
ELSEVIER
Journal of
Phonetics

Journal of Phonetics 41 (2013) 281–295

Contents lists available at SciVerse ScienceDirect

Journal of Phonetics

journal homepage: www.elsevier.com/locate/phonetics

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

Jae Yung Song[a,b,*], Katherine Demuth[c], Stefanie Shattuck-Hufnagel[b], Lucie Ménard[d]

[a] Cognitive, Linguistic, and Psychological Sciences Department, Brown University, Providence, RI 02912, USA
[b] Speech Communication Group, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
[c] Linguistics Department, Macquarie University, Sydney, New South Wales 2109, Australia
[d] Linguistics Department, Université du Québec à Montréal, Center for Research on Brain, Language, and Music, Montreal, Canada H3C 3P8

ARTICLE INFO

Article history:
Received 11 February 2012
Received in revised form
14 March 2013
Accepted 26 March 2013

ABSTRACT

Most studies of phonological development have explored the acquisition of segments, syllables and words using perceptual/transcription methods. Less is known about the articulatory aspects of early speech, or the development of articulatory-acoustic mapping. Recent research on adult speech finds that coarticulation effects are evidenced in both the acoustics and the articulatory gestures, and suggests tighter coarticulation and less variability for monomorphemic compared to polymorphemic segment sequences. The present study explored phonological context and morphological effects in the speech of five adults and five 2-year-olds, combining acoustic and articulatory analysis from ultrasound recordings. The results show that coarticulation effects are found in the word-final consonant cluster (*box*) for both adults and children. For children, these were evidenced only in the articulatory data. In addition, both age groups showed differences in tongue height between the monomorphemic (*box*) and bimorphemic (*rocks*) clusters, suggesting a possible morphological effect. These findings confirm that ultrasound methods can be successfully employed to explore aspects of early gestural development in children as young as 2, and raise many questions regarding the nature of speech planning processes as a function of lexical versus morphological form.

1. Introduction

Both coda consonants and coda consonant clusters present a challenge for young language learners, with simple codas often omitted, and clusters often reduced. Previous studies have explored these issues using impressionistic transcription (e.g., Demuth, Culbertson, & Alter, 2006; Fee, 1995; Kirk & Demuth, 2005; Smith, 1973), but a detailed understanding of the acoustics of child codas has emerged more slowly, and little is known about the articulatory aspects of children's attempted coda consonants. Furthermore, children sometimes make phonological contrasts that, though not perceived as such by adults, are evidenced in the acoustic signal. A full understanding of how and when phonological representations develop must therefore consider possible acoustic 'covert contrasts' in children's early speech (Macken & Barton, 1980; Scobbie, Gibbon, Hardcastle, & Fletcher, 2000). Research on adult speech has also shown that tongue movements do not always map reliably to acoustic events (Browman & Goldstein, 1990; Tiede, Perkell, Zandipour, & Matthies, 2001). Using ultrasound methods, Gick, Michelson, and Radanov (2006) have further demonstrated that some phonological contrasts that are not evidenced in the acoustic signal may be revealed in contrasting articulatory gestures. This suggests that, in both adult speech and that of the developing child, certain phonological contrasts may more easily be found by examining the articulatory record. Ultrasound methods, which afford an easy-to-apply and non-invasive means of collecting such data from young children, provide an ideal opportunity for exploring these issues more fully. In addition, because of the nature of their tongue tissue, young children image very well, resulting in high-quality videos of their tongue movements and articulatory gestures (Stone, 2005).

Research on adult speech has shown that consonant sequences are produced differently as a function of phonological and morphological context. Using data from ElectroMagnetic Midsagittal Articulometer (EMMA) recordings, Tiede et al. (2007) found that coda clusters were less variable in timing than heterosyllabic sequences of the same stop consonants (*pact op* vs. *pack top*). Cho (2002) showed tighter articulatory coordination and less acoustic

---

[*] Corresponding author. Present address: Department of Communication Sciences and Disorders, University of Wisconsin-Milwaukee, Milwaukee, WI 53211, USA. Tel.: +1 414 229 2665; fax: +1 414 229 2620.
E-mail address: songjy@uwm.edu (J.Y. Song).

0095-4470/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.wocn.2013.03.004

variability for Korean consonants produced within versus across a morpheme boundary. These findings suggest that the speech planning process for adults is sensitive to word and morpheme boundaries. This is particularly significant for English, where many word-final clusters are bimorphemic (e.g., *dogs*, *hits*, *picked*). Recent research has also shown that 2-year-olds are less accurate at producing 3rd person singular –*s* when the morpheme is part of a complex coda cluster (e.g., *needs*) than when it is not (e.g., *sees*), suggesting that young children's production of this morpheme is affected by the phonological complexity of a coda (Song, Sundara, & Demuth, 2009). This raises the possibility that children (and adults) may exhibit differences in either acoustics and/or in articulation when producing monomorphemic versus bimorphemic coda clusters (e.g., *box* vs. *rocks*).

Little is known about the gestural underpinnings of child speech, and what these may reveal about early phonological or morphological representations, in part because of the lack of non-invasive methods for collecting articulatory data and in part because articulatory data are not often combined with acoustic analyses. The present study begins to address this issue by demonstrating that portable ultrasound methodology can be used to compare tongue movements in the speech of children as young as 2 years old with those of adults. We investigated tongue movements across different coda consonant conditions, first identifying the period of consonant constriction in the acoustics, and then examining tongue movement trajectories during the constriction using ultrasound techniques. In particular, we wanted to examine tongue configuration as a function of both adjacent phonological segments within the coda cluster and morphological complexity.

Phonological context can affect several aspects of the acoustic and articulatory characteristics of a coda. For instance, consonants are typically shorter in duration in most cluster environments than in singleton environments, and this durational difference suggests that consonantal gestures may partially overlap in time when producing clusters (Klatt, 1976; Browman & Goldstein, 1986). Thus, the production of a complex coda (e.g., *cats*) involves a change in both the acoustic and articulatory characteristics of the individual speech sounds as the articulation of one sound is influenced by the other. Coarticulation in consonant clusters has received considerable attention in the literature, and several studies on the acoustics and articulation of such speech sound sequences have shown robust coarticulation effects in adult speech. For example, acoustic analysis of speech productions from eight American English speakers showed that the formant transitions following the stop release were systematically influenced by the nature of the preceding fricative (e.g., /stɑ/ vs. /ʃtɑ/), demonstrating carryover coarticulation (Repp & Mann, 1982). Similar findings have been reported in articulatory studies. For example, in an ultrasound investigation of the articulation of non-native consonant clusters in five English-speaking adults, Davidson (2005) showed that tongue body position for /s/ was higher in the cluster /sk/ than in the sequence /sək/, demonstrating anticipatory coarticulation. That is, when /s/ is immediately followed by /k/, it coarticulates more closely with the high tongue body position of the /k/ gesture, so that tongue position for the /s/ has a starting position higher in the vocal tract. (For further articulatory (electropalatography (EPG)) studies examining coarticulation effects on the production of stop consonants in clusters, see Hardcastle (1985) and Byrd (1996)).

These findings from studies of adult speech raise interesting questions about how adult-like gestural coordination develops over time. Of particular interest is whether children coarticulate more (or less) than adults. The findings on this point are conflicting. Some studies report that children exhibit less coarticulation than adults, suggesting that coarticulation develops late. For example, using spectral analysis, Sereno, Baum, Marean, and Lieberman (1987) found that adults' production of consonants was significantly affected by the roundness of the following vowel (e.g., *di* vs. *du*). Children aged 3–7 years showed some comparable acoustic coarticulation effects, but there was a considerable individual variability. Furthermore, adult listeners could fairly accurately identify the vowel just by listening to the consonant excised from the consonant–vowel sequence produced by adults, but their performance was significantly worse when listening to child productions. Cheng, Murdoch, Goozée, and Scott (2007) examined the coarticulation effect on the production of consonant clusters /kl/ and /st/ in 4 age groups (6–7 years, 8–11 years, 12–17 years, and adults). Their EPG data of tongue-to-palate contact patterns revealed that the degree of coarticulation increased with age. In particular, the first sign of adult-like coordination between sections of the tongue was observed in the 8–11-year-old group, and the refinement of lingual coordination continued into adolescence. On the other hand, some studies have shown that adults and children around 3–5 years of age exhibit the same amount of coarticulation (Katz, Kripke, & Tallal, 1991; Turnbaugh, Hoffman, Daniloff, & Absher, 1985). Still other researchers report that children coarticulate more than adults. For example, in an acoustic study comparing the production of fricative-vowel syllables between adults and children aged 3–7 years, Nittrouer, Studdert-Kennedy, and McGowan (1989) showed that the degree of fricative (/ʃ/, /s/)-vowel coarticulation decreased with age. Based on the ultrasound images of tongue movement, Zharkova, Hewlett, and Hardcastle (2011) showed that even children as old as 6 to 9-year-olds coarticulate more than adults when producing fricative (/ʃ/)-vowel sequences. In a follow-up ultrasound study, however, Zharkova, Hewlett, and Hardcastle (2012) found significant coarticulatory effects of the vowel on the preceding /s/ in adults, but not in 6–9-year-olds. The authors suggested that children's lack of ability to control tongue tip/blade and body independently might not allow them to anticipate the tongue configuration of a following vowel while producing an initial /s/. In sum, there is yet no coherent picture of the development of adult-like patterns of coarticulation. Some of the conflicting results may be due to the different segments or ages tested, and/or different experimental methodologies and procedures. The fact that children's speech is inherently more variable than that of adults (e.g., Imbrie, 2005) may also have contributed to the different findings across studies.

With these issues in mind, the major focus of the present study was to examine the effect of coarticulation and morphological complexity on coda production. Many inflectional morphemes in English, such as the plural morpheme, appear in coda position, often forming a coda cluster at the end of a word (e.g., *cats*). There have been several attempts to evaluate morphological effects in such contexts in adult speech, across various languages. For example, Walsh and Parker (1983) showed that the duration of English plural –*s* (e.g., *laps*) is systematically longer than that of monomorphemic –*s* (e.g., *lapse*). However, the average difference in length measured between plural –*s* and monomorphemic –*s* was only 9 ms, and the study lacked statistical analysis. Cho (2002) provided articulatory evidence showing how morpheme boundaries affect intergestural timing in Korean. His electromagnetic midsaggital articulography (EMA) and EPG data showed that gestures were coordinated more stably inside a monomorphemic word than across a morphemic boundary, although they were homophonous on the surface. The variability in intergestural timing was measured by standard deviations for measured time intervals between various articulatorily defined points, such as the midpoints of plateaus of consecutive consonant gestures called the C-centers (Browman & Goldstein, 1988). Cho's results suggest that children might also show less variability in the production of monomorphemic coda clusters (e.g., *box*) compared to the same segments in a coda cluster involving a morpheme (e.g., *rocks*).

