year of life, children's vocalizations engage parents (Albert, Schwade, & Goldstein,

2018), whose responses, in turn, encourage infants to vocalize more.

For preschoolers, the effect of parent-child vocal engagement upon speech production outcomes could play out in a number of ways. First, young children who hear more input from caregivers may be more likely to respond and engage in conversations, thereby practicing various components of production (fine motor movement, speech planning). Second, since quantity of caregiver input is one of the strongest predictors of children's vocabulary size (Hoff, 2003), and preschool aged children who have larger vocabularies have more mature phonological outcomes (Stoel-Gammon, 2011; Storkel & Morrisette, 2002), there could be an indirect role of input upon production outcomes: caregiver speech drives children's vocabulary growth which predicts phonological maturity. Finally, children who hear more caregiver speech may simply have greater access to adult speech models and mature speech patterns. Properties of caregiver speech are known to predict phonetic outcomes. In infants, fine-grained differences in caregivers' fricative productions predict /s/ and /ʃ/ discrimination (Cristià, 2011). And in children aged 4;0-8;11, caregivers' bilingual language dominance predicts children's vowel category dispersion (Cychosz, 2020). Thus, caregiver-specific properties of the input may be another mechanism linking adult-child interactions to children's phonetic outcomes.

To summarize, speech production and practice appear to predict phonological development. This production effect starts early, as infants engage in social feedback loops with caregivers that result in increased infant vocal exploration and practice (Franklin et al., 2014; Goldstein & Schwade, 2008; McGillion et al., 2017; Pretzer et al., 2019; Warlaumont et al., 2014). Parental input is still correlated with the frequency of children's vocalizations into toddlerhood (Gilkerson & Richards, 2009; Weisleder & Fernald, 2013) where various mechanisms (motoric practice, vocabulary growth, mature speech models) may explain a connection between adult-child interactions and children's phonetic outcomes. The practice effect also manifests in infants and young toddlers who, during vocal exploration and late babbling stages, produce those sounds that are most salient and frequent in their environments (DePaolis et al., 2011, 2013; Laing & Bergelson, 2020; Vihman, 1993, 2017). Finally, speech practice also appears to facilitate lexical

acquisition, with ramifications for phonological development (Sosa & Stoel-Gammon, 2012; Stoel-Gammon, 2011; Zamuner, 2009): there may be an advantageous effect of production whereby children learn new words faster by articulating them than hearing alone (Icht & Mama, 2015; but cf. Zamuner et al., 2018). The strong predictive role of practice leads us to make several predictions for the factors studied in the current work, which are outlined below.

2   The current study

The goal of the present study is to study the role of three factors that could potentially underlie children's coarticulatory patterns in spoken language: speech planning, vocabulary size (expressive and receptive), and speech practice. Again, the tendency for children to coarticulate is almost certainly due, in part, to their overall inexperience with speech production and immature speech motor control development (Barbier et al., 2020; Goffman et al., 2007; Green et al., 2000; Rubertus & Noiray, 2018; Zharkova et al., 2018). Here we are interested in the potential contribution of factors that may develop more independently of fine motor development.

To test how speech organization may contribute to children's coarticulation, we examined four-year-old children's spoken language production in nonwords and corresponding real words. Each real word had a paired, word-initial CV syllable in the nonword (e.g. [su] in *sudras* and *suitcase*). This lexicality condition allowed us to test how both speech and lexical planning contribute to children's tendency to coarticulate. The children have experience hearing and, crucially, producing, the real words but not the nonwords. How will children's existing lexical representations and articulatory schemata affect their spoken coarticulation patterns? We made the following prediction concerning child coarticulation in real words and paired nonwords:

1. Because children already have experience producing the real words, they will coarticulate less within CV sequences when they are embedded in real words (e.g. *suitcase*) than nonwords (e.g. *sudras*).[3]

———

[3] Hypotheses one and two were originally pre-registered to predict that 1) children would coarticulate more in real words than nonwords and 2) that children with larger vocabularies would coarticulate more than children with

We then correlated the children's spoken language patterns with their vocabulary size (receptive and expressive) and elements from each child's naturalistic language use (measured with daylong audio recordings). We made two predictions concerning these factors:

2. Noiray, Popescu, et al. (2019) found that children with larger expressive vocabularies coarticulated less. As a result, we predict that children who have larger lexicons, estimated as the size of their expressive or receptive vocabularies, will coarticulate less between the segments in CV sequences. We build on the work of Noiray, Popescu, et al. (2019) by comparing the roles of expressive and receptive vocabulary sizes for coarticulation, though we do not make a specific prediction for which measure of vocabulary size will be the stronger predictor.

3. Children who produce more language, quantified as the frequency of child vocalizations during daylong audio recordings, will coarticulate less between segments in CV sequences.

3   Methods

3.1   Participants

One hundred and three children (56 girls, 47 boys) aged 3;3 to 4;4 (years;months, mean=3;5, SD=0;3) participated in the study. All children were monolingual speakers of English participating in a longitudinal study studying children's phonological development. The current data were collected on the second of the child's three scheduled visits, approximately one year after the first visit. All families consented to participate in the research upon their initial visit to the lab. All in-lab tasks were completed at the University of Wisconsin, Madison or the University

---

smaller vocabularies (Cychosz, Edwards, Munson, & Johnson, 2019). However, these hypotheses were registered prior to the publication of Noiray, Popescu, et al. (2019). Those authors found that greater linguistic experience—vocabulary size and phonological awareness—predicted less coarticulation. In light of those results, we eventually adjusted our hypotheses but for the sake of open science, transparency, and replicability, we acknowledge that our original hypotheses actually hypothesized the reverse.

of Minnesota, Twin Cities. Nonword and real word repetition accuracy scores from some of these children (approximately 83%) was previously reported in Cychosz et al., 2020.

Each participant passed a hearing screening in at least one ear at 25dB for 1000, 2000, and 4000Hz. Per parental self-report, 90 (87.4%) of the children had normal speech and hearing development. The 13 remaining children were identified as late talkers by their caregivers. It should be noted that this was a caregiver description, rather than a clinical diagnosis. The caregiver-identified late talkers and caregiver-identified typically-developing children are differentiated in the Results and in statistical modeling.

Socioeconomic status, quantified as mother's education level, was reported by caregivers. 38% of mothers had a graduate degree, 32% a college degree, 21% some college/trade school/associate degree, 7% a high school diploma, and 2% did not have a high school diploma.

3.2   Word repetition tasks

For the in-lab data collection phase, children completed two word repetition tasks: a nonword repetition task, where participants repeated nonce words after an adult model speaker, and a real word repetition task. The experiments and pre-task assessments were completed in two 1-hour test sessions on different days, usually separated by about a week. Children always completed the real word repetition task during the first testing session and the nonword repetition task during the second session. The decision was made to have the children complete the repetition tasks on separate days because real word and nonword tasks are generally given separately (Chiat & Roy, 2007). In addition, we chose not to counterbalance the presentation order of the two word repetition tasks because in our experience conducting speech elicitation paradigms with children of this age, we have found that children sometimes perceive phonotactically-probable nonwords as real words. Furthermore, children of this age do not readily switch back and forth easily between nonwords and real words. See Cychosz et al.(2020) for further justification of the presentation order for the repetition tasks.

3.3   Repetition task materials

The real word stimuli for the word repetition tasks were chosen from lists such as the "Toddler Says" portion of the MacArthur Bates Communicative Development Inventory (Fenson et al., 2007), to ensure familiarity to the majority of children in this age group (see Table 1). Then, the nonword stimuli were designed to match the real word stimuli. Words in both repetition tasks were accompanied by visual stimuli. The visual stimuli in the real word repetition task were color photographs of the objects. Visual stimuli in the nonword repetition task were color photographs of unfamiliar tools, plants, fruit, etc. All stimuli images and sound files are available to use for replication in our Open Science Framework project (`https://osf.io/dcb9p/`).

The entire real word repetition task consisted of 94 trials (4 training). The nonword repetition task consisted of 73 trials (6 training). Each real word had a corresponding nonword elicited in the nonword task (e.g. *suitcase*, *sudras*). The current study analyzes a subset of 23 unique real words and 23 nonwords that were repeated across the two tasks (words that didn't begin with CV sequences were excluded) (Table 1). All words were bisyllabic with penultimate stress. The target CV sequence was always an open syllable, in word-initial position, followed by a consonant (CV.C(C)). The target CV sequences were chosen to include an array of consonant manners (approximant, fricative, stop), plosives that differed by place of articulation (velar, alveolar), and sounds differing in articulatory difficulty for children (e.g. [t] versus [ɹ]). Most previous work on child coarticulation has measured coarticulation within CV sequences (e.g. Nittrouer, 1993; Noiray, Popescu, et al., 2019). To facilitate comparison of this work with previous work, the choice was made to measure coarticulation within CV sequences instead of, for example, VC sequences.

The real word repetition task had more trials than the nonword repetition task because the real word task included additional stimuli that are not relevant for this study. The longer real word repetition task could have resulted in additional child fatigue, though we did not observe this during testing. Furthermore, even if this were the case, the difference between the conditions would only serve to reduce the size of the lexicality effect between the nonword and real word conditions, not change the direction of it.

The target CV sequence was always constant between paired real and nonwords (e.g. [su] in *suitcase* and *sudras*). However, the transitional probabilities between the V of the target CV sequence and the first segment of the remainder of the word (e.g. between [u] and [d] in the nonword *sudras*) differed between the nonword and real word conditions. We computed the phonotactic transitional probability between the CV sequence and the first segment of the remaining word using the Hoosier Mental Lexicon Database (Pisoni, Nusbaum, Luce, & Slowiaczek, 1985). The mean phonotactic transition probability for real words was -7.88 versus -8.22 for nonwords, where a value closer to 0 indicates higher phonotactic probability. The median transitional probability was greater for nonwords (-7.72) than real words (-7.84). Consequently, we included Phonotactic Transition Probability as a covariate in our statistical modeling. It did not improve upon a baseline model fit, suggesting that the lexicality effect was independent of the transitional probability between the CV sequence and the first segment of the remaining word; see Results for details. The script used to calculate phonotactic probability is included in the accompanying github repository. See Edwards et al. (2004) and Cychosz et al. (2020) for details about calculation.

