View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Queen Margaret University eResearch

This article has been published in the Journal of Speech, Language, and Hearing Research. Below is a final personal version of the article, incorporating the changes made during the peer review process. The citation is as follows: Zharkova, N., Hewlett, N., Hardcastle, W.J. & Lickley, R. (2014). Spatial and temporal lingual coarticulation and motor control in preadolescents. *Journal of Speech, Language, and Hearing Research*, **57**, 374-388. doi:10.1044/2014_JSLHR-S-11-0350. The link to the published version: http://jslhr.pubs.asha.org/Article.aspx?articleid=1832512.

Spatial and temporal lingual coarticulation and motor control in preadolescents

Natalia Zharkova, Nigel Hewlett, William J. Hardcastle, Robin J. Lickley

Clinical Audiology, Speech and Language Research Centre

Queen Margaret University

Correspondence to:

Natalia Zharkova

Clinical Audiology, Speech and Language Research Centre

Queen Margaret University

Queen Margaret University Drive

Musselburgh EH21 6UU

East Lothian

United Kingdom

nzharkova@qmu.ac.uk

Abstract

Purpose: Coarticulation and lingual kinematics were compared in preadolescents and adults, in order to establish whether preadolescents had a greater degree of random variability in tongue posture and whether their patterns of lingual coarticulation differed from those of adults.

Method: High-speed ultrasound tongue contour data synchronised with the acoustic signal were recorded from 15 children aged between 10 and 12 years old, and 15 adults. Tongue shape contours were analysed at nine normalised time-points during the fricative phase of schwa-fricative-/a/ and schwa-fricative-/i/ sequences with the consonants /s/ and /ʃ/.

Results: There was no significant age-related difference in random variability. Where a significant vowel effect occurred, the amount of coarticulation was similar in the two groups. However, the onset of the coarticulatory effect on preadolescent /ʃ/ was significantly later than on preadolescent /s/, and also later than on adult /s/ and /ʃ/.

Conclusions: Preadolescents have adult-like precision of tongue control and adult-like anticipatory lingual coarticulation with respect to spatial characteristics of tongue posture. However, there remains some immaturity in the motor programming of certain complex tongue movements.

Introduction

The study reported here investigated variability of tongue posture and lingual anticipatory coarticulation in CV syllables, comparing the performance of a group of adult speakers and a group of children of 10-12 years of age, using high frame-rate ultrasound recording. Throughout

this article, we will use the term "preadolescent" to refer to the younger age group in this study, i.e., 10-12-year-old children, in order to avoid any confusion when discussing the earlier development of lingual motor control and coarticulatory patterns.

Random variability in tongue position

A reduction in random variability has been found to be an indicator of development towards adult-like speech motor control. The findings which link variability and age are most extensive and robust for acoustic measures of segment duration (e.g., Kent & Forner, 1980; Sharkey & Folkins, 1985; Lee et al., 1999) but variability in articulator movement has also been reported to decrease with age[1]. This has been demonstrated in direct articulatory analysis and also inferred from the acoustic information (e.g., Smith & Goffman, 1998; Green et al., 2000; 2002; Walsh & Smith, 2002; Smith & Zelaznik, 2004; Nittrouer et al., 2005; Grigos, 2009; Zharkova et al., 2011; Murdoch et al., 2011). The development of tongue control in speech is thought to be more protracted than that of the jaw and the lips (e.g., Kent, 1992; Nittrouer, 1993; though Nittrouer et al., 2005, suggested that the temporal coordination of jaw movements and phonation may not be fully mastered by 7 years old; see also Smith & Zelaznik, 2004, who showed that the level of consistency in coordinating movements of the jaw and the lips is not yet adult-like even in adolescents). In a recent electromagnetic articulography (EMA) study comparing stability of interarticulator coordination in 6-9-year-old typically developing children and two groups with developmental speech disorders (Terband et al., 2011), the control over the tongue tip has been shown to develop more slowly than control over the lips in both typical and disordered speech.

[1] Here we only refer to the changes in variability from childhood to adulthood, as opposed to changes in variability throughout the lifespan (cf. Stathopoulos et al. 2011).