Some grammatical morphemes are acquired late and are variably omitted in early speech (Brown, 1973), and this is exacerbated in children with language impairment. For example, the age of mastery (indicated by over 90% correct use of a morpheme in obligatory contexts) for even plural –*s*, one of the earliest-acquired morphemes in English, is not until 27–33 months (Brown, 1973). Traditionally, this has been attributed to incomplete or still-developing syntactic and semantic representations (e.g., Wexler, 1994). However, recent studies have shown that the phonological shape of syllables and words, and the prosodic contexts in which they appear, can influence children's production of grammatical function items (Gerken, 1996; Marshall & van der Lely, 2007; Song et al., 2009; Theodore, Demuth, & Shattuck-Hufnagel, 2011). Thus, morphologically complex words are often phonologically (as well as syntactically/semantically) challenging for young children, leading to greater variability in production compared to morphologically simple words. This makes it particularly important for studies to move beyond phonemic transcriptions of children's speech to provide information about the acoustic and articulatory realization of early grammatical morphemes (cf. Theodore et al., 2011).

Ultrasound techniques have been increasingly adopted in studies to address various issues in articulatory phonology and phonetics (Davidson, 2005; Gick, Campbell, Oh, & Tamburri-Watt, 2006; Miller & Finch, 2011). Ultrasound has also been used in remediation studies for those with articulatory problems (Bernhardt, Gick, Bacsfalvi, & Ashdown, 2003). The use of ultrasound in this particular study is ideal since it allows us to examine the movements and shape of the tongue surface, including regions behind the tongue dorsum, in a non-invasive way, contrary to flesh-point tracking methods such as electromagnetometry. Furthermore, it does not require much preparation to use (as compared to, say, training children in MRI procedures), and adequate amounts of data can be collected quite quickly, which is an advantage when dealing with very young speakers. The main issue with ultrasound recordings, as with any imaging technique, however, is measurement artifacts due to the movement of the head relative to the probe. Researchers have developed a variety of methods to ensure head stabilization, as well as consistent orientation of the probe angle to ensure appropriate image capture. Because these stabilization methods can be problematic for very young children at early stages of language development, only a few studies have used ultrasound methods to examine the tongue movements in children below the age of 5 (e.g., Gick et al., 2008; Ménard & Noiray, 2011; Ménard et al., 2010).

The present study used portable ultrasound measures without head stabilization hardware to examine the acoustics and articulation of both adult and child speech, offering detailed information about early word production and phonological representations more generally. As far as we know, this is the first major study using ultrasound to investigate the acoustics and gestural organization of speech in 2-year-olds. To assist with head stabilization without hardware, participants fixated on a visual stimulus as they produced the target words. The probe was held in place by the experimenter or trained adult participant. Importantly, the entire task lasted 10 min or less, increasing the likelihood that the children would sit still through the entire procedure. This is described in more detail in Section 2.

The goal of the present study was to examine whether adults and 2-year-olds would show the effects of adjacent consonants and morphological complexity in the acoustics and articulation of coda cluster consonants during an elicited imitation task. First, we wanted to know if adults and children show similar patterns of coarticulation effects in /ks/ clusters. This might be evidenced in the shorter duration of segments in the cluster compared to singleton coda, and/or in the effects of tongue gestures on the other segment of the cluster (influence of /k/ on /s/, and vice versa). Second, we wanted to determine if there were any morphological effects. The null hypothesis was that there should be no differences in either the acoustics or the articulatory gestures between the monomorphemic cluster in *box* and the bimorphemic cluster in *rocks*. Of course, given the previous literature, we expected that there might be significant differences between the two, perhaps suggesting effects of higher-level representations on the speech planning process.

2. Materials and methods

2.1. Participants

The participants were five adult (4 females, 1 male) and five 2-year-old (3 females, 2 males) monolingual speakers of American English recruited in Providence, RI. The adults ranged in age from 18 to 33 years (mean 23.5 years). The children's ages ranged from 24.8 months to 30 months (mean 27.4 months (2;3 years)). All children had normally developing speech and language skills according to parental report on the MacArthur Communicative Development Inventory (CDI) (Dale & Fenson, 1996). On average, the children produced 93.6 words ($SD = 7.89$) out of 100 words listed in MacArthur CDI (short form). An additional 13 children were not included in the analysis: Eight children failed to complete the experiment because they did not speak or became fussy during the experiment. Five children completed the experiment but the acoustic/ultrasound recordings were not clear enough ($n = 3$) or not saved properly ($n = 2$). The attrition rate was higher than that reported in previous speech production experiments with children of this age (e.g., Gerken, 1996; Song et al., 2009). However, this was a new procedure not previously used with this age group. There was therefore a learning curve with respect to helping child participants feel at ease and pacing the experiment to ensure completion of data collection before the child began to get restless. It is expected that in future experiments, with methodology more closely tailored to the needs and attention span of the child, this attrition rate will decrease.

2.2. Stimuli

We designed the elicited imitation experiment to be as easy and quick to complete as possible, to help ensure the child's cooperation in sitting still and attending to the task.[1] Ideally, we wanted to use 4 familiar, high frequency, monosyllabic, picturable nouns ending with monomorphemic /s/, /k/, and /ks/, and bimorphemic /ks/, and have participants repeat the target words 6 times, for a total of 24 test items. It was critical, however, that the same vowel be used in all the stimulus items, to avoid coarticulation confounds with the coda consonants. Furthermore, low vowels are known to image better than high vowels because tongue surfaces tend to have steep slopes for high vowels (Stone, 2005). This was difficult to achieve with familiar, picturable English words. The compromise was to use three real words (*rock*, *rocks*, *box*) and one (familiarized) nonce word (*das*), all containing the low vowel /ɑ/, but with different coda configurations. *Das* and *rock* each contained a singleton coda /s/ and /k/, respectively. *Box* and *rocks* contained the monomorphemic and bimorphemic consonant cluster /ks/, respectively. The word *rocks* was chosen because the plural form in the word was likely to be productive, rather than a plural dominant or high-frequency plural form such as *socks*.

The target word stimuli were recorded by a female speaker of American English in a child-directed speech register. Acoustic measurement of the 4 target words revealed that the duration of the /k/ closure in each target word ranged from high to low as follows: *rock* (205 ms), *box* (186 ms), *rocks* (182 ms). Thus, consistent with the literature on adult speech (Klatt, 1976), /k/ closure duration was longer in the singleton context compared to the cluster context. On the other hand, the duration of /s/ frication turned out to be similar across different coda contexts: *das* (264 ms), *rocks* (260 ms), *box* (259 ms).

These 4 target audio prompts were then each paired with 6 different pictures to elicit 6 repetitions of each target word, for a total of 24 test items. Six different pictures were used to avoid monotonous presentations and make the experiment as engaging as possible. For the nonce item *das*, a picture of an hourglass was used as the visual stimulus. An hourglass was selected because it has been successfully used as a stimulus for an

---

[1] Although the elicited imitation task was more appropriate in our study, both adults' and children's production characteristics may differ between imitation and spontaneous speech (e.g., Hodson & Paden, 1991).

unfamiliar object in previous studies involving infant speech perception tasks (e.g., White & Morgan, 2008). The pictures were chosen on the basis of being realistic representations of objects, holding similar levels of interest. The 24 items were presented to participants in pseudo-randomized order.

2.3. Procedure

The participants were invited into a sound-attenuated test room to listen to and repeat the prerecorded target words. The entire procedure took approximately 10 minutes. The children sat on their parent's lap and looked at a computer monitor where a puppet invited them to “play a game” involving repeating the names of the items displayed. On each trial, a picture of the target item appeared on the computer monitor along with the auditory prompt (e.g., “box”). This visual fixation helped to minimize head movement while the experimenter held the ultrasound transducer stable under the child's chin (see Fig. 1). The same stimuli and setup were used for adult subjects, except that they were asked to hold the ultrasound probe under their chin themselves while sitting still and looking at the monitor. As it was difficult to completely stabilize the child’s head during the experiment, post hoc corrections in subsequent analyses used ratios and changes in height in order to compare images across frames, rather than absolute tongue positions. This method is discussed in more in detail in Section 2.4.2.

A portable Sonosite 180 Plus ultrasound machine with a C11/7-4 MHz 11-mm broadband curved array transducer was used to collect midsagittal images of the tongue during the production of target words. The probe angle was 84°. A wider probe angle would have impaired image quality, due to the longer time required for one scan. The acoustic signal was recorded using a Shure KSM137 unidirectional microphone that was connected to a M-audio DMP3 preamplifier. Both ultrasound and acoustic signals were recorded through a Sony mini-DV DCR-TRV103 digital camcorder in NTSC format (30 fps). The data were then downloaded to a computer using Adobe Premiere Elements software (www.adobe.com). The audio signal was digitized at a sampling frequency of 48 kHz with 16-bit quantization.

2.4. Analysis

As mentioned earlier, each participant repeated each of the 4 target words 6 times, for a total of 24 tokens. In preparation for the acoustic and ultrasound analyses, each token was excised from a long video file into a separate file using QuickTime Pro software (www.apple.com/quicktime). The 3 acoustically and visually cleanest tokens (i.e., those with less background noise in the audio, and clearer tongue contour in the video) for each target word were then selected for analysis.

2.4.1. Acoustic measures and analysis

Acoustic analysis of the audio was performed using Praat software (Boersma & Weenink, 2005). We coded the data for acoustic landmarks (Stevens, 2002) related to coda consonant constriction: /k/-closure-begin and release burst for /k/, and onset and offset of frication noise for /s/. Performing analyses at the level of individual acoustic landmarks provided a richer and more systematic constellation of observations than simply labeling at the level of the segment. The beginning of /k/ closure was marked at the offset of clear F2 energy in the spectrogram, a point which corresponds to the offset of high-amplitude regularity in the waveform associated with an open vocal tract (vowel) (Turk, Nakai, & Sugahara, 2006). The end of /k/ closure was marked at the left edge of the first burst transient (usually manifested as a sharp spike in the waveform), signaling the release of the pressure build-up for a stop coda consonant. The beginning and end of /s/ frication noise were labeled at the onset and offset of high-frequency aperiodic noise associated with fricative production.

The measure that was used in the acoustic analyses was the duration of coda consonants in various contexts. For the duration of /k/ closure, the time difference between the beginning of the /k/ closure and its release was computed. Likewise, the time difference between the onset and offset of /s/ frication noise was calculated to compute the duration of /s/. Then we examined these durations as a function of /s/ vs. /k/, segmental context, and morphological complexity, using paired *t*-tests.

In addition to acoustic coding for the coda consonants, the midpoint of the vowel /ɑ/ was also identified in each token. To determine the midpoint of the vowel, the beginning and end of the vowel were defined primarily on the bases of the onset and offset of a clear F2 energy in the spectrogram, respectively. For two of the target words *rock* and *rocks*, the preceding liquid made it difficult to identify the beginning of the vowel. Thus, the best estimation of vowel onset was made based on visual information including the changes in amplitude and the formants (especially F3), as well as perceptual information. After determining the midpoint of the vowel, we examined the ultrasound images from the midpoint of the vowel to the end of the coda consonant constriction (i.e., to the end of /k/ closure or /s/ noise), as described in the next section.