Second syllable frequency also differed significantly between the real words and nonwords: real word mean: 2.96, median: 2.94, range: 0-6.4; nonword mean: 1.11, median: 0.69, range:0-3.83 (computed with the Hoosier Mental Lexicon Database (Pisoni et al., 1985)). Second syllable frequency was also added as a covariate during modeling and it did not improve upon a baseline model fit. See Results for further explanation. The second syllables in the real words were more frequent than the second syllables in the nonwords for two reasons. First, it is difficult to concatenate a high-frequency syllable to a CV sequence in English, and generate a phonotactically-probable transition between the syllables, without generating an English word. Second, the real words that determined the target CV sequences were necessarily high-frequency to ensure that children of this age group would recognize them. The use of high-frequency CV sequences further limited the range of possible second syllables in the nonwords.

Two young, female, native speakers, one a speaker of African American English and another

of Mainstream American English, recorded the word stimuli for the tasks. Stimulus recordings were digitized at a sampling rate of 44.1 kHz with a Marantz PMD671 solid-state recorder. To the extent possible, the acoustic signatures of the stimuli were normalized between nonwords and real words. Amplitude was normalized between the conditions, though CV duration was not. The real word stimuli were shorter in duration than the nonword stimuli for both Mainstream American English (real word mean=305.68 ms, SD=73; nonword mean=332.50, SD=87) and African American English (real word mean=245.67, SD=55; nonword mean=264.99, SD=64). The degree of coarticulation within the CV sequences did not differ greatly by lexical status for Mainstream American English (real word mean=29.83, SD=12; nonword mean=29.48, SD=12) or African American English (real word mean=25.77, SD=10, nonword mean=26.23, SD=9); see Cychosz et al. (2020) for details about measurement.

3.4   Repetition task procedure

Each child participant was guided through the word repetition tasks by at least two experimenters. The child was seated in front of a computer screen and presented with a photo while the accompanying word played over external speakers. For the real words, the child was instructed to simply repeat the word. For the nonwords, the child was to repeat the "silly" name of the object as best as possible. Children were encouraged to respond on the first trial. For nonwords, a second trial was allowed if the child did not attempt the word on the first trial. After each trial, the experimenter manually advanced to the next trial. Stimuli were presented in a different random order for each child using E-prime software (Schneider, Eschman, & Zuccolotto, 2012).

Children received the task in their native dialect which was determined by observing mother-child interactions at the beginning of the first study session. N=9 children received the task in African American English and N=94 received the task in mainstream American English. Differences between the African American English and mainstream American English stimuli were in intonation and voice quality, and not at the segmental or morphosyntactic level. Diphthongs

were not monophthongized. We did not observe any relevant dialect differences for the initial CV sequences in children's productions. Dialect was not included as a parameter in our statistical modeling as it was confounded with SES and vocabulary level.

3.5   Data pre-processing

All word productions were segmented into Praat TextGrids (Boersma & Weenink, 2018) and each CV sequence was transcribed by a trained phonetician. Then, the production accuracy of the CV sequence was evaluated. Further details are on the accuracy coding procedure are available in Cychosz et al. (2020), but are summarized briefly here: the place, manner, and voicing (for the consonant) and length, height, and frontness (for the vowel) were evaluated for each CV sequence. Each word's prosodic structure was also evaluated for accuracy (number of syllables, consonant in correct position, and vowel in correct position).

To ensure that we were reliably evaluating repetition accuracy, a second transcriber, also a trained phonetician and native speaker of American English, transcribed a 10% subset of the participants. An intraclass correlation (ICC) statistic measured between-coder agreement. The intraclass correlation between the coders was 0.88, significantly greater than chance (F(374,375)=15.9, p<.001, 95% CI=[0.86, 0.90]) and within an "excellent" range, according to (Cicchetti, 1994).

Only words that were produced entirely correctly, including the prosodic structure, underwent subsequent acoustic analysis. In this way, we could be sure that we were measuring uniformly across the children. For example, measuring the coarticulation between [s] and [ɪ] as in "sister" is different than coarticulation between [t] and [ɪ] for children who pronounce /s/ as [t] in "sister." Word repetition accuracy rates were lowest for stimuli containing late-developing consonants, such as /ɹ/. However, it was still possible to include at least 100 tokens of each CV sequence repetition in the results (Table 2).

Out of a total of 4,738 word repetitions (23 nonwords and 23 real words x 103 participants), we removed 158 (3.3%) because the child's voice in the CV sequence was too breathy to reliably

employ our acoustic measure (measure explained in the following section). This propensity for breathy speech is unsurprising in children of this age (Lee et al., 1999). Of the remaining 4,580 responses, 3,449 (75.3%) were scored as completely accurate meaning that the child produced the correct consonant, vowel, and prosodic structure of the word. The correctly-repeated sequences underwent acoustic analysis. See Table 2 for counts and percentages of items that underwent acoustic analysis, by word type and CV sequence.

For the acoustic analysis, each correctly-produced CV sequence was manually aligned in the previously-generated Praat TextGrid (Boersma & Weenink, 2018) by a native speaker of American English who is a trained phonetician. The audio files were aligned using the visual representation from the waveform and spectrogram in addition to auditory analysis. Coarticulation estimates can be highly sensitive to segmentation decisions so we developed a number of alignment conventions for each CV sequence. For plosive-vowel sequences, such as [kæ], plosive onset corresponded to the burst as it is not possible to mark the closure of word-initial voiceless stops. The start of fricative-vowel sequences, such as [sæ], corresponded to the onset of high-frequency energy in the spectrogram. For both fricative-vowel and stop-vowel sequences, the start of the vowel corresponded to the onset of periodicity and formant structure in the waveform and spectrogram. Delimiting glide-vowel sequences is more gradient: we determined glide offset/vowel onset as a steady state formant. In the very occasional event that we could not identify a steady-state formant in the spectrogram of glide-vowel sequences, we delimited half of the sequence to the glide and half to the vowel. The start of the first consonant in the second syllable of each word marked the target vowel offset. In the case of second syllables that began with plosives, the lack of formant structure and closure marked the end of the vowel. For second syllables with nasal onsets, the presence of anti-formants marked the end of the vowel. And for syllables with fricative onsets, a concentration of high-frequency energy in the spectrogram marked the end of the vowel. Consonant durations in the target CV sequences ranged from 23.40-1159.92 ms and vowel durations ranged from 18.16-574.43 ms.

A second transcriber, who was blind to the study hypothesis concerning real words versus

nonwords, independently aligned a 10% subset of the CV sequences in real words and nonwords. The difference between average consonant duration in the CV sequences aligned by the coders was 16 ms for nonwords and 2 ms for real words. The average difference in vowel duration between coders in the CV sequences was 4 ms for nonwords and 10 ms for real words. In all cases, Pearson correlations between the coders' measurements were significant for all segment types: nonword consonants: r=0.76: p<.001, 95% CI=[0.71, 0.79], real word consonants: r=0.96 p<.001, 95% CI=[0.95, 0.96], nonword vowels: r=0.79 p<.001, 95% CI=[0.76, 0.83], real word vowels: r=0.87 p<.001, 95% CI=[0.85, 0.89]. These statistics suggest high fidelity to the coding conventions.

3.6   Acoustic measures

To calculate the coarticulation between the adjacent C and V phones in the CV sequences, we employed a measure of acoustic distance. Many studies on coarticulation employ acoustic measures of coarticulation such as center of gravity or Peak ERB_N, which are suitable for high-frequency noisy sounds like [s] and [z], or formant frequency-based measurements, which are best suited for vowels. Child formants, however, can be difficult to measure reliably due to the widely-spaced harmonics and breathiness of child speech. Formant-based measures of coarticulation are still valid for children's voices, but they can be less ideal for younger populations (Lee et al., 1999). We employed a single acoustic measurement that is reliable for all manners of articulation: the Euclidean distance between averaged Mel-frequency log-magnitude spectra from adjacent phones, henceforth the Mel spectral distance (Cychosz et al., 2019; Gerosa et al., 2006).[4] For this measurement, a larger spectral distance indicates increased acoustic distance between the phones, meaning less coarticulation (Gerosa et al., 2006). A smaller spectral distance then corresponds to increased spectral similarity—and thus coarticulation—between adjacent speech segments. This measurement has been used previously to measure children's

---

[4] As Lisa Redford pointed out to us, raw frequency scales, such as Hertz, may be better equipped to reflect articulatory relationships between articulation and acoustics in speech while perceptual scales, like Mel, better reflect psychological interpretation of the speech signal. In the current paper, Mel frequency spectra are used as this more closely follows the MFCCs computed in the coarticulation measure originally proposed in Gerosa et al. (2006).

coarticulation (Gerosa et al., 2006) and its performance has been validated (Cychosz et al., 2019). The coarticulation between each C and V phone was calculated automatically using a script running Librosa packages (McFee et al., 2015). Scripts to execute the measure are publicly available (`https://github.com/megseekosh/coartic-experience`).

3.7   Measures of vocabulary size

In addition to the word repetition tasks, participants also completed a series of standard language-related assessments: the Expressive Vocabulary Test, 2nd edition (EVT-2) (Williams, 2007), to quantify expressive vocabulary size, and the Peabody Picture Vocabulary Test, 4th edition (PPVT-4) (Dunn & Dunn, 2007), to quantify receptive vocabulary size. For the expressive vocabulary test, the child was presented with an image and either asked to name the item or provide a synonym for it (depending on the stimulus item). For the receptive vocabulary test, the tester named a word for the child and the child had to choose the corresponding image from an array of four images. Summary statistics for both these tests are presented in the Results.

3.8   Measures of speech practice

To calculate children's naturalistic language use and experience, participating families completed a daylong audio recording in the home. Of the 103 children who completed the experimental tasks, 81 (79%) also completed a daylong audio recording. The remaining families opted not to participate in the daylong recording part of the research program; they were not asked to elaborate upon their decision not to participate.

The daylong recordings were made using the Language ENvironment Analysis (LENA) system (Greenwood, Thiemann-Bourque, Walker, Buzhardt, & Gilkerson, 2011). The LENA system consists of a small (2"x3") digital language processor, similar to a small audio recorder, that a child wears inside of a specialized vest pocket throughout a day. The processor then tracks elements of the child's language environment. This method permits maximally naturalistic observation of the child's language experience, thereby limiting sampling and observational bias. The LENA software program then calculates relevant environmental measures such as the number

of adult words spoken, conversational turns that the target child and adult participate in, and child vocalizations. Families were instructed to turn the recorders on in the morning when the child awoke and record throughout a typical day for the child. Each family completed one recording.

For each recording, we computed the average number of child vocalizations per hour and the average number of adult words spoken per hour. We follow the methods of (Mahr & Edwards, 2018) (2018) for the calculation of these environmental measures and divide the total number of adult words and child vocalizations by the length of the audio recording in hours.