Typically developing children aged between 6 and 9 years old have been shown by Zharkova et al. (2011; 2012) to have a greater variability of tongue position than adults, in the fricatives /s/ and /ʃ/. Variability of formant patterns, which is certain to be at least partly attributable to variability of tongue placement, has been shown to reach adult levels after the age of 14 years (Lee et al., 1999). Tongue placement variability is also reflected in the data from tongue-to-jaw coupling. Cheng et al. (2007b), in an EMA study of tongue and jaw movements in children aged 6-7 and 8-11 years old, in 12-16-year-olds, and in adults, showed that the 6-7-year-old group had a weaker coupling of the tongue tip and the jaw (as evidenced by the difference in the time of achieving maximum speed by these two articulators during the production of /t/) than all the other age groups. The 6-7-year-olds were also significantly less consistent than adults in the coupling of tongue and jaw movement for /k/. No significant differences between older child groups and adults were reported. However the authors noted a trend of increasing temporal coupling of the tongue tip and the jaw during /t/ with age, as well as a trend of increasing consistency of tongue-to-jaw spatiotemporal coupling patterns in /k/ with age. Reduction in variability of articulator movement with age was reported in Murdoch et al. (2011), for the lips, the tongue tip and tongue body. In their EMA study, 6-7-year-olds were found to be significantly more variable than both 12-17-year-olds and adults. Additionally, the 8-11-year-old group in Murdoch et al. (2011) was significantly more variable than adults. In the present study, we tested the hypothesis that the preadolescents, like younger children, would have significantly more token-to-token variability in tongue position than adults.

Development of coarticulation

In developmental research, a central question concerns the units of speech production that children acquire and then employ when they build up utterances. A number of theories propose a variety of forms and sizes of stored phonetic plans, learned in infancy, which form the basis for the later development of speech production skills (Locke, 1983; Davis & McNeilage, 1995; Guenther & Perkell, 2004; Serkhane et al., 2007; Vihman et al., 2009; Nam et al., 2013; also, for reviews, see Vihman, 1996; Smith & Goffman, 2004; Ziegler & Maassen, 2004). Research comparing coarticulation in children and adults has largely focused on whether motor programs for speech are segment-based or syllable-based, with the predominance of evidence pointing to the syllable or even larger units being favoured by young children. Evidence from acoustical analysis of formant movement in the vowel in the word "box" has been interpreted as support for the segmental view (Kent, 1983). In Kent's study, adult speech displayed a rise in F2 during the vowel, while four-year-old children's F2 remained rather flat throughout the vowel. It was concluded that children had a more 'segmental' approach to speech than adults, since they appeared not to display the same coarticulatory effects. Evidence contrary to this conclusion was suggested by Nittrouer and colleagues (e.g., Nittrouer et al., 1989; 1996; for comparable results, see also Siren & Wilcox, 1995; Nijland et al., 2002). Their experiments examined F2 within fricative consonants (30 ms before vowel onset) and pairs of contrasting vowels. Their results showed greater coarticulation at this point in 3, 5 and 7 year olds than in adults. They explained the apparent discrepancy between Kent's (1983) results and their own by suggesting that the tongue may have been already raised before the start of the vowel in "box" and that coarticulation was therefore already present. On the basis of their evidence, Nittrouer et al. proposed that young children use syllables rather than segments as early units of speech production. Under their interpretation, young children initially display a greater degree of

coarticulation than adults, and this decreases as production becomes more segment-based in adulthood.