2.4.2. Ultrasound measures and analysis

To analyze tongue movements during the production of target coda consonants, the individual video frames between the acoustically-defined midpoint of the vowel and end (release) of the coda were determined for each token and extracted from the ultrasound video using VirtualDub software (www.virtualdub.org). These extracted still images were then loaded into EdgeTrak (Li, Kambhamettu, & Stone, 2003), a program that tracks the tongue contours in the ultrasound image and extracts them as a series of *x*–*y* coordinate points which are saved in a text file. One hundred coordinate points were extracted for each tongue contour and then used as input data for Lingua (Aubin & Ménard, 2006; Ménard, Aubin, Thibeault, & Richard, 2012),

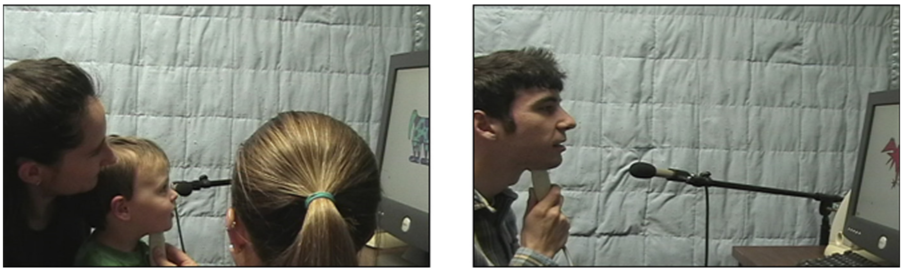

Fig. 1. The experiment setup with the child (left) and adult (right) participant.

a Matlab application developed at the Phonetics Laboratory at University of Quebec, Montreal. Using Lingua, we extracted three parameters quantifying the shape and position of tongue contours: asymmetry, curvature, and height of the highest point of the tongue. These are discussed in detail below.

Two of the parameters characterized the tongue shape (asymmetry, curvature) and one reflected relative tongue position in the vertical dimension (change in the height of the highest point of the tongue). These parameters were derived from the points on a triangle formed by connecting the two ends of the tracked tongue contour (Points A and B in Fig. 2) and the highest point on the tongue contour (Point C in Fig. 2), from which a vertical line segment forms a 90° angle with the triangle base.

Asymmetry is a measure of the position of the mass of the tongue relative to the whole tongue. This is defined as the ratio of the distance AD over the distance DB (AD/DB) in Fig. 2. A bigger value means that the mass of the tongue is positioned toward the front of the vocal tract. Curvature is a measure of how bunched the tongue is. This is defined as CD/AB in Fig. 2. A bigger value indicates a more bunched tongue, whereas a smaller value indicates that the tongue is flatter. It must be mentioned that the curvature and asymmetry measures are based, in part, on the beginning and end points of the extracted tongue contour. The x and y-coordinates of those two points are variable on the ultrasound image, due to jaw shadow, hyoid bone shadow, or image quality, for instance. This is a well-known problem inherent to the ultrasound technique. In order to test the robustness of the asymmetry and curvature measures, we simulated various ultrasound movements and artifacts using an articulatory model of speech production (Ménard et al., 2012). It was shown that ultrasound probe movements had no significant effect on curvature degree and asymmetry values (referred to as “curvature position” in Ménard et al., 2012), contrary to vowel quality. Thus, these measures are robust to variation in tongue surface extraction. The third measure was the change in height of the highest point of the tongue (Point E in Fig. 2) from the midpoint of the vowel to the highest point of the tongue in the following coda consonant. The highest point of the tongue was defined as the point of the tongue contour whose y-value is the maximum in the x–y space. This point was not necessarily identical to Point C, which was the peak of a triangle, i.e. the point on the tongue surface with the greatest distance from the triangle base. Also, like the other points, Point E could not with certainty be identified with the same physical point on the tongue contour across frames. However, as the height of Point E is known to reflect tongue height in vowels (Aubin & Ménard, 2006), it was expected to provide useful information about the tongue height for coda consonants as well.

In order to examine the change in the highest point of the tongue from the midpoint of the vowel to the coda consonant, we first identified the height of Point E in the image that was extracted at the temporal midpoint of the vowel. The tongue height at the midpoint of the vowel was then subtracted from that at each subsequent frame until the end of /k/ closure (i.e., until the release of /k/) or the end of /s/ frication noise. Zero indicated no change in the height of Point E between the vowel midpoint and coda. Positive values indicated an increase in height from the midpoint of the vowel. This method allowed us to examine how tongue height changes over time during /k/ closure and /s/ noise compared to its value at the midpoint of the vowel. Thus, for each token, there was a set of values from individual frames during /k/ closure and /s/ noise indicating the change in tongue height during coda consonant constriction. These sets of values were later used in linear mixed-effects regression analysis to compare the overall difference in the change in height between tokens.

The method of comparing the tongue height of the preceding vowel /ɑ/ against the tongue height for the coda consonant was primarily used because we were not able to completely stabilize the child's head during the experiment, so that it may have varied across target word utterances (although this is unlikely during a single target word utterance). Thus, we chose to have the same vowel /ɑ/ in all target words, and the tongue height at the middle of the vowel /ɑ/ served as a consistent reference point for comparison of height for the following coda consonant within each target word utterance across trials. This method helped to achieve consistency of data within each word, as well as between-speakers. Note also that the first two ratio measures, asymmetry and curvature, were independent of absolute position of the tongue.

As was mentioned earlier, linear mixed-effects regression models were used to test how tongue shape and relative tongue height/position change in individual speakers as a function of different coda consonant configurations. We carried out linear mixed-effects regression analyses on individual speakers because we were interested in the individual patterns in articulation, which were expected to vary considerably across individuals.

A mixed-effects regression analysis was used to provide insight into the full structure of the data by incorporating both fixed- and random-effects (for further information, see Baayen (2008) and Baayen, Davidson, and Bates (2008) and Johnson (2008)). A factor is considered to be fixed if the levels of the factor are selected by the researcher with the purpose of examining the effects of the levels. In contrast, a factor is considered to be random if the levels of the factor are viewed to be randomly sampled from a larger population. Although fixed effects are usually the primary interest, we were interested in the variances accounted for by the random effects so that the fixed effects could be properly evaluated. Analyses were carried out using the *R* statistical computing software (R Development Core Team, 2011), and in particular the lme (linear mixed-effects models) package. In each mixed-effects regression model, the dependent variable was one of the three measures of tongue shape and relative position during /k/ closure and /s/ noise in each of the four target words (*box*, *das*, *rock*, *rocks*): asymmetry, curvature, and change in tongue height from the midpoint of the vowel to the coda during coda consonant constriction. For the independent variables, the models included one fixed-effect factor, coda configuration (e.g., /s/ in monomorphemic *box* vs. /s/ in bimorphemic *rocks*), and one random-effect factor, the repetition of the word. The random-effect factor was included to control for any repetition-specific differences. To this end, the values obtained from all frames during the /k/ closure and /s/ frication noise from the three

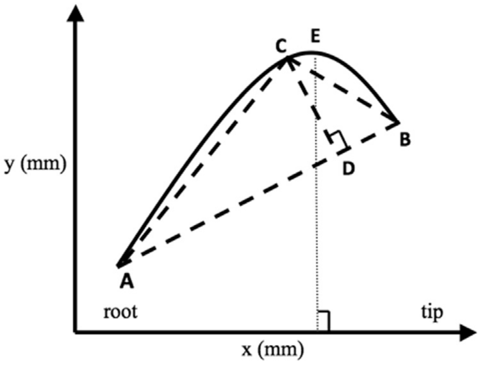

Fig. 2. A schematized representation of the tongue contour (solid line) and parameters based on a triangle fitted underneath the contour (dashed lines).

repetitions were used as individual data points in each mixed-effects regression analysis. The mean number of frames analyzed for children was greater than that for adults due to children's slower speaking rate: the average number of frames for /k/ closure across three words (*rock*, *box*, *rocks*) from 5 subjects in each group was 15 for adults and 26 for children, while the average number of frames for /s/ noise across three words (*das*, *box*, *rocks*) from 5 subjects in each group was 24 for adults and 32 for children. In Section 4, we present the results of mixed-effects analysis in each condition.

2.5. Predictions

We made several comparisons to examine how segmental context and morphological complexity affected the acoustics and articulation of coda consonants: (1) a baseline measure: /s/ in *das* [dɑs̲] vs. /k/ in *rock* [ɹɑk̲]; (2) a cluster coarticulation measure for /s/:/s/ in *das* [dɑs̲] vs. /s/ in *box* [bɑks̲]; (3) a cluster coarticulation measure for /k/:/k/ in *rock* [ɹɑk̲] vs. /k/ in *box* [bɑk̲s]; (4) a morphological complexity measure for /s/:/s/ in *box* [bɑks̲] vs. /s/ in *rocks* [ɹɑks̲]; and (5) a morphological complexity measure for /k/:/k/ in *box* [bɑk̲s] vs. /k/ in *rocks* [ɹɑk̲s]. For the baseline comparison in acoustic analysis, it was expected that /s/ would be generally longer than /k/. Based on the literature on adult speech, we also predicted that the duration for /k/ and /s/ would be longer in singleton coda contexts than in complex coda contexts (Klatt, 1976), and longer in the bimorphemic target word than in the monomorphemic word (Walsh & Parker, 1983).

For the baseline comparison in the articulatory analysis, the singleton coda /k/ and /s/ were expected to show systematic differences in all three measures. A bigger asymmetry for vowels indicates that the mass of the tongue is positioned toward the front (Aubin & Ménard, 2006). Thus, the more front constriction for coronal /s/ predicted a bigger asymmetry for this consonant than for velar /k/. However, if asymmetry is independent of the consonant constriction, there should be no consistent difference between /k/ and /s/. The raised dorsum for /k/ predicted a larger curvature and increase in tongue height for this consonant compared to /s/. Furthermore, it was predicted that, unlike singleton /k/ and /s/, /k/ and /s/ in the /ks/ coda cluster would show coarticulation effects, with each consonant exhibiting articulatory influences of the other. That is, on all three measures, /k/ in /ks/ in *box* would show characteristics of the adjacent /s/ to some extent, whereas singleton /k/ would not show such characteristics (e.g., smaller curvature for /k/ in /ks/ compared to singleton /k/; for full specific predictions on individual measures, see Appendix A). Likewise, /s/ in /ks/ in *box* would show characteristics of the adjacent /k/ to some extent, whereas singleton /s/ in *das* would not. Finally, based on previous studies showing a tighter coarticulation in monomorphemic sequences (Cho, 2002), we predicted a more robust coarticulation effect for monomorphemic /ks/ in *box* than for bimorphemic /ks/ in *rocks*. Specifically, if /ks/ in *box* is more tightly coarticulated, then the /k/ in /ks/ in *box* would be more like singleton /s/ than /k/ in /ks/ in *rocks* would be on all three articulatory measures. Likewise, the /s/ in the /ks/ in *box* would be more like singleton /k/ than /s/ in the /ks/ in *rocks* would be. This is consistent with the expectation that the acoustic duration of monomorphemic /ks/ would be shorter than that of bimorphemic /ks/ (Walsh & Parker, 1983).