Our exclusionary criteria was as follows: for each recording, we excluded all data recorded after midnight. We also removed one recording that was less than three hours in length as the LENA algorithm performs markedly worse on recordings under 5 hours (Xu, Yapanel, & Gray, 2009) and short recordings may not be representative of the child's entire day. The 80 remaining recordings were at least 9 hours in length (range=9.1-16, mean=14.6, SD=1.5).

4   Results

4.1   Descriptive statistics

Summary statistics for all participants' vocabulary scores and measures from the daylong recording are listed in Table 3. We report both standard scores and growth scale values (linear transformations of raw scores). Both standard scores and growth scale values are linear, so they can be used in statistical analyses. But growth scale values have the advantage of employing a scale that grows linearly with age.

We also report a summary on the relevant measurements from the daylong recording. The average child vocalization count per hour varied greatly. Similarly large ranges have been reported for total (not hourly) child vocalizations in children aged 1;0-1;8 (N=11-5611 vocalizations for 12-hour recordings in Greenwood et al., 2011). For the 12-hour recordings in Greenwood et al. (2011), then, the range of hourly child vocalizations would be 0-467.6, though the authors did not calculate this statistic or hourly child vocalizations. Previous work has also demonstrated large

variability in child vocalization counts in older children aged 3;7 (similar to the child age in the current study): mean=2446, SD=1138 for the 12-hour recordings in Gilkerson and Richards (2008). The average hourly child vocalization count for the 12-hour recordings in Gilkerson and Richards (2008) would be approximately 203.8 (SD=95). (The authors only reported the mean and standard deviation of child vocalization and adult word counts, not ranges.)

The caregivers of 13 children identified their children as late talkers (see Methods). On average, the caregiver-identified late-talkers produced fewer child vocalizations per hour (220.3 versus 246.6) (Table 4). The caregiver-identified children also had lower average standard scores for receptive vocabulary (PPVT-4) (110 versus 123) and expressive vocabulary (EVT-2) (110 versus 121), although it should be noted that mean standard scores for both vocabulary measures were above the standardized mean of 100 for both groups of children . As a result of these score discrepancies between groups, we included late talker status as a covariate in our statistical model fitting procedure

4.2   Vocabulary knowledge and coarticulation

We first evaluated our hypotheses concerning the roles of speech planning and vocabulary size for coarticulation. Here we hypothesized that children would coarticulate less in real words than nonwords and that children with larger vocabularies were coarticulate less (Noiray, Popescu, et al., 2019). A mixed effects linear regression model was fit to predict the Mel spectral distance between each CV sequence. Recall that in these models, a larger outcome variable—the Mel spectral distance—indicates less gestural overlap, and thus less coarticulation between the segments. Model fitting was conducted using `lme4` (Bates, Maechler, Bolker, & Walker, 2015) and `lmerTest` (Kuznetsova, Brockhoff, & Christensen, 2017) in the R studio programming environment (R Core Team, 2020, version 1.3.1056). Model summaries were made with `Stargazer` (Hlavac, 2018). All continuous variables were mean-centered to facilitate model interpretation. The decision to include a predictor in a model was determined by a log-likelihood test comparing model fits with and without the predictor. A significant log-likelihood test indicated that the model with the

predictor improved upon the model without. Given the highly similar results for models with Expressive Vocabulary and Receptive Vocabulary, but the theoretical interest in distinguishing between the two measure of vocabulary, best model fit for the vocabulary measures was instead determined by the model with the lowest AIC/BIC value; this is explained in further detail below. All modeling and analysis scripts are publicly available in the accompanying GitHub repository (`https://github.com/megseekosh/coartic-experience`).

The baseline model included random effects of Word and Participant. Each participant contributed one repetition of each item which does not permit random slopes of Participant by Word. We followed a forward model fitting testing procedure with the following parameters tested: Word Duration (to control for speaking rate differences between real words and nonwords), Phonotactic Transitional Probability (between the CV sequence and the first consonant of the remainder of the word), Second Syllable Frequency (e.g. [dras] from the item *sudras*), Word Type (real word versus nonword), Late Talker (yes or no), Maternal Education Level, Gender (reported by caregiver), Receptive Vocabulary Size (PPVT-4 growth scale value), and Expressive Vocabulary Size (EVT-2 growth scale value).

The predictor Word Duration improved upon a baseline model containing only the random effects. Next, we added Word Type (real word versus nonword) to the model; it did not improve upon model fit. We additionally checked the roles of Phonotactic Transitional Probability between the target CV sequence and the first consonant of the remainder of the word and Second Syllable Frequency; neither variable improved upon model fits with or without the predictor Word Type. This result suggests that children coarticulate similarly between CV sequences in real words and nonwords after controlling for word duration.

Including Word Duration as a parameter in the modeling is important since the children produced the real words faster (mean=874.61 ms, SD=279.50), on average, than the nonwords (mean=1101.10, SD=269.77). Moreover, coarticulation tends to increase in faster speech, at least for adults. Controlling for Word Duration by word type in our modeling may therefore have masked the (much smaller) effect of Word Type upon coarticulation. We return to this point in

the Discussion.

Next, we tested Maternal Education Level and Gender as these factors have been shown to predict children's speech and language outcomes. Neither Maternal Education Level nor Gender improved model fit so we excluded the variables from further analysis. Late Talker status also did not improve model fit and was excluded.

Best model fit included main effects of Word Duration and Receptive Vocabulary Size (Figure **??**; Table 5). Expressive Vocabulary Size also improved model fit, but Receptive Vocabulary Size resulted in a marginally better model fit. The improvement of Receptive Vocabulary Size over Expressive Vocabulary Size was so slight—the AIC estimations for the models was the same and the BIC was slightly smaller for the model with Receptive Vocabulary Size—that the difference between the different vocabulary estimates for coarticulation in this study is almost inconsequential. The significant effect of Receptive Vocabulary Size suggests that children with larger vocabularies tend to coarticulate less between phones (β=.02, t=2.65, p=.01).

4.3   Spoken language experience and coarticulation

After evaluating the roles of speech planning and vocabulary size for coarticulation, we turned to how speech practice may predict the children's speech patterns and, more specifically, how speech practice could interact with speech planning and vocabulary. We predicted that the frequency of children's vocalizations, quantified in daylong audio recordings, would mediate their coarticulation patterns: children who vocalize more often in daylong recordings will coarticulate less within CV sequences. To evaluate this hypothesis, an additional mixed effects linear regression model was fit to predict the Mel spectral distance between phones in each CV sequence. As before, a larger spectral distance indicates less gestural overlap—less coarticulation—between the segments.

To address the role of speech practice on coarticulation, we included measures of the children's daily language experiences from the 80 of the 103 children (77.7%) who completed a daylong audio recording that was at least 9 hours long. The baseline model included the random

effects of Word and Participant. We then repeated the forward model fitting procedure as before
with the following parameters added: Word Duration, Phonotactic Transitional Probability,
Second Syllable Frequency, Word Type, Late Talker, Maternal Education Level, Gender, Receptive
Vocabulary Size (PPVT-4 growth scale value), Expressive Vocabulary Size (EVT-2 growth scale
value), and Child Vocalization Count (the average number of hourly child vocalizations from the
daylong recording). Predictors that did not improve model fit were removed from the analysis.

As in the model fit to the full dataset, Phonotactic Transitional Probability, Second Syllable
Frequency, and Word Type did not improve upon a baseline model fit with just the random effects
and Word Duration. We likewise did not find significant model improvements after adding
Maternal Education Level, Gender, or Late Talker so we removed those factors from the model.

Table 6 shows the final model fit with significant main effects of Word Duration and
Expressive Vocabulary Size (β=.02, t=2.57, p=0.01).[5] Child Vocalization Count was also
significant, indicating that children who tended to vocalize more throughout the daylong recording
coarticulated less within CV sequences (β=.003, t=2.74, p=0.01), whether the sequences were
embedded in nonwords or real words (Figure 2). This result provides some support for a "practice
effect" for children's coarticulation patterns.

Child Vocalization Count and Receptive Vocabulary Size were not significantly correlated
(Pearson r=.02, t=.16, p=.88). Nor were Child Vocalization Count and Expressive Vocabulary
Size (r=-0.05, t=-0.47, p=.64). This lack of correlation between these variables indicates that
children's language skills and vocabulary sizes do not simply increase as they talk/vocalize more
frequently, a point which is further addressed in the Discussion.

To further explore the patterns between vocabulary, child vocalizations, and coarticulation,
we performed a median split on the 80 children who had completed a daylong recording, dividing

———

[5] In the dataset containing children who completed a daylong recording, Expressive Vocabulary Size proved a
marginally better fit than Receptive Vocabulary Size. This fact further underscores the lack of distinction between
Expressive and Receptive Vocabulary Sizes for coarticulation: we do not find strong evidence for one vocabulary
measure over the other.

them into smaller (n=52 children) and larger expressive vocabulary groups (n=28 children). Dividing the children into these groups allowed us to see that the effect of Child Vocalization Count appears to be especially prominent for the children in the smaller vocabulary group (left panel of Figure 2). For the children with larger vocabularies, Child Vocalization Count does not appear to be strongly correlated with coarticulation. Despite these differences by vocabulary size, the interaction of Child Vocalization Count and Expressive Vocabulary was not significant. Finally, while the relationship between Child Vocalization Count and coarticulation may have been weaker in the larger vocabulary group the effect of word type—though not significant in the modeling—appears to be stronger in the larger vocabulary group. Children with larger vocabularies tend to coarticulate just slightly more in nonwords than real words. This is, however, just a trend in the data as the interaction of Word Type and Expressive/Receptive Vocabulary Sizes was not significant.

4.4 Language input and coarticulation

Although we found that children who vocalized more in the daylong recordings coarticulated less within the CV sequences, these results could also be explained by parent-child engagement. Specifically, children who hear more input from caregivers may, in turn, speak more so there could be an indirect role of input upon children's coarticulation. In our sample, we did find that the average number of hourly child vocalizations was positively correlated with the average number of hourly adult words (r=.44, t=4.35, p<.001; Figure 3). This positive correlation between the two measures led us to hypothesize that the quantity of adult speech in the child's environment could predict the child's coarticulation patterns, potentially having a stronger role than the child's own productions.

We tested this exploratory hypothesis in the same model of the 80 children who completed a daylong recording. Now instead of adding Child Vocalization Count, we instead added Adult Word Count, or the average number of words spoken by adults per hour in the daylong recording. Adult Word Count did not improve model fit, either with a model that included Child

Vocalization Count (χ^2=1.99, df=1, p=.16) or one that did not (χ^2=.01, df=1, p=.92). Therefore, we can confidently conclude that Adult Word Count does not directly predict children's coarticulation within the CV sequences.