Despite disagreements in the literature concerning differential amounts of lingual coarticulation in children and adults, findings from various studies, including those using articulatory methodologies, generally agree that some adaptation of the tongue position in the consonant to that of the following vowel is present from an early age. For example, in Sussman et al. (1992), coarticulatory effects were demonstrated in 3- to 5-year-old children, using locus equations (i.e., straight lines characterising the extent of anticipatory vowel-on-consonant coarticulation, with F2 values taken from mid-vowel plotted along the *x* axis, and F2 values at the vowel onset plotted along the *y* axis). Similar findings for 4- to 5-year-olds were reported in Noiray et al. (2013), using locus equations calculated on both acoustic and ultrasound tongue imaging data. In Katz et al.'s (1991) study, however, 3-year-olds failed to demonstrate a measurable coarticulatory effect. The fact that some anticipatory lingual coarticulation within a CV syllable has been reported in children since at least the age of 4 years old suggests that children, like adults, can plan and program (see Van der Merwe, 2009, for a description of planning and programming) the CV as one "chunk" (in the sense of, e.g., Klapp, 2003). In fact, there is evidence from a study of labial coarticulation that 4-5-year-old children can anticipate lip rounding as far as several segments in advance of the rounded vowel (Goffman et al., 2008). However it has also been shown that coarticulation in children can be present or not, depending on the segment. Such was our finding in a previous experiment using ultrasound analyses of tongue movements. Children aged 6-9 years displayed greater lingual coarticulatory effects than adults in /ʃ/+vowel syllables, but only adults showed a coarticulatory effect with /s/+vowel (Zharkova et al., 2011; 2012). It may be that coarticulation development depends partly on the

articulatory characteristics of a sound or sound sequence, rather than simply the number of sounds in a chunk. This possibility is pursued in the following subsection in relation to /s/ and /ʃ/.

Coarticulation of /s/ and /ʃ/

The consonant-specific findings reported in Zharkova et al. (2011; 2012) might be explained with reference to development of the articulatory musculature. Less flexibility in tongue movement than in typical adult speech is characteristic of developmental disordered speech, and, to a certain extent, of typical speech development (Gibbon, 1999; Cheng et al., 2007a; Gick et al., 2008). For example, in a study using electropalatography (EPG), the coordination of tongue tip and tongue back in the consonant cluster /kl/ was found to be underdeveloped in children aged 6-7 years old (Cheng et al., 2007a), and its refinement continued in the 8-11-year-old group and in the 12-17-year-old group. It is possible that 6-9 year old children from the studies described in Zharkova et al. (2011; 2012) have not yet gained adult-like ability to vary the position of the tongue dorsum (relatively) independently of the tongue tip, the position of the latter being crucial of course in the production of /s/. It would consequently have been impossible for the children to adapt the tongue position for /s/ to that of the upcoming vowel without compromising the identity of /s/, as coarticulation is constrained by the need to successfully convey the identity of the sound to a listener. In accordance with the findings of immaturity in the speech of children up to the age of 14 years (Lee et al., 1999, described above), in the present study we expected more coarticulation in preadolescents than in adults for /ʃ/, and less for /s/.

The differential ability of speech sounds to adapt towards an upcoming or preceding segment, due to their different articulatory characteristics, is at the centre of the Degree of Articulatory Constraint (DAC) model of lingual coarticulation (e.g., Recasens et al., 1997). The concept of DAC is closely related to that of coarticulation resistance, cf. Bladon & Al-Bamerni, 1976; we have adopted the term coarticulation resistance in this report. The postalveolar fricative /ʃ/ is claimed to have greater coarticulation resistance than the alveolar fricative /s/ (Recasens & Espinosa, 2009; see also Pouplier et al., 2011; Niebuhr & Meunier, 2011; Niebuhr et al., 2011) because its primary articulation involves the comparatively massive tongue dorsum. However, the allegedly greater possibilities for adaptation to an adjacent segment by /s/ depend on the ability to control different parts of the tongue to some extent independently. The fricatives /s/ and /ʃ/ were chosen as target consonants for the present study partly to explore the role of these factors in speech development and partly to facilitate comparison with previous studies. In the present study, we wished to confirm, using tongue surface contour data, that syllable-initial /s/ does indeed have a lower coarticulation resistance than /ʃ/, in adult speech (cf. Recasens & Espinosa, 2009), by establishing that /s/ exhibits a greater adaptation to a following vowel than does /ʃ/, during the consonant. For the preadolescents, we predicted a similar pattern of outcomes with respect to coarticulation of /s/ and /ʃ/ as we found in our previous study of younger children, namely less tongue adaptation during /s/ than during /ʃ/.