3. Acoustic results

For the acoustic analyses, we examined the duration of /k/ closure and /s/ noise as a function of (1) /s/ vs. /k/ (baseline), (2) singleton vs. consonant cluster for /s/, (3) singleton vs. consonant cluster for /k/, (4) morphological complexity for /s/, and (5) morphological complexity for /k/. In each condition, the durations were compared between each pair of words using paired *t*-tests. As there were 5 comparisons made within each group, a Bonferroni correction was used to adjust the alpha level (.05/5 = .01). Fig. 3 summarizes the results for both adults and children. Although we used the same keys for significance levels ($*p<05$, $**p<01$, $***p<001$) across all figures, only the results with *p*-value equal to or less than .01 were considered significant here.

Although no direct statistical comparison was made between children and adults, an examination of the acoustic durations of /s/ and /k/ revealed that the durations were overall longer and more variable in child speech than in adult speech, as expected. Below we highlight the acoustic results of

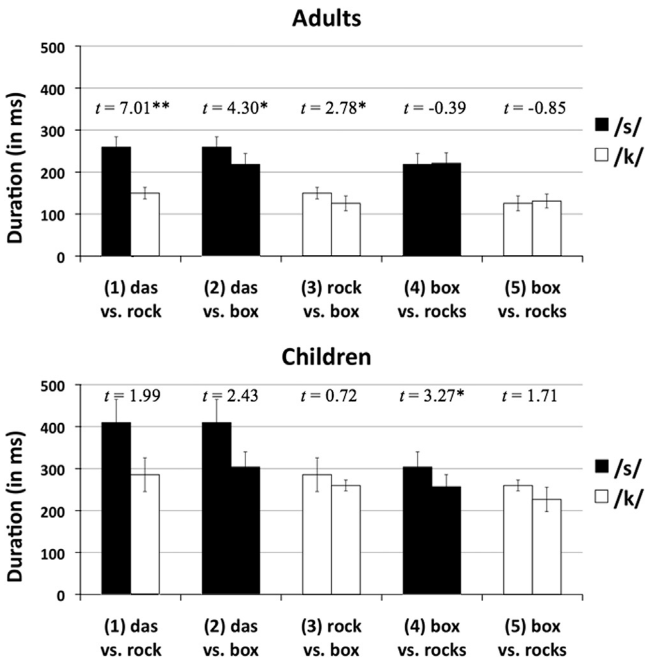

Fig. 3. The average duration of /k/ closure and /s/ noise for each comparison: (1) *das* [dɑs̲] vs. *rock* [ɹɑk̲], (2) *das* [dɑs̲] vs. *box* [dɑks̲], (3) *rock* [ɹɑk̲] vs. *box* [bɑks̲], (4) *box* [bɑks̲] vs. *rocks* [ɹɑks̲], (5) *box* [bɑk̲s] vs. *rocks* [ɹɑk̲s]. Error bars represent standard error. *T*-values were obtained from paired *t*-tests. As there were 5 subjects in each group, degree of freedom was 4 in all comparisons (*Note*: $* = p < .05$ (not significant), $** = p < .01$, $*** = p < .001$).

the five comparisons (1)–(5): (1) Baseline results show that singleton coda /s/ in *das* had a significantly longer duration than singleton coda /k/ in *rock* for adults. Although children showed the same tendency, the differences were not significant, probably due to the large variability in their production of /s/. (2, 3) With respect to the effects of coda complexity, as predicted, the duration of both /s/ noise and /k/ closure in adult's speech was generally longer in singleton coda contexts (*das*, *rock*) than in complex coda contexts (*box*). However, the differences did not reach the .01 level of significance, although the *p*-values were less than .05. For children, there was only a non-significant trend in the same direction for both /s/ and /k/ durations. Again, this appears to be due to a considerable amount of variability in their segment durations. (4, 5) With respect to the effects of morphology, adults did not show any difference in duration of /s/ and /k/ as a function of the morphological context. Thus, in contrast to Walsh & Parker (1983), we found no evidence of longer duration for plural /s/ as compared to monomorphemic /s/ in adults' production. Interestingly, for children, there was a tendency for the duration of /s/ and /k/ in monomorphemic *box* to be longer than in bimorphemic *rocks*, but the differences did not reach signficance. Finally, we also compared the duration of monomorphemic vs. bimorphemic /ks/ as a whole. If there were more gestural overlap for monomorphemic /ks/, it should be shorter in duration than bimorphemic /ks/. The results showed no significant effect of morphological complexity on the duration of /ks/ in adult speech, $t(4)=-.96$, $p=.39$. For children, the duration of monomorphemic /ks/ was significantly longer than that of bimorphemic /ks/, $t(4)=4.45$, $p<.05$, largely due to the longer duration of monomorphemic /s/. Thus, neither adults nor children exhibited shorter duration for monomorphemic /ks/.

We then examine the ultrasound results to determine if there were any differences in articulatory gestures as a function of coarticulation and morphological context. Although the acoustic analysis was not revealing in this respect, it is possible that there might be articulatory covert contrasts that would provide evidence of morphological organization in the speech planning process.

4. Ultrasound results

4.1. Baseline: /s/ in das [dɑs̲] vs. /k/ in rock [ɹɑk̲]

Fig. 4 shows how tongue asymmetry, curvature, and height change differ between the /s/ in *das* and the /k/ in *rock* for both adults and children when repetition differences were taken into account. This figure shows that /s/ and /k/ differed significantly in asymmetry in most of the adults' speech, but the direction of difference was counter to our prediction; that is, the mass of the tongue was positioned more toward the front when producing /k/ than /s/. This result may be explained by the fact that the asymmetry measure is independent of the actual location of the constriction for the coda consonant. For example, Fig. 5 shows ultrasound images extracted at the onset of /s/ frication noise (top right) and /k/ closure (bottom right), and the image frames immediately preceeding them (left counterparts) (*note*: tongue tip is at the right). The two tokens were produced by the same adult speaker (A3). In the frame corresponding to the acoustic start of the /s/ noise, there is a sudden upward movement at the front part of the tongue (a region indicated by the arrow) when compared to the tongue contour in the frame before, suggesting that a constriction for /s/ is formed at the

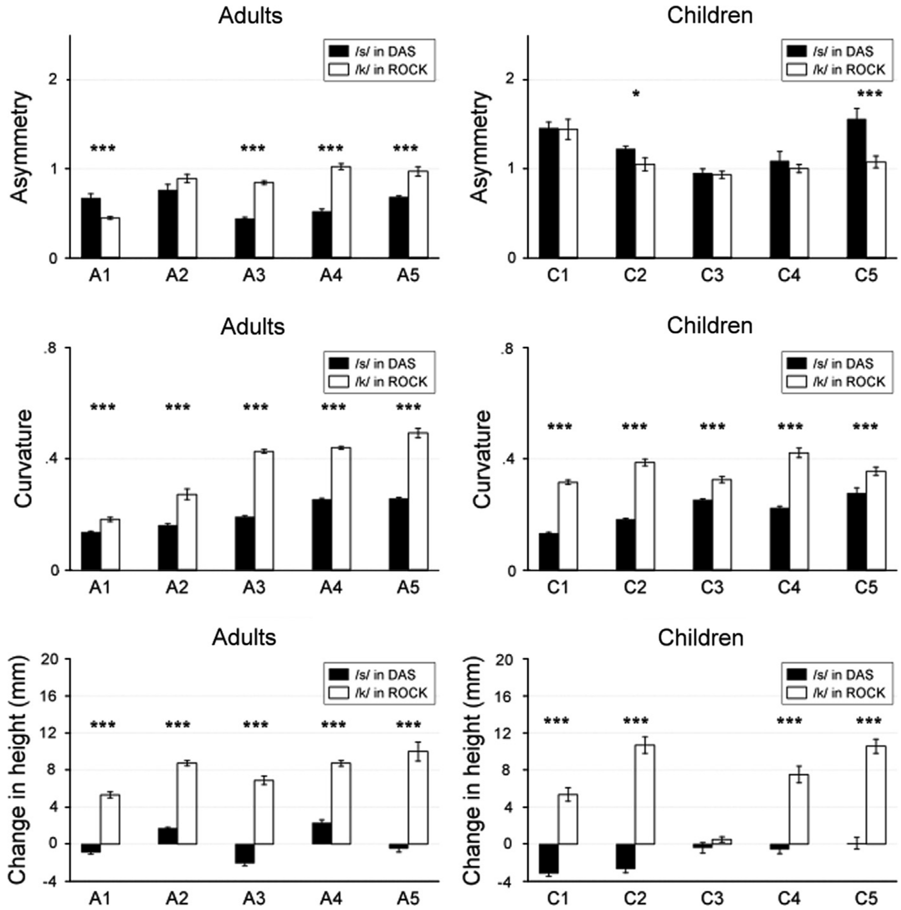

Fig. 4. Asymmetry (first row), curvature (middle row), and change in tongue height (bottom row) for /s/ noise and /k/ closure. Error bars represent standard error. (*Note*: $*=p<.05$, $**=p<.01$, $***=p<.001$).