While we did not find a direct effect of Adult Word Count, the relationship between Adult Word Count and Child Vocalization Count, and in turn the relationship between Child Vocalization Count and degree of coarticulation, suggests an indirect role of Adult Word Count on children's coarticulation: more adult words spoken in the environment leads to a greater number of child vocalizations, which, in turn, predicts the degree of children's coarticulation. This relationship between the quantity of ambient adult language and coarticulation is indirect: Adult Word Count is not related to the children's coarticulatory outcomes without factoring in Child Vocalization Count.

4.5   Coarticulation by consonant manner

The acoustic measure employed here is validated for all consonant manners. However, we additionally analyzed the potential effect of consonant manner on the relationship between coarticulation, vocabulary size, and child vocalizations. In all cases, analysis was limited to glides, fricatives, and stops because there was only 1 stimulus pair with an affricate-vowel sequence. Broadly, the results by consonant manner replicated those conducted over all segments together: there was little to no effect of word type (Figure 4 in Appendix A) and children with larger vocabularies tended to coarticulate less. The vocabulary effect was most pronounced in fricative-vowel and glide-vowel sequences. Children who vocalized more also tended to coarticulate less between CV sequences, independent of consonant manner (Figure 5 in Appendix A).

5   Discussion

5.1   Coarticulation in real words and nonwords

Previous work has primarily attributed young children's propensity to coarticulate to two factors: children's immature, developing speech motor control (Barbier et al., 2020; Rubertus &

Noiray, 2018; Zharkova et al., 2018) and the size of their representational units (Nittrouer et al., 1989, 1996; Noiray, Popescu, et al., 2019; Zharkova et al., 2011). Past work has used cross-sectional experimental designs to distinguish between motoric and representational explanations for coarticulation. But children's motor skills are tightly tied to their chronological age. This study instead evaluated how several factors—speech planning, vocabulary size, and speech practice—could predict coarticulation in a large cohort of four-year-old children. Crucially, each of these factors develops more independently of speech motor control than children's age. Thus, evaluating the relationships between these factors and the degree of children's coarticulation allows us to determine some additional, underlying causes of coarticulation, beyond children's developing motor control abilities. In the following sections, we discuss our findings concerning each of the three factors and their relationship with four-year-old children's coarticulation patterns.

To evaluate the role of speech planning, the degree of the children's coarticulation was measured in two environments: real words and corresponding nonwords (e.g. [su] in *suitcase* and *sudras*). We hypothesized that children would coarticulate less between CV sequences embedded in real words than nonwords because children have more experience accessing and articulating the real words making nonword repetition a more demanding task. For example, while the CV sequence was matched across each real word and nonword pair, nonword production requires repetition without any phonetic, lexical, or semantic support. Speakers must encode and produce a novel combination of syllables without reinforcement. In real word repetition, on the other hand, speakers encode a semantic and phono-lexical representation and then execute the associated articulatory schema. Nonword repetition is also often considered a proxy for phonological representations in working memory (e.g. Edwards & Lahey, 1998; Gathercole, 2006): children with larger vocabularies perform better on the task (Edwards et al., 2004; Munson et al., 2005) and speakers of all ages perform worse on longer, multisyllabic words (Byrd, Vallely, Anderson, & Sussman, 2012).[6] The task demands of nonword repetition, in addition to the fact

———————

[6] Nonword repetition is a proxy for phonological working memory, but there are myriad factors, including those

that children cannot rely upon established motor schemata to articulate the nonwords, could then affect the children's coarticulation by not permitting the same degree of planning permitted for real word repetition.

Our analyses did not support the prediction that children would coarticulate more in nonwords: there was no effect of word type on coarticulation after controlling for word duration. Children with larger expressive vocabularies did tend to coarticulate less between CV sequences in real words than nonwords, but this did not approach significance in the modeling. There are several ways to interpret this result. First, there could simply be no effect of speech planning on children's coarticulation. Second, there could be an effect of speech planning but it is only apparent under certain experimental manipulations and lexical status is not one of them. Finally, there could be an effect of speech planning on coarticulation as implemented in this experimental design, but the much stronger effect of speaking rate on coarticulation may have masked the weaker, correlated effect of word type. We address this possibility in the remainder of this section.

Word duration was controlled for in the modeling both because the nonwords tended to be longer in duration than the real words and also because, more generally, coarticulation and speaking rate are tightly linked. Adult speakers coarticulate more in faster speech (Gay, 1981; Matthies, Perrier, Perkell, & Zandipour, 2001). The relationship between coarticulation and speaking rate may be reversed in children whose coarticulation reflects different pressures. In adults, coarticulation is usually planned (Whalen, 1990) and reflects an equilibrium between speaker efficiency and listener comprehension (Bradlow, 2002). But, as stated throughout this work, child coarticulation is instead generally attributed to a lack of fine motor control and entrenched articulatory schema that may be required to differentiate between adjacent phones during speech (Goffman et al., 2007; Green et al., 2000; McAllister Byun & Tessier, 2016) or phonological reorganization (Nittrouer et al., 1989, 1996; Noiray, Popescu, et al., 2019; Redford, 2018; Zharkova et al., 2011). The result is that children might coarticulate more in *slower* speech;

—————

pertaining to the stimuli like phoneme frequency and sub-syllabic frequency, that predict task performance (see Szewczyk, Marecka, Chiat, and Wodniecka (2018) for comprehensive overview).

but their coarticulation and speaking rate are, nevertheless, intertwined.

Because of this relationship between speaking rate and coarticulation, it was essential to include a proxy for speaking rate (word duration) in the modeling. However, word duration, which covaries tightly with word type, might have masked any effect of word type. There are two reasons why word duration would have this effect. First, the effect of speaking rate on coarticulation is quite robust (Gay, 1981; Matthies et al., 2001), regardless of the direction of the effect by age. Second, any effect of word type on degree of coarticulation is likely to be quite small. Even employing a fine-grained measure of coarticulation such as that used here might not be capable of capturing differences due to word type that are independent of the strong link between word duration and coarticulation.

We summarize the results of speech planning on coarticulation by saying that there is no reliable effect of word type upon coarticulation that is independent of speaking rate. An effect of word type would have provided evidence of extra-motoric influences on coarticulation, such as those stemming from phonological reorganization. In the following sections we interpret the findings from the other two variables studied, vocabulary size and speech practice, which do suggest an effect of phonological reorganization upon children's coarticulation.

5.2   Language experience and coarticulation

The role of vocabulary size on coarticulation.   Our second hypothesis concerned the role of vocabulary size on child coarticulation patterns in the real words and nonwords. Using our full sample size of 103 children, we found that children with larger receptive vocabularies coarticulated less between the segments in CV sequences. Receptive vocabulary size was only marginally more predictive of coarticulation degree than expressive vocabulary size for the full dataset. The difference between the models with the different vocabulary measures was marginal. Furthermore, expressive and receptive vocabulary scores are correlated, and in the sub-analysis of the 80 children who completed a daylong recording, we instead found that expressive vocabulary better predicted coarticulation outcomes. As a result, we do not make strong claims concerning

the role of receptive versus expressive vocabulary for coarticulation: both estimates predict coarticulation to a relatively similar degree.

Given the results of Noiray, Popescu, et al. (2019), the finding in this paper that children with larger vocabularies coarticulate less is unsurprising. Noiray, Popescu, et al. (2019) evaluated a relationship between vocabulary size and degree of coarticulation in children aged 4;0-7;0 and found that children with larger expressive vocabularies showed less anticipatory coarticulation in [stop]-V sequences. The results from the current paper replicate this finding. Noiray, Popescu, et al. (2019) attributed the relationship between vocabulary size and coarticulation to the primacy of the lexicon for phonetic and phonological development as phonological representations emerge from generalizations that children make over their vocabularies. This theoretical approach argues that as the lexicon grows, phonological representations reorganize from larger sub-lexical units like syllables and feet into smaller, phoneme-sized ones (Beckman & Edwards, 2010; Edwards et al., 2004; Metsala, 1999; Metsala & Walley, 1998; Storkel & Hoover, 2011; Storkel & Morrisette, 2002). Children with larger vocabularies are then expected to have more abstract speech segments. Phonological reorganization manifests in production accuracy when children with larger vocabularies perform better on nonword repetition (3;2-8;10: Edwards et al., 2004; 3;0-6;0: Munson et al., 2005) and novel word learning tasks (2;11-6;0: Storkel & Hoover, 2011). The phonological abstraction stemming from the growth of the lexicon could also play out in children's phonetic development as children with larger vocabularies coarticulate less. Still, the relationship between receptive vocabulary size and coarticulation observed here is not as robust as the relationship between vocabulary size and those other outcome measures (repetition accuracy, novel word learning) (Edwards et al., 2004; Metsala, 1999; Storkel & Hoover, 2011). The different result found here could be because there are myriad other factors, beyond vocabulary size, that predict children's spoken phonetic outcomes, and coarticulation in particular.

Throughout this work, we have entertained two primary explanations for children's coarticulation patterns: speech motor control and phonological reorganization. The finding that children with larger vocabularies coarticulate less supports the idea that coarticulation reflects

phonological reorganization from larger sub-lexical units into phonemes because previous work has concluded that children with larger vocabularies have more segmental phonological representations (e.g. Metsala & Walley, 1998). However, the weaker relationship between coarticulation and vocabulary size, compared to previous work evaluating the relationship between word learning or repetition accuracy and vocabulary size, suggests that there is also a strong role of another factor for children's coarticulation—fine motor control—that may not explain as much variability in novel word learning or nonword repetition accuracy. Vocabulary size may predict coarticulation, but the effect is weaker because fine motor control exerts a strong influence on phonetic outcomes (Barbier et al., 2020) in a way that it does not for word learning and other lexical tasks.