The present study included determination of the time-course of the coarticulatory effect on the fricative. Since Recasens et al. (1997), in a study of coarticulation in VCV syllables, concluded that the temporal and spatial domains of coarticulation were closely and positively related, we predicted analogous outcomes for spatial and temporal measures of our data (with

earlier versus later onset of coarticulation corresponding to a larger versus smaller spatial coarticulatory effect, respectively).

Summary of hypotheses

H1. There is greater random variability in tongue position in preadolescents than in adults.

H2. For /ʃ/, the following vowel affects the tongue posture more in preadolescents than in adults.

H3. For /s/, the following vowel affects the tongue posture less in preadolescents than in adults.

H4. In adults, the following vowel affects the tongue posture of /s/ more than that of /ʃ/.

H5. In preadolescents, the following vowel affects the tongue posture of /s/ less than that of /ʃ/.

Method

Participants, stimuli and data collection

Participants consisted of 15 typically developing preadolescents (six female and nine male) and 15 adults (12 female and three male). All of them were native speakers of Scottish Standard English (Scobbie et al., 2007). The mean age for the preadolescents was 11;2 ([years;months]) and the age range was between 10;0 and 12;4. The mean age of the adults was 37 years and the range was between 18 and 58 years[2]. Ages of all participants are presented in Table 1.

[2] Most participants were aged between 21 and 48 years old, with two people above 50 years old and one person under 20 years old. In order to address possible concerns about age range of the adults, an investigative sensitivity analysis was carried out on the data. Linear mixed models were run (see subsection "Comparing size of coarticulatory effect, across consonant and across age group" in Method), excluding the data from the two oldest adults and one youngest adult, and the significant vowel-related difference in tongue shape was present for both consonants already at the consonant onset. Separate analyses were also conducted for the younger speakers (up to 30 years old). The significant vowel-related difference in tongue shape was present already at the consonant onset, for both consonants, in the three youngest participants, as well as in the five youngest, the six youngest and the seven youngest

The stimuli were the syllables /si/, /sa/, /ʃi/ and /ʃa/, embedded in the carrier phrase "It's a … Pam". The coarticulatory effects reported for our previous study (Zharkova et al., 2011; 2012) applied to /i/ as contrasted with /a/ and to /u/ as contrasted with /a/. There was no vowel effect for /i/ versus /u/, in either the adults or the children, for either fricative. This was probably because the tongue positions for /i/ and /u/ are quite similar in Scottish English (Scobbie et al., 2007; Zharkova & Lickley, 2010), /u/ being somewhat fronted in this accent. Therefore, in the study reported below, only the vowels /i/ and /a/ were used. The target syllables were spelled as "sea", "Sah", "she" and "shah". The sentences were shown to the participants on a computer screen, accompanied by images corresponding to the target words (the syllable /sa/ was introduced as the name of an imaginary creature; this nonsense word was combined with real words in order to facilitate comparison of the results with findings from previous studies, including our own). Every target was repeated six times, and the order of presentation was randomised. The total number of tokens recorded was 720 (4 syllable types x 6 repetitions x 30 participants). The Queen Margaret University high speed ultrasound system (Wrench & Scobbie, 2008) was used for recording tongue movements in the midsagittal plane, via a transducer located beneath the lower jaw. The acoustic signal was synchronised with the ultrasound recording. The frame rate for ultrasound was 100 Hz. This frame rate allowed us to carry out a detailed analysis of changes in the tongue surface contour over time. Participants wore a headset for stabilising the transducer in relation to the head (Articulate Instruments Ltd, 2008; Scobbie et al., 2008). They were seated in a sound-treated studio, isolated from any noise-emitting

---

speakers. In an analysis including the two oldest participants, a significant vowel-related difference in tongue shapes was also present at the consonant onset. In summary, the values of the youngest and the oldest subgroups of participants did not stand out from the values of the group as a whole.

instrumentation. Ethical approval for the research was obtained using standard procedures at Queen Margaret University.