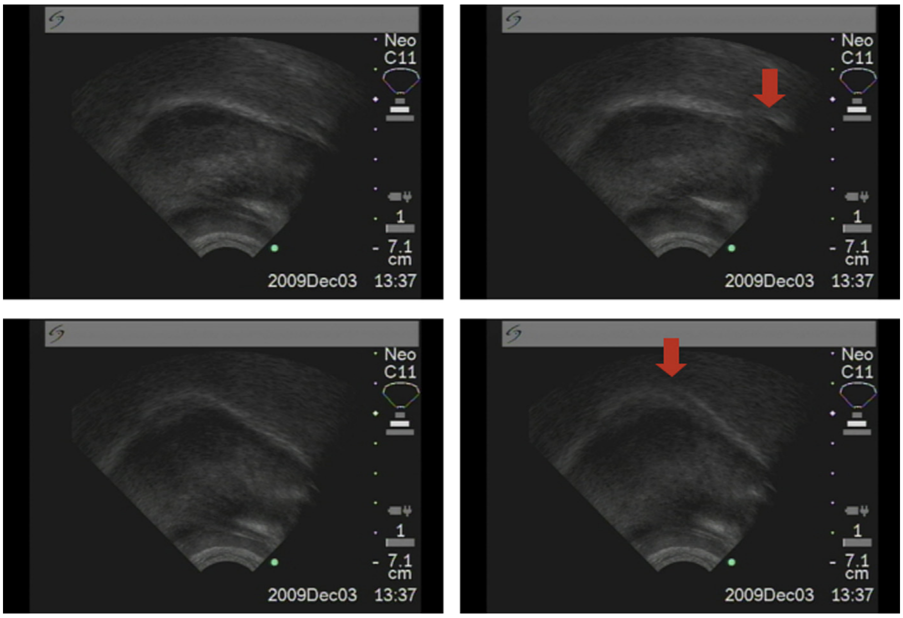

Fig. 5. Two consecutive frames extracted immediately before and at the onset of /s/ frication noise (top) and /k/ closure (bottom).

tongue tip region. On the other hand, at the beginning of /k/ closure the tongue shows a sudden upward movement around the middle part of the tongue contour (the region indicated by the arrow). This observation is in line with the fact that the location of the constriction for velar /k/ is more back than that of coronal /s/. Although this predicted difference in constriction location seemed not to be captured by the triangle that is used to derive the asymmetry value, that triangle-based analysis nevertheless revealed an interesting and unpredicted systematic difference in asymmetry between /s/ and /k/ in adult speech; that is, for most speakers (A3, A4, A5), the relative mass of the tongue for /k/ was more forward than for /s/ (i.e., the peak of the triangle fitted under the tongue contour is more forward for /k/). Only one adult subject (A1) had greater asymmetry for /s/ than /k/. As shown in Fig. 4, children overall had greater asymmetry values than adults, suggesting that the mass of the tongue is positioned more toward the front of the tongue in children than in adults. This finding appears to be consistent with the observation from Kent (1990) that the infant has a relatively anterior tongue mass compared to the adult. For three children (C1, C3, C4), the relative mass of the tongue was forward for both /s/ and /k/, and for two children (C2, C5), /s/ was more forward than /k/. Thus, none of the children showed greater asymmetry for /k/, like the majority of adults.

As expected, the curvature values were greater for /k/ than /s/ for both adults and children, meaning that the tongue was more bunched for /k/ (see Fig. 4). Likewise, for both adults and children, tongue height increase from the midpoint of the vowel to the consonant was greater when producing /k/. There was one child (C3) whose result was marginally not significant ($p = .056$), but she showed a difference in the right direction. To summarize, /s/ in *das* and /k/ in *rock* differed in all three measures for adults, with asymmetry revealing an unexpected type of difference between /s/ and /k/. Children overall showed adult-like patterns in curvature and change in tongue height, but many of them did not differentiate /k/ and /s/ in terms of the asymmetry of the tongue.

With these baseline results in mind, we now turn to the effects of segmental context and morphological complexity on the production of /s/ and /k/. Unlike curvature and change in height, asymmetry showed mixed results in the baseline comparison. As the mixed results on the asymmetry measures made it difficult to make specific predictions for /s/ and /k/ in monomorphemic and bimorphemic /ks/ clusters, we focused on the other two measures (curvature and change in height) in the subsequent analyses examining segmental context and morphological effects.

4.2. Coarticulation effect: monomorphemic items

4.2.1. Singleton /s/ in das [dɑs̲] vs. /s/ in the consonant cluster /ks/ in box [bɑks̲]

To examine how the articulation of /s/ is influenced by the preceding consonant /k/, we compared the articulatory gestures of /s/ in *das* [dɑs̲] vs. monomorphemic /s/ in *box* [bɑks̲] (see Fig. 6). Overall, adults' production of /s/ was affected by the preceding /k/. That is, the tongue was more bunched (curvature) and the increase in tongue height (change in height) was greater when /s/ was preceded by /k/ than when it appeared in a singleton coda. In general, most children (4/5) showed the same pattern as adults, with greater curvature and change in height for /s/ in the /ks/ cluster than in /s/ alone. Thus, for both the adults and children, the gestural characteristics of /s/ were generally modified by the preceding /k/ within the cluster, suggesting coarticulation effects. Despite this overall similarity, children overall exhibited greater variability than adults. For example, one child failed to show the difference in curvature (C5) and another child showed a significant difference in the change in height (C3) between singleton /s/ and /s/ in the /ks/ cluster in the opposite direction to adults.

4.2.2. Singleton /k/ in rock [ɹɑk̲] vs. /k/ in consonant cluster /ks/ in box [bɑk̲s]

To examine how the articulation of /k/ is influenced by the following consonant /s/, we compared /k/ in *rock* [ɹɑk̲] vs. /k/ in monomorphemic *box* [bɑk̲s] (see Fig. 7). Unlike the production of /s/, which was affected by the tongue shape and position of the preceding /k/ within the cluster, there was no clear evidence that the production of /k/ was affected by the following /s/. For adults, no clear differences were seen across participants for either of the two measures, suggesting there was little articulatory effect of the /s/ on the preceding /k/.

Similarly, children showed no clear differences in tongue height change between singleton coda /k/ and /k/ in the /ks/ cluster. However, for three children (C1, C2, C4), the tongue was more bunched when producing /k/ alone in *rock* than when producing /k/ in *box*, as indicated by larger curvature values for singleton /k/. This result suggests a coarticulation effect between /k/ and /s/ in the /ks/ cluster for *box* in these children. We now examine the /ks/ clusters to determine if there is any morphological effect.

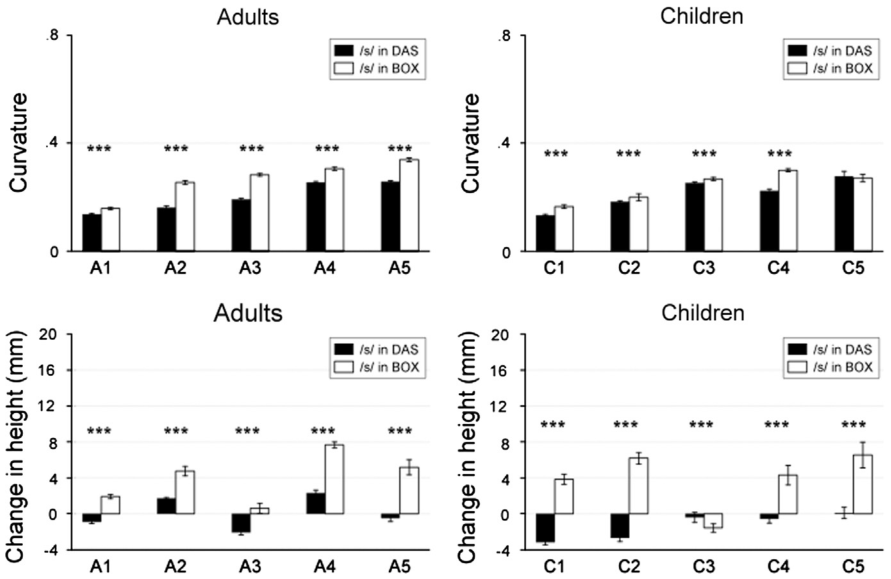

Fig. 6. Curvature (upper row) and change in tongue height (bottom row) for singleton coda /s/ vs. /s/ in consonant cluster /ks/. Error bars represent standard error. (*Note*: *=*p*<.05, **=*p*<.01, ***=*p*<.001).

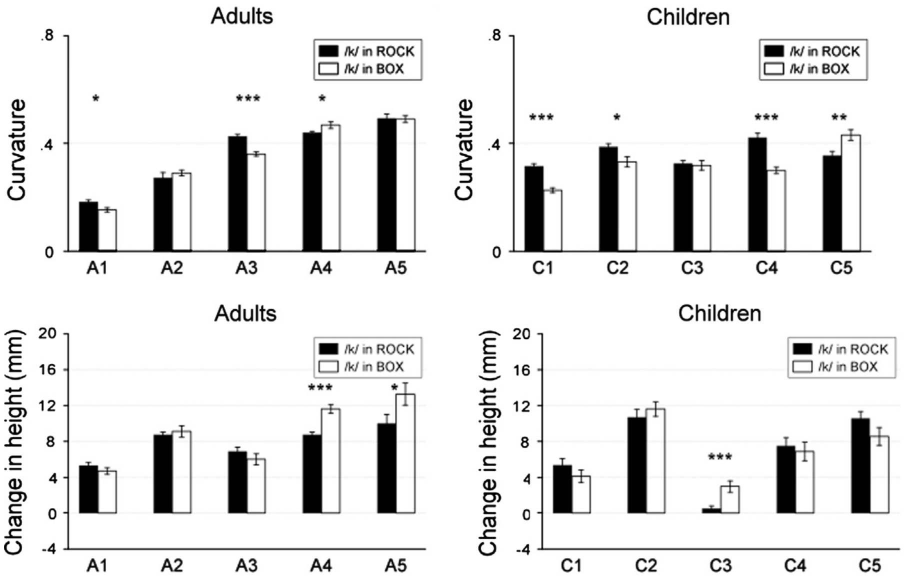

Fig. 7. Curvature (upper row) and change in tongue height (bottom row) for singleton coda /k/ vs. /k/ in consonant cluster /ks/. Error bars represent standard error. (*Note*: *=*p*<.05, **=*p*<.01, ***=*p*<.001).

4.3. Morphological effect

4.3.1. /s/ in monomorphemic box [bɑks̲] vs. bimorphemic rocks [ɹɑks̲]

As shown in Fig. 8, most children (C1, C2, C3, C4) had a more bunched tongue for the /s/ in *rocks* than for the /s/ in *box*, and three children (C1, C4, C5) demonstrated a height change that was greater for the /s/ in *box* than for the /s/ in *rocks*. Two adults also showed these patterns of tongue curvature (A1, A4) and height change (A2, A4). Thus, unlike the acoustic measures, these different articulatory measures begin to reveal some morphological differences in the production of coda /s/ in monomorphemic vs. bimorphemic words.