The role of production practice on coarticulation.   The final potential predictor of children's coarticulation that we evaluated in this study was the frequency of children's vocalizations. The role of children's vocalization frequency was examined in the subset of children who completed a daylong audio recording (80/103 or 77.7%) and we found that children who vocalized more during the daylong recording coarticulated less, both in real words and nonwords. This study was the first to examine this "practice effect" in older children, aged 4;0, using naturalistic recordings of the children's everyday environments. Our results corroborate previous evidence in support of a practice effect for speech development (DePaolis et al., 2011, 2013; Icht & Mama, 2015; Keren-Portnoy et al., 2010; Majorano et al., 2014; Redford, 2014 cf. Zamuner et al., 2018) whereby children who practice more spoken language, or articulate words during word learning, could learn to encode articulatory movements with their acoustic representations. This articulatory practice, in turn, may further entrench phonological representations and may make the task of accessing a given word or phone easier, and smoother, for future productions. The results found here also support previous findings on a different measure of children's speech practice—articulation rate. That work has demonstrated how adults perceive children who speak faster to have clearer speech, again suggesting that children who talk more have more mature

speech production outcomes (Redford, 2014).[7]

The practice effect found here is potentially even more interesting since it works independently from vocabulary size: we did not find a positive relationship between children's vocalization frequency and their vocabulary size. The lack of a relationship between vocalization frequency and vocabulary size suggests that speech practice predicts speech development without the mediating factor of vocabulary size.

It should be noted that our estimate of hourly child vocalizations was relatively coarse. While the child vocalization count estimator of the LENA speech parsing algorithm performs reliably well (Cristia et al., 2020) (better than other more language- and context-specific estimations from LENA), the algorithm does not distinguish between speech-like vocalizations and crying or laughing. Consequently, this study is an important first step in establishing the role of self-practice for children's coarticulation and speech development. Future work now needs to evaluate children's naturalistic speech productions in this age range in a more fine-grained manner to better understand what characteristics of children's vocalizations predict their speech outcomes. For example, while we did not find that an estimate of receptive language experience, adult word count, directly predicted the degree of coarticulation, adult word count and child vocalization count were positively correlated. Thus, it could be that interactions with adult interlocutors in particular indirectly predict coarticulatory development. Delving deeper into the role of self-practice and the language learning environment is thus an important area for future research on phonetic development.

This study only measured coarticulation between CV sequences in words that the children repeated correctly. Words in which the children made phonological substitutions or deleted

---

[7] Given the results of Redford (2014), a supplementary analysis was performed to evaluate the relationship between articulation rate (duration of repeated words) and Child Vocalization Count for the children in the current study who completed a daylong recording. Results showed a significant, but weak correlation between the duration of the repeated real words and Child Vocalization Count (r= -.07, t= -2.72, p=0.007) suggesting that children who speak more throughout the day also speak faster, at least in a lab-based word repetition task (there was no relationship between the duration of nonwords and Child Vocalization Count).

syllables were removed from analysis. The decision to remove incorrect words was made for practical reasons: there was no way to control the errors that children made as they repeated the words, which made it difficult to perform an individual differences analysis. However, children probably do coarticulate differently in words that they repeat incorrectly. For example, deletion of a syllable during word repetition could indicate that greater planning demands were placed on the child, meaning they might have been more likely to coarticulate. For substitution errors, children who regularly substitute [w] for /r/ might coarticulate more in [we] sequences in words like "raisin" than they do in phonemically /we/ sequences in words like "wait" or "waste." Consequently, measuring coarticulation in correctly- and incorrectly-repeated words seems to be a potentially informative line of future research on speech development which we encourage on the basis of the results from correctly-repeated words in the current study.

5.3 On data exclusion and daylong recordings

As daylong recording methodologies increase in popularity in developmental language research, ethical issues surrounding participant exclusion will become more apparent (Casillas & Cristia, 2019). In this study, 22 families opted not to participate in the daylong recording portion of the study. We still included the word repetition and vocabulary data for these children rather than excluding them completely from the study.

Daylong recordings require that children wear recorders over extended periods. Sound within an approximately five foot radius of the child is captured in the recording. Some families may, understandably, not be comfortable with this methodology. One solution for studies like ours that incorporate covariates derived from daylong recordings could be to exclude all children whose families opt out of the daylong recording. However, we do not encourage this option. The decision to contribute a daylong audio recording to a research study may depend on several external factors like socioeconomic status, race, and familial structure (i.e. single-parent households). For example, in a hypothetical example that relates to the current study, some caregivers may have unpredictable work schedules making it difficult to complete a 12 hour recording. Given these

challenges, it would hardly be surprising if some individuals from under-served communities were more likely to opt out of daylong recording methodologies. In fact, in our sample, we found a different SES distribution in families who completed the daylong recording (92.59% of mothers had more than a high school diploma) versus those who did not (86.36% of mothers had more than a high school diploma).[8] Thus, excluding all families who could not complete a daylong recording could inadvertently exclude some groups from representation in social science research. The exclusion of these families would not be random and could, inadvertently, serve to further bias language development findings for upper middle class North American children.

5.4   The status of children's phonological representations

An overarching goal of this study was to contribute to our understanding of children's phonological representations. By showing novel evidence for the roles of vocabulary size and speech practice—but not speech planning—on coarticulation, this study corroborates Noiray, Popescu, et al. (2019) in suggesting that coarticulation is dually explained by speech motor control maturation and phonological reorganization. In particular, the predictive relationship of vocabulary size for coarticulation found in this study suggests that children with larger lexicons have more segmental phonological units. This finding supports phonological reorganization.

The ramifications for the role of child vocalization frequency may be less straightforward because algorithmically-derived speech measures, such as LENA's child vocalization count or adult word count, do not necessarily map onto linguistically-coded categories. Children who vocalized more often during the daylong recording may exhibit motor practice effects, resulting in more entrenched or even abstracted phonological categories. Of course children who vocalize more could

---

[8] The SES breakdown of the two groups was as follows: for families who did complete a daylong recording n=1 did not have a high school diploma, n=1 had the equivalent of a high school diploma (e.g. GED), n=4 had a high school diploma, n=0 had <2 years of college, n=15 had 2+ years of college, n=27 had completed college, and n=33 had graduate degrees . For families who did not complete a daylong recording, n=1 did not have a high diploma, n=0 had the equivalent of a high school diploma, n=2 had a high school diploma, n=1 had <2 years of college, n=6 had 2+ years of college, n=6 had completed college, and n=6 had graduate degrees.

also have greater motor control because they speak more frequently. The hourly child vocalization count derived from the LENA recordings was more predictive than vocabulary size (receptive or expressive) for coarticulation in the participants where we could test both vocabulary size and child vocalization count. On the basis of the comparison of these two factors, it appears that the frequency of speech practice is the strongest predictor of children's coarticulation patterns.

6   Conclusion

This study measured four-year-old children's coarticulation between phones in CV sequences. Previous work on this topic has struggled to disentangle the roles of fine motor control and the potential role of children's larger representational units for children's coarticulation patterns. Three factors that develop more independently of fine motor control than children's age—speech planning, vocabulary size, and speech practice—were evaluated for their role on coarticulation. We did not find an effect of speech planning on coarticulation: children coarticulated similarly in real words and matched nonwords after controlling for speaking rate. We did, however, find effects of vocabulary size and speech practice. Children with larger vocabularies coarticulated less, as did those who vocalized more during a daylong audio recording. This daily practice effect was more predictive of coarticulatory outcomes than the quantity of adult language in the ambient environment and the child's vocabulary size, suggesting the strong role of self-practice for speech development. Overall, the observed effects of vocabulary size and speech practice suggest that both fine motor control and children's representational units independently explain the development of coarticulation. (Word count: 11955)

7   Acknowledgements

8   Declaration of interest

The authors have no conflicts of interest to report.

9   References

Albert, R. R., Schwade, J. A., & Goldstein, M. H. (2018, September). The social functions of babbling: Acoustic and contextual characteristics that facilitate maternal responsiveness. *Developmental Science*, *21*(5), e12641. doi: 10.1111/desc.12641

Barbier, G., Perrier, P., Payan, Y., Tiede, M. K., Gerber, S., Perkell, J. S., & Ménard, L. (2020). What anticipatory coarticulation in children tells us about speech motor control maturity. *PLoS ONE*, *15*(4), e0231484. doi: 10.1371/journal.pone.0231484

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

Beckman, M. E., & Edwards, J. (2010, January). Generalizing over lexicons to predict consonant mastery. *Laboratory Phonology*, *1*(2). doi: 10.1515/labphon.2010.017

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.

Bruderer, A. G., Danielson, D. K., Kandhadai, P., & Werker, J. F. (2015, November). Sensorimotor influences on speech perception in infancy. *Proceedings of the National Academy of Sciences*, *112*(44), 13531–13536. doi: 10.1073/pnas.1508631112

Byrd, C. T., Vallely, M., Anderson, J. D., & Sussman, H. (2012, September). Nonword repetition and phoneme elision in adults who do and do not stutter. *Journal of Fluency Disorders*, *37*(3), 188–201. doi: 10.1016/j.jfludis.2012.03.003

Casillas, M., & Cristia, A. (2019, May). A step-by-step guide to collecting and analyzing long-format speech environment (LFSE) recordings. *Collabra: Psychology*, *5*(1), 24. doi: 10.1525/collabra.209

Chiat, S., & Roy, P. (2007, April). The Preschool Repetition Test: An Evaluation of Performance in Typically Developing and Clinically Referred Children. *Journal of Speech, Language, and Hearing Research*, *50*(2), 429–443. doi: 10.1044/1092-4388(2007/030)

Cicchetti, D. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment*, *6*, 284–290.

Cristià, A. (2011, May). Fine-grained variation in caregivers' /s/ predicts their infants' /s/ category. *The Journal of the Acoustical Society of America*, *129*(5), 3271–3280. doi: 10.1121/1.3562562

Cristia, A., Lavechin, M., Scaff, C., Soderstrom, M., Rowland, C., Räsänen, O., . . . Bergelson, E. (2020). A thorough evaluation of the Language Environment Analysis (LENA) system. *Behavior Research Methods*.

Cychosz, M. (2020). *Phonetic development in an agglutinating language* (Unpublished doctoral dissertation ed.). Berkeley, CA: University of California.

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

Cychosz, M., Erskine, M., Munson, B., & Edwards, J. R. (2020). A lexical advantage in four-year-old children's word repetition. *Journal of Child Language*.

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

DePaolis, R. A., Vihman, M. M., & Keren-Portnoy, T. (2011, December). Do production patterns influence the processing of speech in prelinguistic infants? *Infant Behavior and Development*, *34*(4), 590–601. doi: 10.1016/j.infbeh.2011.06.005

DePaolis, R. A., Vihman, M. M., & Nakai, S. (2013, December). The influence of babbling patterns on the processing of speech. *Infant Behavior and Development*, *36*(4), 642–649. doi: 10.1016/j.infbeh.2013.06.007

Dunn, L. M., & Dunn, D. M. (2007). *PPVT-4: Peabody picture vocabulary test.* Pearson Assessments.

Edwards, J., Beckman, M. E., & Munson, B. (2004). The interaction between vocabulary size and

phonotactic probability effects on children's production accuracy and fluency in novel word repetition. *Journal of Speech Language and Hearing Research*, *57*, 421–436.