<Table 1 about here>

Annotation and tongue curve fitting

Annotation and tongue tracing from ultrasound images were carried out in Articulate Assistant Advanced (Articulate Instruments Ltd, 2010). For every token by each speaker, the onset and offset of the target consonant were annotated from the acoustic record. The consonant onset was identified as the onset of the frication noise on the waveform. If preaspiration of the consonant was present in the preadolescents (it is a relatively common phenomenon in Scottish English speaking children; cf. Gordeeva, 2005), it was not counted as part of the consonant. Preaspiration was identified initially by auditory detection. In the acoustic recording, preaspiration was revealed by an interval of noticeably lower amplitude at the beginning of the noise phase. The consonant onset in such a case was marked at the abrupt increase of the noise amplitude on the acoustic waveform. The consonant offset was determined as the end of the frication noise, which coincided with the onset of the periodic waveform of the following vowel.

In each consonant token, tongue curves were traced (using an edge detection algorithm applied along each of 42 equally spaced radial axes superimposed on the imaged space) for every ultrasound frame, at 10 ms intervals, between onset and offset of the consonant. Each curve was examined to check the fit between the imaged contour and the superimposed curve, and

manually corrected where necessary[3]. Manual correction of the fit of the curve was required in most adult tokens, however some part of the curve was often traced accurately by the software, and in those cases only the gaps in the automatically traced curve were traced by hand. Less correction was required for preadolescent than for adult tongue contours, because the preadolescents generally had a brighter image of the tongue surface, due to the shorter distance from the transducer to the tongue surface, i.e., due to smaller heads. During visual checking of each curve, tongue surface contours in neighbouring ultrasound frames were consulted to facilitate the decision on whether corrections were necessary.

The software then exported each curve as a sequence of x-y values, based on the 42 coordinates from the radial axes, with two additional interpolated values between each pair of radial axes. The number of x-y values per curve could potentially reach 124 (including the original 42 values and the interpolated values), though in practice it was always lower, because the radial grid, which was the same for every speaker, was wider than the imaged tongue contour. The tongue contours corresponding to the onset and offset of the consonant contained articulatory information collected up to 9 ms before the annotated consonant onset, or up to 9 ms after the annotated consonant offset, respectively.

[3] Problems with automatic tracing of tongue curves from ultrasound images have been described in the literature; cf. Stone (2005: 30): "Automatic edge extraction and tracking that allows fast accurate measurements of tongue contours would be ideal for ultrasound tongue analysis. Unfortunately, because the ultrasound image is very noisy, the brightest edge may not be the important edge and it has been found that even automatic systems need to be supervised by a human experimenter to be sure that the selected edges are accurate." The same article cited previous studies which have assessed within-transcriber and between-transcriber reliability for hand measurements, and reported measurement errors of 0.7 mm or less; it is also stated that "human measurement error can be reduced with training to within 1 pixel. One pixel typically represents 0.25-0.5mm depending on the depth setting used and the machine" (Stone 2005: 30).

Normalisation for time

A major purpose of the study was to compare size of coarticulation at comparable time-points, which required a normalisation for time. In each consonant token, x-y values for the tongue curve data from nine equally spaced time points (henceforth Normalised Time Points, NTPs) for each speaker were extracted from Articulate Assistant Advanced. The first and the last of the nine NTPs coincided with the onset and offset of the consonant, respectively, and were used as the reference points for the normalisation. The software automatically constructed tongue curves at each NTP using a linear interpolation between the two surrounding tongue curves originally traced at 10 ms intervals throughout the consonant (the curves for the first and the last NTPs were the same as the first and the last originally traced curves, respectively). The number of NTPs was chosen to be one less than the number of ultrasound frames in the shortest consonant token (the shortest consonant token, an /s/ produced by Adult 7, was 97 ms long). This ensured that the tongue curves constructed for any two consecutive NTPs did not share data from the same two ultrasound frames. Since there were nine NTPs for each token collected, this yielded 720 x 9 = 6480 curves, in total.

Quantitative comparison of tongue curves

The approach taken to quantifying coarticulatory influence on the initial consonant from the following vowel was to measure distances between (for example) the tongue curves of s_i and the tongue curves of s_a and to compare them with the distances among the tongue curves of s_i and with the distances among the tongue curves of s_a[4]. The nearest neighbour method was used to determine the distance separating a pair of curves, A and B. It was implemented in Python (Lutz,

[4] The abbreviation s_a is used for /s/ in the context of /a/, and s_i denotes /s/ in the context of /i/.