4.3.2. /k/ in monomorphemic box [bɑk̲s] vs. bimorphemic rocks [ɹɑk̲s]

Adults showed no consistent differences in curvature for /k/ as a function of morphological complexity of the coda cluster (see Fig. 9). However, three of the adults (A2, A4, A5) showed a greater increase in height for /k/ in monomorphemic /ks/. For most of the children (C1, C2, C3, C4), the tongue was more bunched for /k/ in *rocks*, suggesting less coarticulation with the /s/. Similar to the adults, the change in the tongue height was greater for /k/ in *box* for two of the children (C1, C4). This suggests that /ks/ in *box* is overall higher than /ks/ in *rocks* for these subjects. To confirm this possibility, we conducted an additional linear mixed-effects regression analysis of the tongue height change as a function of monomorphemic and bimorphemic /ks/ as a whole. The results showed that the 3 subjects in each group (60%) (A2, A4, A5 and C1, C4, C5) had bigger height change values for the /ks/ in *box* than for the /ks/ in *rocks*; one child (C3) showed a difference in the opposite direction, and two adults (A1, A3) and one child (C2) showed no difference in height change between the two clusters. This result demonstrates that the /ks/ in *box* was higher than the /ks/ in *rocks* for

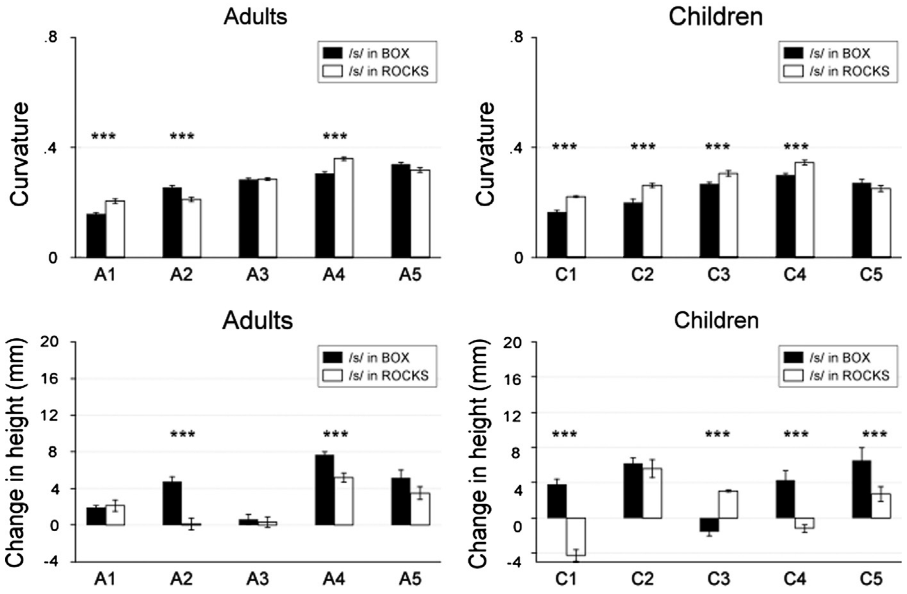

Fig. 8. Curvature (upper row) and change in tongue height (bottom row) for /s/ in monomorphemic coda vs. /s/ in bimorphemic coda. Error bars represent standard error. (*=p<.05, **=p<.01, ***=p<.001).

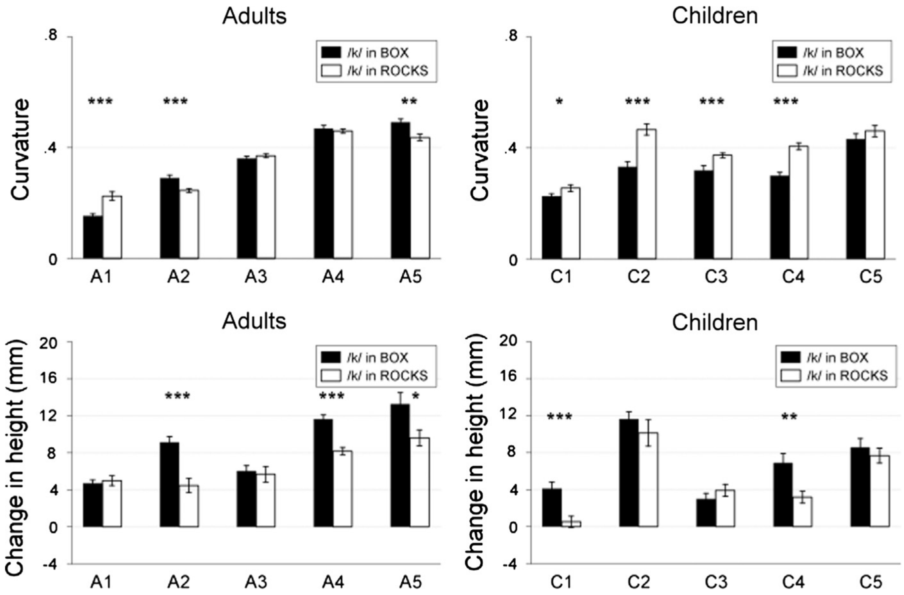

Fig. 9. Curvature (upper row) and change in tongue height (bottom row) for /k/ in monomorphemic coda vs. /k/ in bimorphemic coda. Error bars represent standard error. (*Note*: *=p<.05, **=p<.01, ***=p<.001).

the majority of the subjects. This is an interesting possibility because it suggests that the primary articulatory 'target' may be the /k/ in *box*, but the morphemic /s/ in *rocks*. This is explored further in Section 5.

Figs. 10 and 11 show representative examples of the change in highest point of the tongue from one of the adults (A4) and one of the children (C5), highlighting the major findings in the change in height. The graphs show how the height of the highest point of the tongue changes from the middle point of the vowel to the end of the word. As mentioned earlier, the highest point of the tongue need not be the same physical point on the tongue contour across frames. The x-axis shows the number of frames; zero is the middle point of the vowel, and frames are 33 ms apart. The y-axis shows the difference in height of the highest point of the tongue between the midpoint of the vowel and the subsequent individual frames. The triangles represent the frames for the vowel, i.e., between the midpoint and end of the vowel. The diamonds show the frames for a stop-closure-like silence before /s/ in the word *das*; this is discussed more in detail below. The circles show the frames corresponding to /k/ closure, i.e., from the /k/ closure to release. The open circles represent the frames corresponding to the post-release noise of /k/. The square markers indicate the frames corresponding to /s/ frication noise.

Several differences can be seen between the images of the different target words, and also between the adult and child. Consistent with the baseline results, the images show that, for both speakers, the tongue height increase from the vowel was greater when producing the /k/ in *rock* compared to the /s/ in *das*. For /k/ in the adult tokens, there was usually a sudden increase in height of the highest point of the tongue just before the moment of acoustic closure, and then it stayed at the same level until around the release. The height change of the highest point at the moment of acoustic release was not as abrupt as that at closure. In contrast, the child's tongue continued to go up during the /k/ closure, usually showing a couple of jumps in tongue height during the closure period. Similar behavior was observed for many tokens across all children in the study.

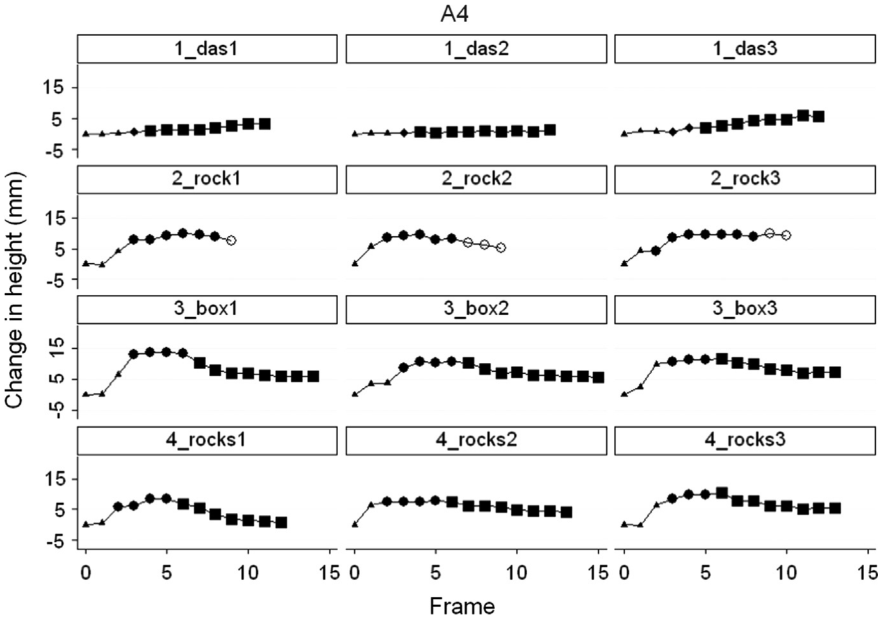

Fig. 10. Frame-by-frame change tongue height from the midpoint of the vowel to the subsequent individual frames in one of the adult subjects (A4). Legend for markers is as follows: triangles: vowel, diamonds: stop-closure-like silence before /s/ in *das*, circles: /k/ closure, open circles: post-release noise of /k/, squares: /s/ noise.

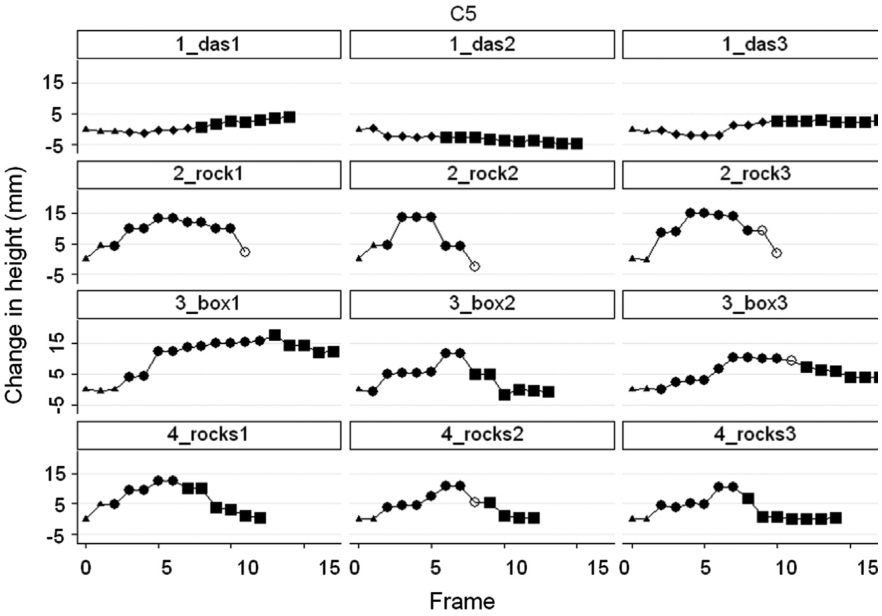

Fig. 11. Frame-by-frame change tongue height from the midpoint of the vowel to the subsequent individual frames in one of the child subjects (C5). Legend for markers is as follows: triangles: vowel, diamonds: stop-closure-like silence before /s/ in *das*, circles: /k/ closure, open circles: post-release noise of /k/, squares: /s/ noise.

In addition, as indicated by the diamond markers between the vowel and /s/ noise in *das*, both adults and children sometimes exhibited a stop-closure-like silence before /s/ noise in *das*. Interestingly, although the silence was not acoustically distinct from the silence in stop closures on the measures reported here, its articulation was clearly different from /k/ closure, in that there was no sudden increase in tongue height before or during the silence preceding /s/ noise in *das*. This suggests that the silence preceding /s/ might be just a period of voicelessness before the aerodynamic requirements are met for turbulence noise for the /s/. The frames during the silence in *das* are shown in Figs. 10 and 11, but the silent intervals were not considered as part of our acoustic measurements; they were also not included in any of the acoustic and articulatory analysis. Congruent with the coarticulation effect found for /s/ in the /ks/ cluster, it can be seen in the images that the height of the highest point is overall greater for /s/ in *box* than for /s/ in *das*. Finally, the images confirm that the height of the highest point is overall greater for /ks/ in *box* than for /ks/ in *rocks*. They also show the different height signatures for /k/ closure across *rock*, *box*, and *rocks*, for the adult, and especially for the child.