Edwards, J., & Lahey, M. (1998, April). Nonword repetitions of children with specific language impairment: Exploration of some explanations for their inaccuracies. *Applied Psycholinguistics*, *19*(2), 279–309. doi: 10.1017/S0142716400010079

Engel de Abreu, P. M., Conway, A. R., & Gathercole, S. E. (2010, November). Working memory and fluid intelligence in young children. *Intelligence*, *38*(6), 552–561. doi: 10.1016/j.intell.2010.07.003

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.

Ferguson, C., & Farwell, C. (1975). Words and sounds in early language acquisition. *Language*, *51*(2), 419–439.

Ferjan Ramírez, N., Lytle, S. R., Fish, M., & Kuhl, P. K. (2019, May). Parent coaching at 6 and 10 months improves language outcomes at 14 months: A randomized controlled trial. *Developmental Science*, *22*(3), e12762. doi: 10.1111/desc.12762

Franich, K. (n.d.). The effect of cognitive load on tonal coarticulation. In *Proceedings of the 18th International Congress of Phonetic Sciences.*

Franklin, B., Warlaumont, A. S., Messinger, D., Bene, E., Nathani Iyer, S., Lee, C.-C., . . . Oller, D. K. (2014, July). Effects of Parental Interaction on Infant Vocalization Rate, Variability and Vocal Type. *Language Learning and Development*, *10*(3), 279–296. doi: 10.1080/15475441.2013.849176

Gathercole, S. E. (2006, October). Nonword repetition and word learning: The nature of the relationship. *Applied Psycholinguistics*, *27*(4), 513–543. doi: 10.1017/S0142716406060383

Gathercole, S. E., Willis, C., Emslie, H., & Baddeley, A. (1991). The influences of number of syllables and wordlikeness on children's repetition of nonwords. *Applied Psycholinguistics*, *12*(3), 349–367.

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

Gilkerson, J., & Richards, J. (2009). *Impact of Adult Talk, Conversational Turns, and TV During the Critical 0-4 Years of Child Development* (2nd Edition ed.) (No. ITR-01-2).

Gilkerson, J., & Richards, J. A. (2008). The LENA Natural Language Study. , 26.

Goffman, L., Gerken, L., & Lucchesi, J. (2007, April). Relations Between Segmental and Motor Variability in Prosodically Complex Nonword Sequences. *Journal of Speech, Language, and Hearing Research*, *50*(2), 444–458. doi: 10.1044/1092-4388(2007/031)

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)

Goldstein, M. H., & Schwade, J. A. (2008, May). Social Feedback to Infants' Babbling Facilitates Rapid Phonological Learning. *Psychological Science*, *19*(5), 515–523. doi: 10.1111/j.1467-9280.2008.02117.x

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.

Green, J. R., Moore, C. A., Higashikawa, M., & Steeve, R. W. (2000). The Physiologic Development of Speech Motor Control: Lip and Jaw Coordination. *Journal of Speech Language and Hearing Research*, *43*(1), 239–255.

Greenwood, C. R., Thiemann-Bourque, K., Walker, D., Buzhardt, J., & Gilkerson, J. (2011, February). Assessing Children's Home Language Environments Using Automatic Speech Recognition Technology. *Communication Disorders Quarterly*, *32*(2), 83–92. doi: 10.1177/1525740110367826

Hart, B., & Risley, T. (1995). *Meaningful differences in the everyday experience of young American children.* Baltimore, MD: Paul H. Brookes Publishing.

Hitch, G. J., & Halliday, M. S. (1983). Working memory in children. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *302*(1110), 325–340.

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

Huttunen, K. H., Keränen, H. I., Pääkkönen, R. J., Päivikki Eskelinen-Rönkä, R., & Leino, T. K. (2011). Effect of cognitive load on articulation rate and formant frequencies during simulator flights. *The Journal of the Acoustical Society of America*, *129*(3), 1580–1593. doi: 10.1121/1.3543948

Icht, M., & Mama, Y. (2015, September). The production effect in memory: A prominent mnemonic in children. *Journal of Child Language*, *42*(5), 1102–1124. doi: 10.1017/S0305000914000713

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.

Kent, R. (1976). Anatomical and neuromuscular maturation of the speech mechanism: Evidence from acoustic studies. *Journal of Speech Language and Hearing Research*, *19*, 421–447.

Kent, R. (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$_4$

Keren-Portnoy, T., Vihman, M. M., DePaolis, R. A., Whitaker, C. J., & Williams, N. M. (2010, October). The Role of Vocal Practice in Constructing Phonological Working Memory. *Journal of Speech Language and Hearing Research*, *53*(5), 1280. doi:

10.1044/1092-4388(2009/09-0003)

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

Laing, C. E., & Bergelson, E. (2020). From babble to words: Infants' early productions match words and objects in their environment. *Cognitive Psychology*, *122*, e101308.

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

Lively, S. E., Pisoni, D. B., Summers, W. V., & Bernacki, R. H. (1993). Effects of cognitive workload on speech production: Acoustic analyses and perceptual consequences. *The Journal of the Acoustical Society of America*, *93*(5), 2962–2973.

Mahr, T., & Edwards, J. (2018, November). Using language input and lexical processing to predict vocabulary size. *Developmental Science*, *21*(6), 1–14. doi: 10.1111/desc.12685

Majorano, M., Vihman, M. M., & DePaolis, R. A. (2014, April). The Relationship Between Infants' Production Experience and Their Processing of Speech. *Language Learning and Development*, *10*(2), 179–204. doi: 10.1080/15475441.2013.829740

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.

McAllister Byun, T., Inkelas, S., & Rose, Y. (2016). The A-map model: Articulatory reliability in child-specific phonology. *Language*, *92*(1), 141–178. doi: 10.1353/lan.2016.0000

McAllister Byun, T., & Tessier, A.-M. (2016, September). Motor Influences on Grammar in an Emergentist Model of Phonology: Motor and Grammar in Emergentist Phonology. *Language and Linguistics Compass*, *10*(9), 431–452. doi: 10.1111/lnc3.12205

McBride-Chang, C., Bialystok, E., Chong, K. K., & Li, Y. (2004, October). Levels of phonological awareness in three cultures. *Journal of Experimental Child Psychology*, *89*(2), 93–111. doi: 10.1016/j.jecp.2004.05.001

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

McGillion, M., Herbert, J. S., Pine, J., Vihman, M., dePaolis, R., Keren-Portnoy, T., & Matthews, D. (2017, January). What Paves the Way to Conventional Language? The Predictive Value of Babble, Pointing, and Socioeconomic Status. *Child Development*, *88*(1), 156–166. doi: 10.1111/cdev.12671

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).

Metsala, J. (1999). Young children's phonological awareness and nonword repetition as a function of vocabulary development. *Journal of Educational Psychology*, *91*(1), 3–19.

Metsala, J., & Walley, A. (1998). Spoken vocabulary growth and the segmental restructuring of lexical representations: Precurors to phonemic awareness and early reading ability. In *Word recognition in beginning literacy* (pp. 89–120). Lawrence Erlbaum Associates.

Munson, B., Edwards, J., & Beckman, M. E. (2005, February). Relationships Between Nonword Repetition Accuracy and Other Measures of Linguistic Development in Children With Phonological Disorders. *Journal of Speech, Language, and Hearing Research*, *48*(1), 61–78. doi: 10.1044/1092-4388(2005/006)

Nittrouer, S. (1993, October). The emergence of mature gestural patterns is not uniform: Evidence from an acoustic study. *Journal of Speech Language and Hearing Research*, *36*, 959–972.

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$_J SLHR - S - 17 - 0148$

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$_J SLHR - S - CSMC7 - 18 - 0208$

Orena, A. J., Byers-Heinlein, K., & Polka, L. (2019). Reliability of the Language Environment Analysis (LENA) in French-English Bilingual Speech. *Journal of Speech Language and Hearing Research*, *67*(2), 2491–2500. doi: 10.31234/osf.io/3xcvu

Perkell, J. S. (2013, December). Five Decades of Research in Speech Motor Control: What Have We Learned, and Where Should We Go From Here? *Journal of Speech, Language, and Hearing Research*, *56*(6), 1857–1874. doi: 10.1044/1092-4388(2013/12-0382)

Pierrehumbert, J. B. (2003, June). Phonetic Diversity, Statistical Learning, and Acquisition of Phonology. *Language and Speech*, *46*(2-3), 115–154. doi: 10.1177/00238309030460020501

Pisoni, D. B., Nusbaum, H. C., Luce, P. A., & Slowiaczek, L. M. (1985, August). Speech perception, word recognition and the structure of the lexicon. *Speech Communication*, *4*(1-3), 75–95. doi: 10.1016/0167-6393(85)90037-8

Pretzer, G. M., Lopez, L. D., Walle, E. A., & Warlaumont, A. S. (2019, November). Infant-adult vocal interaction dynamics depend on infant vocal type, child-directedness of adult speech, and timeframe. *Infant Behavior and Development*, *57*, 101325. doi: 10.1016/j.infbeh.2019.04.007

Redford, M. A. (2014). The perceived clarity of children's speech varies as a function of their default articulation rate. , *135*(5), 2952-2963.

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: $10.1044/2018_JSLHR - S - 17 - 0126$

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

Rubertus, E., & Noiray, A. (2020, June). Vocalic activation width decreases across childhood: Evidence from carryover coarticulation. *Laboratory Phonology: Journal of the Association for Laboratory Phonology*, *11*(1), 7. doi: 10.5334/labphon.228

Schneider, W., Eschman, A., & Zuccolotto, A. (2012). *E-Prime.* Psychology Softwre Tools, Inc. Pittsburgh, PA.