2008), using the method described in Zharkova & Hewlett (2009). This method involves computing the mean of the Euclidean distances from each point on curve A to the point nearest to it on curve B, and from each point on curve B to that nearest to it on curve A. Nearest neighbour distances from each consonant curve in one vowel context to each consonant curve in the other vowel context were calculated. These were called Across-Set (AS) distances. For each speaker and each consonant, there were 36 AS distances (36 being the square of the six repetitions). So, for example, for the preadolescent group there were a total of 36 x 15 = 540 AS distances between the s_i consonant curves and the s_a consonant curves, at each NTP. The Within-Set (WS) distances (i.e., the distances among the consonant curves produced within the same vowel context) were calculated in analogous fashion. For each speaker for each consonant for each vowel context, there were 15 WS distances for the six repetitions (N*(N-1)/2, where N = 6). So, for each group of speakers there was a total of 15 x 15 = 225 WS distances for s_i and 225 WS distances for s_a.

Normalisation for tongue size

Magnetic Resonance Imaging studies of vocal tract growth in humans have shown that in preadolescents, the vocal tract length is somewhere between 60% and 80% of that in adults (e.g., Fitch & Giedd, 1999; Vorperian et al., 2009), raising the possibility of an effect from inequality of tongue size in our data. A normalisation based on imaged tongue length was therefore undertaken in this study, using the procedure described in Zharkova et al. (2011) in order to ensure that comparisons between adults and preadolescents were not obscured by vocal tract differences. Tongue lengths were measured, in mm, for every speaker. The tongue contour at /s/ offset in the context of /a/ (for every repetition by each speaker) was selected for measurement,

because the tongue tip in this position is low, and therefore likely to be present in the ultrasound scan, and the hyoid bone is also low, allowing for more of the contour at the tongue back to be imaged. On the imaged tongue contour, represented as a sequence of x-y points, the Euclidean distance between each consecutive pair of x-y points was measured, and the sum of all the distances was computed, representing the tongue length. The six tongue length values for each speaker were averaged. The mean tongue length of the adults was 75.5 mm (ranging between 59.7 mm and 91.6 mm), and for the preadolescents it was 67.3 mm (ranging between 57.2 mm and 76.9 mm). Mean tongue length values for each participant were represented as a proportion of the tongue length value for the speaker with the longest imaged tongue surface, and all AS and WS distances for each speaker were divided by the proportionate tongue length values. In statistical testing which included comparisons of spatial measures between age groups, the normalised values were used. Elsewhere, the non-normalised values were used, on the principle that analyses should be applied to data in their rawest form possible.

Cross-group comparison of random variability

Since WS distances reflect random variability in tongue positioning, they provide a convenient measure to compare variability of tongue placement in adults versus preadolescents. For each consonant for each group, the normalised WS distances in both vowel contexts and every NTP were pooled and compared across Age Group. Two linear mixed models (LMMs), one for each consonant, were carried out in R, using *lmer* software package (Baayen, 2008), with Speaker as a random factor. The significance of the main effect was established by subjecting the results of each LMM to an ANOVA in R. Because establishing denominator degrees of freedom in LMMs is somewhat problematic (Baayen et al., 2008), for this and other LMMs reported here we

adopted the approach of Reubold et al. (2010), also used in Beňuš (2012). In this approach, the denominator degrees of freedom are set at 60, the main effect is deemed significant at the 0.05 level if the *F* value in the ANOVA exceeds 7.2, and at 0.01 level if the *F* value exceeds 8.49.

Comparing size of coarticulatory effect, across consonant and across age group

AS distances reflect size of coarticulatory effect but only on condition that they significantly exceed the corresponding WS distances (otherwise, the two sets of curves, one from each context, must be assumed to have similar shape and orientation). The first step, therefore, involved testing for a significant difference between AS and corresponding WS distances, at each NTP. Then, for NTPs at which a significant difference is found, the size of coarticulatory effect, across consonant or across age group, could be statistically compared using the AS distances only. These two steps are described in turn, below.