5. Discussion

It has long been known that consonantal gestures overlap within clusters (Hardcastle & Roach, 1979; Zsiga, 1994; Byrd, 1996). It has also been shown that identical sequences of segments may show different articulatory signatures across a morpheme boundary (Cho, 2002). Yet, little is known about how and when these aspects of gestural organization develop. This paper set out to explore the acoustics and gestural organization in both monomorphemic and bimorphemic clusters, examining coarticulatory effects in the speech productions of adults and 2-year-olds. We found that both adults and children showed strong coarticulatory effects of /k/ on the following /s/ in the lexical item *box* (although for children the acoustics were not as clear as the ultrasound data). The lack of effect of /s/ on /k/ in *box* is rather surprising, because anticipatory coarticulation is one of the most common types of coarticulation (Crystal, 2003). Furthermore, in Catalan consonant clusters, /s/, which involves tongue dorsum lowering, was shown to have greater coarticulatory resistance than /k/, which involves tongue dorsum raising (Recasens & Pallarès, 2001; Recasens, 2004). This finding is in agreement with the suggestion that some consonants have the greater resistance to coarticulatory effects than others, due to the differences in articulatory constraints. One possible explanation is that the primary articulatory target for this coda cluster is the /k/.

In contrast, morphemic /s/ in *rocks* exerted strong coarticulatory effects on lexical /k/; for the /s/ in *rocks* the tongue was lower than for the /s/ in *box*, for both adults and children. The /k/ in *rocks* showed the same effects, tending to be lower than the /k/ in *box*—for some adults, and even more so for children. Thus, the most important articulatory target for the bimorphemic cluster in *rocks* appeared to be the plural morpheme /s/. Although these results were somewhat unexpected, they are also intriguing, suggesting some type of morphological effect on these different items.

One possible explanation for the difference in /ks/ cluster articulation across the different morphemic conditions could be that there is a lexical frequency difference between the two target words *rocks* and *box*, and that this was realized as a difference in articulatory gestures. The effects of phonological reduction on high frequency lexical items have been well documented in the literature, often leading to processes of phonological change (e.g., Bybee, 2002). Reduction has also been found in a recent study exploring articulatory gestures using ultrasound, where articulatory gestures were less likely to meet their articulatory targets in high frequency words, at least for some speakers (Lin, Beddor, & Coetzee, 2011). Perhaps, then, *rocks* is more frequent, relaxing the requirement for a fully raised tongue body for the /k/. However, an examination of lexical frequencies in child-directed speech at 24 months reveals that *box* is substantially the more frequent: out of a 1 million word corpus from the CHILDES database (MacWhinney, 2000), *box* occurred 767 times compared to only 53 for *rocks*. Thus, although the relative frequency of the inflected plural is a good predictor of children's use of the morpheme (Zapf, 2004), it is not clear how a lexical frequency effect could account for the articulatory differences found in the present study.

Our current results, along with other studies, suggest semantics/information content might be differentially encoded in articulatory gestures. Recent reports have found that when children simplify a coda stop+s cluster to a singleton in utterance medial position, they tend to preserve the plural morpheme (with variable voicing) (e.g., *pigs* [pɪs]~[pɪz]) (Theodore et al., 2011). Perhaps, then, the primary articulatory target in plurals tends to be the plural morpheme itself. This is an interesting area for further research.

In sum, both adults and children in this study showed articulatory effects of coarticulation between the consonants within the coda cluster for both the monomorphemic and the bimorphemic lexical items. Interestingly, however, the directionality of the coarticulatory effects within the cluster differed as a function of morphological form: in the monomorphemic cluster, the first consonant appeared to be the primary articulatory target, whereas in the bimorphemic cluster, the second consonant/morpheme was the primary articulatory target. If this is so, then the gestural organization of clusters is strongly influenced by morphological form. If confirmed by further studies, this finding will provide important insight into the processes of articulatory planning in speech production (cf. Levelt, 1989).

That these effects were evidenced in 2-year-olds suggests that some aspects of articulatory timing, at least at the level of the word, are being acquired quite early. Although there is evidence that adults exhibit different articulatory patterns with respect to morphological boundaries, very little is known about how this develops in children, where both lexical and morphological representations may still be somewhat fragile. It is then perhaps not surprising that children might exert more articulatory effort than adults in producing inflectional morphemes, exhibiting morphological effects like those found in this study. Some support for this comes from studies indicating that children are more accurate in the production of words that are more frequently inflected (e.g., Zapf, 2004). This is an interesting area for further research.

Although these findings and implications are important, the small sample size of the present study does not allow firm conclusions to be drawn about phonological development processes in all children. Also, a considerable amount of individual variation was observed across speakers, especially with respect to the pattern of articulatory differences conditioned by the morphological structure of the coda. Therefore, the results found here should be considered tentative and hypothesis-generating at this stage, pending verification with more subjects. Further investigations are warranted as the ultrasound techniques for the study of language acquisition continues to develop.

6. Conclusions

The goal of the present study was to examine the acoustic and articulatory effects of segmental context and morphological complexity in the coda consonant productions of 2-year-olds and adults. In particular, we wanted to know if both children and adults would show the same evidence of coarticulation, and if both would exhibit morphological effects; in addition, we sought to determine whether ultrasound methods can be successfully used to study gestural development in children as young as age 2.

The answer to all these questions is yes. Both children and adults showed coarticulation effects within both monomorphemic and bimorphemic /ks/ coda clusters, and although this was less clear in the acoustic measures, where there was no support for previous claims that bimorphemic /s/ is longer than monomorphemic /s/, children and adults showed similar articulatory effects of coarticulation within the clusters, at least to some extent. This is evidenced by how the monomorphemic vs. bimorphemic clusters were produced, with the /ks/ in monomorphemic *box* articulated with much higher relative tongue height than the /ks/ in morphologically complex *rocks*. This suggests that some of the articulatory planning processes are different for these two items—especially for children, as shown by the fact that there were a greater number of children than adults who demonstrated these effects. These findings also demonstrate that ultrasound methods can be useful for exploring different aspects of language production in children as young as two—an age where little is known about the mechanisms of production. The results therefore suggest many directions for further research involving the nature of the speech planning processes, and how they develop.

Acknowledgments

We thank Elena Tenenbaum for recording the stimuli, Melanie Cabral, Karen Evans, Clara Kilman-Silver, Jill Thorson and Matt Vitorla for research assistance, and Suyong Song and Ivan Yuen for discussion and suggestions. This work was supported by NIH Grant R01HD057606 to Katherine Demuth and Stefanie Shattuck-Hufnagel.

Appendix A. Summary of ultrasound results

Note: Tables A1–A5 summarize predictions for each articulatory measure, and the number of subjects whose results were consistent with predictions (CP), opposite to predictions (OP), and not significant (NS). If over 50% of the total number in each group met the qualifications for any of the categories, the cells were highlighted in gray.

(1) *Baseline comparison*: The results showed that significant differences were found between /k/ and /s/ on all three measures. As predicted, the curvature and tongue height increase were greater for /k/ than for /s/. The asymmetry showed mixed results; the majority (3/5) of adults had greater asymmetry values for /k/ than for /s/, suggesting more anterior placement of the tongue mass for /k/. In contrast, for one adult and 2 children, /s/ showed greater asymmetry than /k/.

Table A1
/s/ in *das* vs. /k/ in *rock*.

Measures	Predictions	Adults				Children			
		CP	OP	NS	Sum	CP	OP	NS	Sum
Asymmetry	/s/>/k/	1	3	1	5	2	0	3	5
Curvature	/s/</k/	5	0	0	5	5	0	0	5
Height change	/s/</k/	5	0	0	5	4	0	1	5

(2) and (3) *Coarticulation effects*: /s/ in the /ks/ cluster exhibited articulatory characteristics of the adjacent /k/. In contrast, /k/ in the /ks/ cluster overall did not exhibit articulatory characteristics of adjacent /s/.

Table A2
/s/ in *das* vs. /s/ in *box*.

Measures	Predictions	Adults				Children			
		CP	OP	NS	Sum	CP	OP	NS	Sum
Curvature	/s/ in /ks/>/s/ alone	5	0	0	5	4	0	1	5
Height change	/s/ in /ks/>/s/ alone	5	0	0	5	4	1	0	5

Table A3
/k/ in *rock* vs. /k/ in *box*.

Measures	Predictions	Adults				Children			
		CP	OP	NS	Sum	CP	OP	NS	Sum
Curvature	/k/ alone>/k/ in /ks/	2	1	2	5	3	1	1	5
Height change	/k/ alone>/k/ in /ks/	0	2	3	5	0	1	4	5

(4) and (5) *Morphological effects*: Although the prediction that monomorphemic /ks/ would show a more robust coarticulation effect than bimorphemic /ks/ was not supported, the results revealed a couple of interesting differences between monomorphemic /ks/ and bimorphemic /ks/: First, the increase in tongue height was greater for the /ks/ in *box*. Second, for most children, the curvature was greater for /ks/ in *rocks*.

Table A4
/s/ in *box* vs. /s/ in *rocks*.

Measures	Predictions	Adults				Children			
		CP	OP	NS	Sum	CP	OP	NS	Sum
Curvature	/s/ in *box*>/s/ in *rocks*	1	2	2	5	0	4	1	5
Height change	/s/ in *box*>/s/ in *rocks*	2	0	3	5	3	1	1	5

Table A5
/k/ in *box* vs. /k/ in *rocks*.

Measures	Predictions	Adults				Children			
		CP	OP	NS	Sum	CP	OP	NS	Sum
Curvature	/k/ in *box* </k/ in *rocks*	1	2	2	5	4	0	1	5
Height change	/k/ in *box* </k/ in *rocks*	0	3	2	5	0	2	3	5

References

Aubin, J., & Ménard, L. (2006). Compensation for a labial perturbation: An acoustic and articulatory study of child and adult French speakers. In *Proceedings of the seventh international seminar on speech production* (pp. 209–216). Ubatuba, Brazil.

Baayen, R. H. (2008). *Analyzing linguistic data: A practical introduction to statistics*. Cambridge: Cambridge University Press.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*, 390–412.

Bernhardt, B., Gick, B., Bacsfalvi, P., & Ashdown, J. (2003). Speech habilitation of hard of hearing adolescents using electropalatography and ultrasound as evaluated by trained listeners. *Clinical Linguistics and Phonetics*, *17*, 199–217.

Boersma, P., & Weenink, D. (2005). PRAAT: Doing phonetics by computer (Version 4.4.07) [Computer program]. Available from 〈http://www.praat.org/〉.