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

Shu, H., Peng, H., & McBride-Chang, C. (2008, January). Phonological awareness in young Chinese children. *Developmental Science*, *11*(1), 171–181. doi: 10.1111/j.1467-7687.2007.00654.x

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

Sosa, A. V., & Stoel-Gammon, C. (2012, April). Lexical and Phonological Effects in Early Word Production. *Journal of Speech Language and Hearing Research*, *55*(2), 596. doi: 10.1044/1092-4388(2011/10-0113)

Stoel-Gammon, C. (2011, January). Relationships between lexical and phonological development in young children*. *Journal of Child Language*, *38*(01), 1–34. doi: 10.1017/S0305000910000425

Storkel, H. L. (2002, May). Restructuring of similarity neighbourhoods in the developing mental lexicon. *Journal of Child Language*, *29*(2), 251–274. doi: 10.1017/S0305000902005032

Storkel, H. L., & Hoover, J. R. (2011, June). The influence of part-word phonotactic probability/neighborhood density on word learning by preschool children varying in expressive vocabulary. *Journal of Child Language*, *38*(03), 628–643. doi: 10.1017/S0305000910000176

Storkel, H. L., & Maekawa, J. (2005, November). A comparison of homonym and novel word learning: The role of phonotactic probability and word frequency. *Journal of Child Language*, *32*(4), 827–853. doi: 10.1017/S0305000905007099

Storkel, H. L., & Morrisette, M. L. (2002, January). The Lexicon and Phonology: Interactions in Language Acquisition. *Language, Speech, and Hearing Services in Schools*, *33*(1), 24–37. doi: 10.1044/0161-1461(2002/003)

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

Szewczyk, J. M., Marecka, M., Chiat, S., & Wodniecka, Z. (2018, October). Nonword repetition depends on the frequency of sublexical representations at different grain sizes: Evidence from a multi-factorial analysis. *Cognition*, *179*, 23–36. doi: 10.1016/j.cognition.2018.06.002

Vihman, M. M. (1993, January). Variable paths to early word production. *Journal of Phonetics*,

21(1-2), 61–82. doi: 10.1016/S0095-4470(19)31321-X

Vihman, M. M. (2017). Learning words and learning sounds: Advances in language development. *British Journal of Psychology*, *108*(1), 1–27. doi: 10.1111/bjop.12207

Vorperian, H. K., Kent, R., Lindstrom, M. J., Kalina, C. M., Gentry, L. R., & Yandell, B. S. (2005). Development of vocal tract length during early childhood: A magnetic resonance imaging study. *The Journal of the Acoustical Society of America*, *117*(1), 338–350. doi: 10.1121/1.1835958

Walsh, B., & Smith, A. (2002). Articulatory movements in adolescents: Evidence for protracted develop- ment of speech motor control processes. *Journal of Speech Language and Hearing Research*, *45*(6), 1119–1133.

Warlaumont, A. S., Richards, J. A., Gilkerson, J., & Oller, D. K. (2014, July). A Social Feedback Loop for Speech Development and Its Reduction in Autism. *Psychological Science*, *25*(7), 1314–1324. doi: 10.1177/0956797614531023

Weisleder, A., & Fernald, A. (2013, November). Talking to Children Matters: Early Language Experience Strengthens Processing and Builds Vocabulary. *Psychological Science*, *24*(11), 2143–2152. doi: 10.1177/0956797613488145

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

Williams, K. (2007). *Expressive Vocabulary Test, Second Edition* (Second ed.). San Antonio, TX: Pearson Education.

Xu, D., Yapanel, U., & Gray, S. (2009). *Reliability of the LENA Language Environment Analysis System in young children's natural home environment* (Technical Report lTR-05-2). Boulder, CO: LENA Research Foundation.

Zamuner, T. S. (2009, February). Phonotactic Probabilities at the Onset of Language Development: Speech Production and Word Position. *Journal of Speech, Language, and Hearing Research*, *52*(1), 49–60. doi: 10.1044/1092-4388(2008/07-0138)

Zamuner, T. S., Strahm, S., Morin-Lessard, E., & Page, M. P. (2018, July). Reverse production effect: Children recognize novel words better when they are heard rather than produced.

Developmental Science, *21*(4), e12636. doi: 10.1111/desc.12636

Zharkova, 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: $10.1044/2014_J SLHR - S - 11 - 0350$

Table 1

Stimuli used in word repetition tasks

REAL WORD	NONWORD	REAL WORD	NONWORD
candle [kændəl]	'kamig' [kæmɪg]	sidewalk [saɪdwak]	'sighprote' [saɪpɹot]
chicken [tʃɪkən]	'chihmig' [tʃɪmɪg]	sister [sɪstəɹ]	'sihplok' [sɪplok]
coffee [kafi]	'kahsep' [kasɛp]	suitcase [sutkes]	'soodros' [sudɹas]
cutting [kʌtɪŋ]	'kuhfeem' [kʌfim]	sunny [sʌni]	'suhbith' [sʌbɪθ]
kitchen [kɪtʃən]	'kihpon' [kɪpon]	tiger [taɪgəɹ]	'tieblor' [taɪbloɹ]
raisins [ɹezɪnz]	'raebith' [ɹebɪθ]	toaster [tostəɹ]	'toezell' [tozɛl]
rabbit [ɹæbɪt]	'rapoin' [ɹæpoɪn]	toothbrush [tuθbɹʌʃ]	'toografe' [tugɹaɪf]
reading [ɹidɪŋ]	'reefras'[ɹifɹas]	waiting [wetɪŋ]	'waymahg' [wemag]
rocking [ɹakɪŋ]	'rahlide'[ɹalaɪd]	washer [waʃəɹ]	'wahkrad' [wakɹæd]
running [ɹʌnɪŋ]	'ruhglok' [ɹʌglok]	water [watəɹ]	'wahprote' [wapɹot]
sandwich [sændwɪtʃ]	'samell' [sæmɛl]	window [wɪndo]	'wihmell' [wɪmel]
sharing [ʃeɹɪŋ]	'shaevahs' [ʃevas]		

Table 2

Correctly-produced words analyzed in results, by word type and CV sequence

CV sequence	Correct nonwords	Correct real words	Total
kæ	N=76/103 (73.79%)	N=86/103 (83.50%)	N=162
ka	80 (77.67)	91 (88.35)	171
kɪ	75 (72.82)	81 (78.64)	156
kʌ	81 (78.64)	95 (92.23)	176
ɹæ	64 (62.14)	70 (67.96)	134
ɹe	58 (56.31)	65 (63.10)	123
ɹi	52 (50.49)	64 (62.14)	116
ɹa	45 (43.69)	55 (53.40)	100
ɹʌ	46 (44.66)	55 (53.40)	101
sæ	66 (64.08)	75 (72.82)	141
ʃe	56 (54.37)	65 (63.10)	121
saɪ	77 (74.76)	88 (85.44)	165
sɪ	77 (74.76)	88 (85.44)	165
su	75 (72.82)	76 (73.79)	151
sʌ	57 (55.33)	89 (86.41)	146
taɪ	75 (72.82)	86 (83.50)	161
tʃɪ	63 (61.17)	77 (74.76)	140
to	82 (79.61)	95 (92.23)	177
tu	88 (85.44)	95 (92.23)	183
we	81 (78.64)	89 (86.41)	170
wa	72* (69.90)	91 (88.35)	163
wa	67† (65.05)	93 (90.29)	160
wɪ	76 (73.79)	91 (88.35)	167
Total	1589	1860	3449

* Reports the pair "washer" and [wakɹæd]

† Reports the pair "water" and [waprot]

Table 3

Summary statistics for vocabulary scores (N=103 children) and daylong recording measures (N=80 children)

	MEAN(SD)	RANGE
Adult word count/hour	1280.4(494)	322.8-2748.3
Child vocalization count/hour	244.7(91)	12.1-495.8
PPVT-4 Growth scale value	129(17)	85-160
PPVT-4 Standard score	119(17)	76-151
EVT-4 Growth scale value	134(15)	85-163
EVT-4 Standard score	117(19)	70-156

Table 4

Summary statistics for caregiver-identified late-talkers and not late-talkers

	Late-talkers (N=13)		Not late-talkers (N=90)	
	MEAN(SD)	RANGE	MEAN(SD)	RANGE
Adult word count/hour	1275(210)	1043-1551	1280.9(510)	322.8-2748.3
Child vocalization count/hour	220.3(91)	115.9-351.8	246.6(91)	12.1-495.8
PPVT-4 Growth scale value	118(12)	87-135	133(16)	85-160
PPVT-4 Standard score	110(13)	81-132	123(13)	76-151
EVT-4 Growth scale value	127(11)	85-139	138(14)	94-163
EVT-4 Standard score	110(14)	70-130	121(18)	75-156

Table 5

The effects of vocabulary and word duration upon coarticulation

Intercept	6.67***
	(5.06, 8.28)
	t = 8.13
	p = 0.00
Receptive vocabulary score (PPVT-4)	0.02**
	(0.004, 0.03)
	t = 2.65
	p = 0.01
Word duration (ms)	0.001*
	(0.0000, 0.001)
	t = 2.02
	p = 0.05
Log Likelihood	−9,163.64
Akaike Inf. Crit.	18,339.27
Bayesian Inf. Crit.	18,376.15
Note:	*p<0.05; **p<0.01; ***p<0.001

Table 6

The effects of speech practice and vocabulary upon coarticulation

Intercept	7.99***
	(7.20, 8.79)
	t = 19.71
	p = 0.00
Hourly child vocalization count	0.003**
	(0.001, 0.01)
	t = 2.74
	p = 0.01
Word duration (ms)	0.001*
	(0.0001, 0.001)
	t = 2.43
	p = 0.02
Expressive vocabulary score (EVT-2)	0.02*
	(0.005, 0.04)
	t = 2.57
	p = 0.02
Log Likelihood	−7,454.78
Akaike Inf. Crit.	14,923.57
Bayesian Inf. Crit.	14,965.08