At each NTP, LMMs were run, with Speaker being a random factor. Significance testing was carried out using a method first published in Zharkova & Hewlett (2009). In total, 36 tests were run, separately for adult /s/, adult /ʃ/, preadolescent /s/ and preadolescent /ʃ/, at every consecutive NTP starting from the consonant onset. A significant difference between AS and WS distances, depending on the upcoming vowels, was deemed to be present if the non-normalised AS distances were significantly greater than both sets of non-normalised WS distances. The main effect of Distance Type (three levels: AS, WS_a and WS_i) needed to be significant, as well as the results of two *post hoc* comparisons: between AS and WS_a, and between AS and WS_i[5]. Tukey

[5] The abbreviation WS_a is used for WS distances in the context of /a/, and WS_i refers to WS distances in the context of /i/.

post hoc tests were used to compare AS values to the two sets of WS values. For the results of the *post hoc* comparisons, Bonferroni adjustment was applied (for 36 models). Therefore if after finding a significant main effect of Distance Type, both *p* values in a Tukey test were lower than 0.00139 (this number was obtained by dividing 0.05 by 36), then the tongue curves for the consonant across vowel environments were deemed to be significantly different from each other.

For each NTP at which both groups of speakers had a significant vowel-related difference between consonant tongue contours for both consonants, a LMM was carried out on the normalised AS distances, with Consonant and Age Group as fixed factors and Speaker as a random factor. Two separate LMMs on these data were also carried out, one with Consonant as a fixed factor and one with Age Group as a fixed factor.

Time of onset of coarticulation: testing for significance

Comparison of AS and WS distances, at each NTP, as described at the beginning of the previous section, yields information concerning the development of the coarticulatory effect over the consonant, for each consonant and each age group. However, finding a difference, between age groups or between consonants, in the NTP at which a significant vowel-related effect was first apparent does not necessarily imply a *significant* difference in timing of onset of the effect and testing for significant differences in timing was not possible within the statistical framework described above. In order to answer this question, therefore, a different approach was used, based on individual speaker means within each group, for each consonant. For each speaker and each consonant, it was established at which NTP the AS distance value first exceeded both WS distance values (the data from six repetitions were pooled within speaker), with the proviso that both WS values continued to be exceeded at all subsequent NTPs. The resulting NTP values (15

values per consonant per group), representing the onset of coarticulatory effect for each speaker, were compared across the two groups of speakers, separately for /s/ and for /ʃ/, using two Mann-Whitney U-tests, with a Bonferroni adjustment for two tests. For the comparison of the time of onset of coarticulatory effect across consonants, two Wilcoxon signed-ranks tests were carried out, one for each group of speakers (also with a Bonferroni adjustment for two tests). A virtue which arises from the necessity of applying quite different statistical treatments to the same data is that it allows comparison of outcomes: the NTP values for onset of coarticulation arrived at using the method described here should not be radically different from the NTP values for onset of the vowel-related difference between mean group AS and WS distances, arrived at under the procedure described in the previous section.

Results

A qualitative description of the tongue movement patterns over the consonant is presented first, illustrated by four figures. This is followed by the results on within-speaker variability. The findings on size of effect are then presented. Finally, comparison of the time of onset of the vowel effect across age groups is described.

Figure 1 displays all consecutive tongue curves during one repetition of each consonant, separately for the two vowel contexts, in two representative speakers, one adult and one preadolescent. In order to give some idea of the nature of the tongue movement over the time course of the fricative, the tongue curves collected from the first and second half of the segment are shown by red solid and black dotted lines, respectively. The figure shows that there was in fact very little movement during /s/ where the following vowel was /a/. In the context of the vowel /i/, on the other hand, the tongue moved noticeably more than in the context of /a/, in both

speakers; this difference in tongue movement over the consonant across the two vowel contexts was very consistent across both adult and preadolescent speakers. Figures 2 and 3 present tongue curves for /ʃ/ in the two vowel contexts, for the nine NTPs, in an adult and a preadolescent, respectively. For each of the nine NTPs, all the curves of the six repetitions are shown, for each vowel context, with those from the /i/ context represented by black dotted lines and those from the /a/ context represented by red solid lines. The figures show that the back of the tongue was further forward before /i/ than before /a/ in both speakers, and, by NTP 9, at least, the front was slightly higher.