Browman, C. P., & Goldstein, L. (1986). Towards an articulatory phonology. *Phonology Yearbook*, *3*, 219–252.

Browman, C. P., & Goldstein, L. (1988). Some notes on syllable structure in articulatory phonology. *Phonetica*, *45*, 140–155.

Browman, C. P., & Goldstein, L. (1990). Tiers in articulatory phonology, with some implications for casual speech. In: T. Kingston, & M. E. Beckman (Eds.), *Papers in laboratory phonology I: Between the grammar and physics of speech* (pp. 341–376). Cambridge, England: Cambridge University Press.

Brown, R. (1973). *A first language: The early stages*. Cambridge, MA: Harvard University Press.

Bybee, J. (2002). Word frequency and context of use in the lexical diffusion of phonetically conditioned sound change. *Language Variation and Change*, *14*, 261–290.

Byrd, D. (1996). Influences on articulatory timing in consonantal sequences. *Journal of Phonetics*, *4*, 59–74.

Cheng, H. Y., Murdoch, B. E., Goozée, J. V., & Scott, D. (2007). Electropalatographic assessment of tongue-to-palate contact patterns and variability in children, adolescents, and adults. *Journal of Speech, Language, and Hearing Research*, *50*, 375–392.

Cho, T. (2002). Effects of morpheme boundaries on intergestural timing: Evidence from Korean. *Phonetica*, *58*, 129–162.

Crystal, D. (2003). *A dictionary of linguistics and phonetics* (5th ed.). Oxford: Blackwell.

Dale, P. S., & Fenson, L. (1996). Lexical development norms for young children. *Behavioral Research Methods, Instruments, and Computers*, *28*, 125–127.

Davidson, L. (2005). Addressing phonological questions with ultrasound. *Clinical Linguistics and Phonetics*, *19*, 619–633.

Demuth, K., Culbertson, J., & Alter, J. (2006). Word-minimality, epenthesis, and coda licensing in the acquisition of English. *Language and Speech*, *49*, 137–174.

Fee, J. (1995). Segments and syllables in early language acquisition. In: J. Archbald (Ed.), *Phonological acquisition and phonological theory* (pp. 43–61). Hillsdale, NJ: Lawrence Erlbaum.

Gerken, L. A. (1996). Prosodic structure in young children's language production. *Language*, *72*, 683–712.

Gick, B., Bacsfalvi, P., Bernhardt, B. M., Oh, S., Stolar, S., & Wilson, I. (2008). A motor differentiation model for liquid substitutions in children's speech. *Proceedings of Meetings on Acoustics*, *1*(060003), 1–9.

Gick, B., Campbell, F., Oh, S., & Tamburri-Watt, L. (2006). Toward universals in the gestural organization of syllables: A cross-linguistic study of liquids. *Journal of Phonetics*, *34*, 49–72.

Gick, B., Michelson, K., & Radanov, B. (2006). Articulation without acoustics: A case study of Oneida utterance-final forms. *Journal of the Acoustical Society of America*, *120*, 3377.

Hardcastle, W. J. (1985). Some phonetic and syntactic constraints on lingual coarticulation during /kl/ sequences. *Speech Communication*, *4*(1-3), 247–263.

Hardcastle, W. J., & Roach, P. (1979). An instrumental investigation of coarticulation in stop consonant sequences. In: H. Hollien, & P. Hollien (Eds.), *Current issues in the phonetic sciences* (pp. 531–540). Amsterdam: John Benjamins.

Hodson, B., & Paden, E. (1991). *Targeting intelligible speech: A phonological approach to remediation* (2nd ed.). Austin: TX: Pro-Ed.

Imbrie, A. K. K. (2005). *Acoustical study of the development of stop consonants in children*. Cambridge, MA.: Harvard-MIT Division of Health Sciences and Technology, Massachusetts Institute of Technology [Ph.D. thesis].

Johnson, K. (2008). *Quantitative methods in linguistics*. Oxford: Blackwell.

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

Kent, R. D. (1990). The biology of phonological development. In: C. A. Ferguson, L. Menn, & C. Stoel-Gammon (Eds.), *Phonological development: Models, research, implications* (pp. 65–90). Timonium, MD: York Press.

Kirk, C., & Demuth, K. (2005). Asymmetries in the acquisition of word-initial and word-final consonant clusters. *Journal of Child Language*, *32*, 709–734.

Klatt, D. H. (1976). Linguistic uses of segmental duration in English: Acoustic and perceptual evidence. *Journal of the Acoustical Society of America*, *59*, 1208–1221.

Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. *Cambridge, MA*. MIT Press.

Li, M., Kambhamettu, C., & Stone, M. (2003). EdgeTrak: A program for band-edge extraction and its applications. Paper presented at the sixth IASTED international conference on computers, graphics and imaging, Honolulu, HI, August 13–15.

Lin, S., Beddor, P.S., & Coetzee, A.W. (2011). Gestural reduction and sound change: an ultrasound study. *Proceedings of the 17th International Congress of Phonetic Sciences* (pp. 1250–1253). Hong Kong.

Macken, M., & Barton, D. (1980). The acquisition of the voicing contrast in English: A study of voice onset time in word-initial stop consonants. *Journal of Child Language*, *7*, 41–74.

MacWhinney, B. (2000). *The CHILDES project*. Mahwah, NJ: Erlbaum.

Marshall, C., & van der Lely, H. (2007). The impact of phonological complexity on past tense inflection in children with Grammatical-SLI. *Advances in Speech Language Pathology*, *9*, 191–203.

Ménard, L., Aubin, J., Thibeault, M., & Richard, G. (2012). Comparing tongue shapes and positions with ultrasound imaging: A validation experiment using an articulatory model. *Folia Phoniatrica et Logopaedica*, *64*, 64–72.

Ménard, L., & Noiray, A. (2011). The development of lingual gestures in speech: Comparing synthesized vocal tracts with natural vowels. *Faits de langue*, *37*, 189–202.

Ménard, L., Toupin, C., Thibeault, M., Noiray, A., Giroux, M., & Rousseau, M.C. (2010). A study of VCV coarticulation in 4-year-old children and adults, *Poster presented at the Ultrafest V conference, Haskins Laboratory.* New Haven, CT.

Miller, A., & Finch, K. (2011). Corrected high frame rate anchored ultrasound with software alignment. *Journal of Speech, Language, and Hearing Research*, *54*, 471–486.

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 and Rearing Research*, *32*, 120–132.

Vienna, Austria: R foundation for statistical computing〈http://www.R-project.org〉[Computer program]. Available from.

Recasens, D. (2004). The effect of syllable position on consonant reduction (evidence from Catalan consonant clusters). *Journal of Phonetics*, *32*(3), 435–453.

Recasens, D., & Pallarès, M. D. (2001). Coarticulation, assimilation and blending in Catalan consonant clusters. *Journal of Phonetics*, *29*273–301.

Repp, B. H., & Mann, V. A. (1982). Fricative-stop coarticulation: Acoustic and perceptual evidence. *Journal of the Acoustical Society of America*, *71*, 1562–1567.

Scobbie, J. M., Gibbon, F., Hardcastle, W. J., & Fletcher, P. (2000). Covert contrast as a stage in the acquisition of phonetics and phonology. In: M. B. Broe, & J. B. Pierrehumbert (Eds.), *Papers in laboratory phonology V: Acquisition and the lexicon* (pp. 194–207). Cambridge, United Kingdom: Cambridge University Press.

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

Smith, B. L. (1978). Temporal aspects of English speech production:A developmental perspective. *Journal of Phonetics*, *6*, 37–67.

Song, J. Y., Sundara, M., & Demuth, K. (2009). Phonological constraints on children's production of English third person singular –s. *Journal of Speech, Language, and Hearing Research*, *52*, 623–642.

Stevens, K. N. (2002). Toward a model for lexical access based on acoustic landmarks and distinctive features. *Journal of the Acoustical Society of America*, *111*(4), 1872–1891.

Stone, M. (2005). A guide to analysing tongue motion from ultrasound images. *Clinical Linguistics and Phonetics*, *19*(6/7), 455–501.

Theodore, R., Demuth, K., & Shattuck-Hufnagel, S. (2011). Acoustic evidence for position and complexity effects on children's production of plural –s. *Journal of Speech, Language, and Hearing Research*, *54*, 539–548.

Tiede, M., Perkell, J., Zandipour, M., & Matthies, M. (2001). Gestural timing effects in the "perfect memory" sequence observed under three rates by electromagnetometry. *Journal of the Acoustical Society of America*, *110*, 2657.

Tiede, M., Shattuck-Hufnagel, S., Johnson, B., Ghosh, S., Matthies, M., Zandipour, M., et al. (2007). Gestural phrasing in /kt/ sequences contrasting within and across word contexts. *Proceedings of the 16th International Congress of Phonetic Sciences, Saarbrü*cken (521–524), Germany.

Turk, A., Nakai, S., & Sugahara, M. (2006). Acoustic segment durations in prosodic research: A practical guide. In: S. Sudahoff, D. Lenertova, R. Meyer, S. Pappert, P. Augurzsky, I. Mleinek, N. Richter, & J. Schliesser (Eds.), *Methods in empirical prosody research* (pp. 1–28). Berlin: de Gruyter.

Turnbaugh, K., Hoffman, P. R., Daniloff, R. G., & Absher, R. (1985). Stop-vowel coarticulation in 3-year-old, 5-year-old, and adult speakers. *Journal of the Acoustical Society of America*, *77*, 1256–1257.

Walsh, T., & Parker, F. (1983). The duration of morphemic and non-morphemic /s/ in English. *Journal of Phonetics*, *11*, 201–206.

Wexler, K. (1994). Optional infinitives, head movement and the economy of derivations in child grammar. In: D. Lightfoot, & N. Hornstein (Eds.), *Verb movement* (pp. 305–350). Cambridge, MA: Cambridge University Press.

White, K. S., & Morgan, J. L. (2008). Sub-segmental detail in early lexical representations. *Journal of Memory and Language*, *59*, 114–132.

Zapf, J. (2004). Frequency in the input and children's mastery of the regular English plural. In A. Brugos, L. Micciulla, & C. E. Smith (Eds.), *Proceedings of the 28th Annual Boston University Conference on Language Development* (pp. 669–680). Somerville, MA: Cascadilla.

Zharkova, N., Hewlett, N., & Hardcastle, W. J. (2011). Coarticulation as an indicator of speech motor control development in children: An ultrasound study. *Motor Control*, *15*, 118–140.

Zharkova, N., Hewlett, N., & Hardcastle, W. J. (2012). An ultrasound study of lingual coarticulation in /sV/ syllables produced by adults and typically developing children. *Journal of the International Phonetic Association*, *42*(2), 193–208.

Zsiga, L. (1994). Acoustic evidence for gestural overlap in consonant sequences. *Journal of Phonetics*, *22*, 121–140.