Note: *p<0.05; **p<0.01; ***p<0.001

figures/figure$_1$.jpeg

Figure 1. CV coarticulation by receptive vocabulary score (PPVT-4)
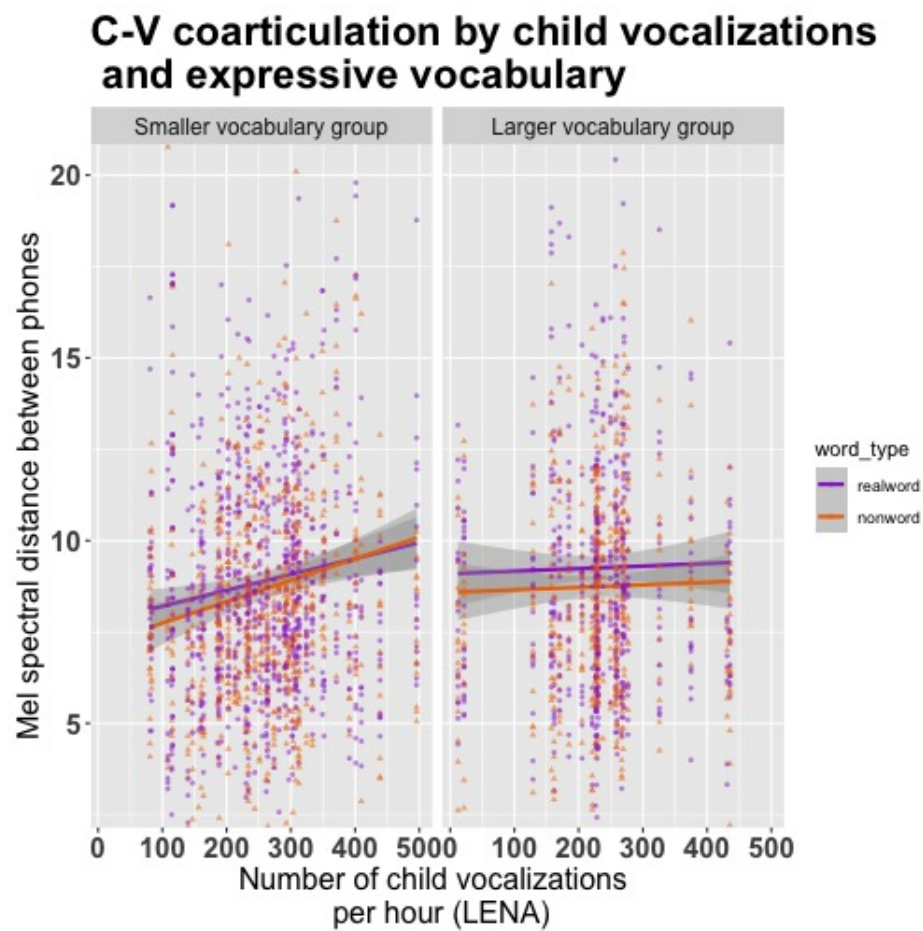
Figure 2. Child vocalization count by Mel spectral distance between CV phones with a median split by expressive vocabulary size (EVT-2).
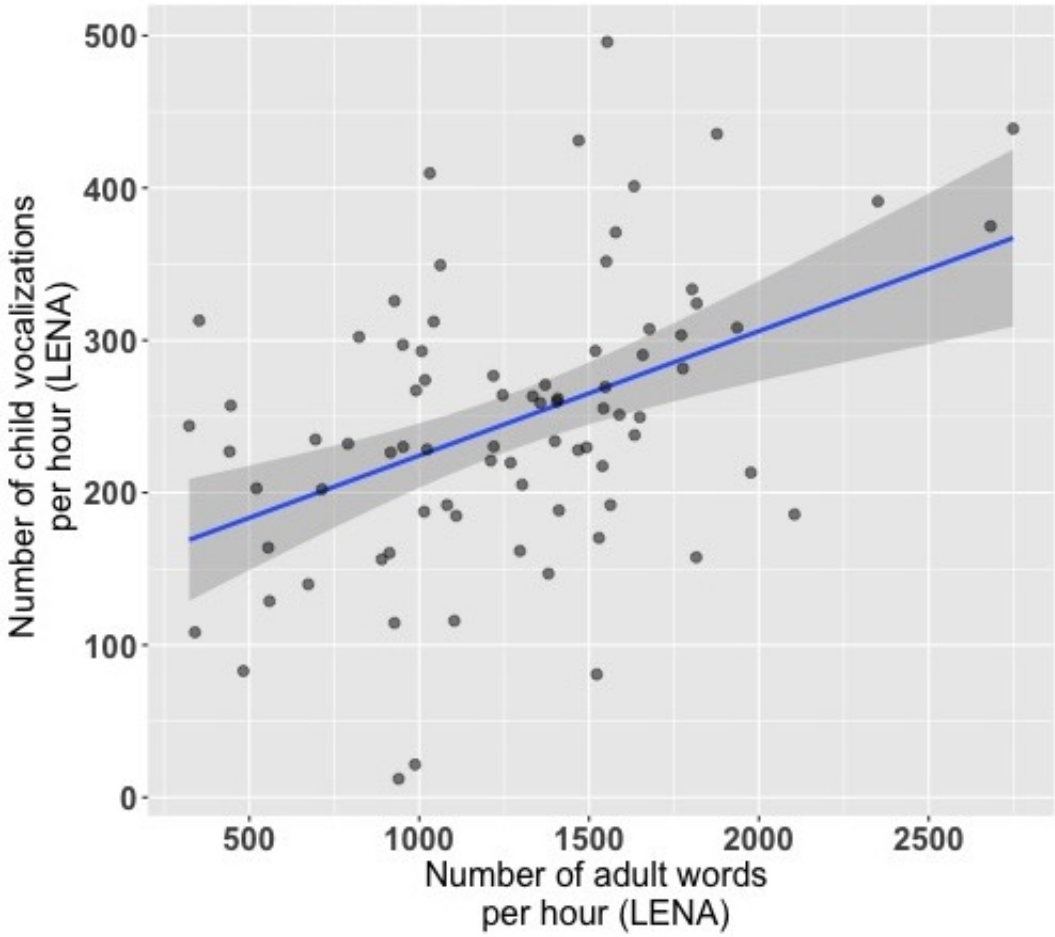
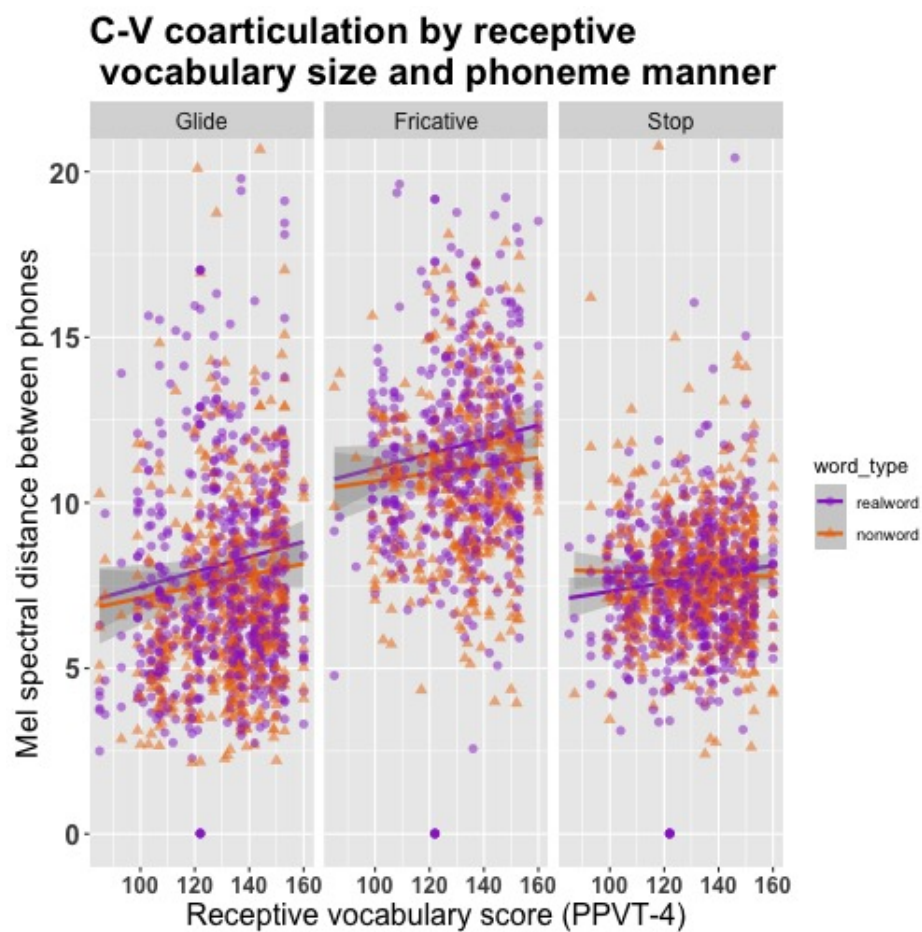*Figure 3*. Relationship between adult word count and child vocalization count

Figure 4. CV coarticulation by receptive vocabulary size and consonant manner
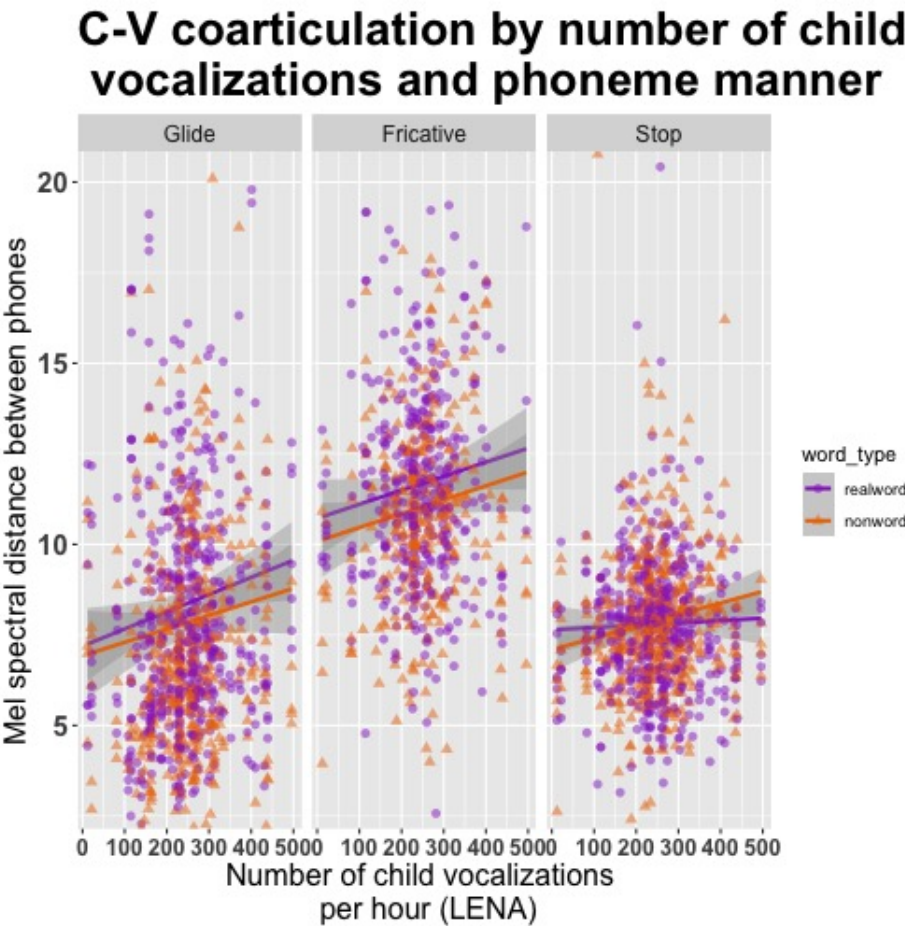
Figure 5. CV coarticulation by number of child vocalizations and consonant manner