<Figure 1 about here>

<Figure 2 about here>

<Figure 3 about here>

Figure 4 shows mean (non-normalised) AS distances and WS distances at each time point, for both consonants, for adults (on the left) and preadolescents (on the right). It illustrates increasing AS distance over time, while, as would be expected, WS distances, which reflect random variation in tongue position, remain more or less static. Broadly similar patterns are apparent in the adults and the preadolescents, although in /ʃ/ the AS distances appear to rise more slowly in the preadolescents than in the adults, during the earlier NTPs.

<Figure 4 about here>

Cross-group comparison of random variability

Mean normalised WS distances for /s/ were 1.26 mm in the adults and 1.59 mm in the preadolescents; for /ʃ/, they were 1.18 mm in the adults and 1.47 mm in the preadolescents. The results of the LMMs for both /s/ and /ʃ/ showed that in neither case was the main effect of Age Group significant.

Comparison of size of coarticulatory effect

Table 2 shows *F* values from LMMs testing for the presence of a significant vowel-related difference between consonant tongue contours at each of the nine NTPs. The adult speakers showed a significant difference between the two sets of tongue contours from NTP1 onwards (i.e., from the consonant onset), in both /s/ and /ʃ/. The preadolescents showed such a difference from NTP2 onwards in /s/ and from NTP4 onwards in /ʃ/.

<Table 2 about here>

Since both groups of speakers had a significant vowel-related difference between AS and WS distances by NTP 4 in both consonants, significance testing to compare size of effect was carried out on the normalised AS distances for NTPs 4-9. The distance values at these NTPs are reported in Table 3, and the results of the LMMs with Consonant and Age Group as fixed factors can be found in Table 4. The models showed that there was a significant interaction of these two factors between NTP 4 and NTP 7 (with the relative difference between coarticulatory patterns of /s/ and /ʃ/ being greater in the preadolescents than in the adults), but not at the last two NTPs. The LMMs comparing adults and preadolescents, also at NTPs 4-9, showed that the effect of

Age Group was not significant in the case of either consonant, at any NTP. The LMMs comparing /s/ with /ʃ/ at NTPs 4-9 showed that in the preadolescents, the coarticulatory effect was significantly greater in /s/ than in /ʃ/ at every NTP. For the adults, the difference between the two consonants was significant starting from NTP 5.

<Table 3 about here>

<Table 4 about here>

Time of onset of coarticulation

Using the criteria described in the last sub-section of 'Method', above, the mean NTP at which onset of coarticulation occurred in /s/ for the adult participants was 1.8, and for the preadolescents it was 2.0. For /ʃ/, the corresponding means were 1.7 and 3.5 for the adults and preadolescents, respectively. Figure 5 shows plots for one adult and one preadolescent, for /s/ and for /ʃ/. In the Mann-Whitney U-test for /s/, there was no significant difference between adults and preadolescents. For /ʃ/, there was a significant age-related difference ($p < 0.05$, two-tailed, after Bonferroni adjustment), showing that in the preadolescents the onset of the vowel effect was significantly later than in the adults. In the Wilcoxon Signed Ranks Test for the adults, there was no significant difference between the two consonants. For the preadolescents, the difference between /s/ and /ʃ/ was significant ($p < 0.01$, two-tailed, after Bonferroni adjustment).

<Figure 5 about here>

Discussion

This section begins with some observations on the contrasting tongue movement trajectories of the fricatives, induced by differing vowel contexts, as captured in Figure 1, above. These are followed by discussion of the fate of the hypotheses, in the order in which they are summarised at the end of the Introduction. The only one of these which received support is that which posited a greater coarticulatory effect on /s/ than on /ʃ/ in the case of the adult speakers (H4). The remaining hypotheses were rejected. In one case (H2) the findings would support an opposite conclusion to that predicted (in that the preadolescents started to coarticulate /ʃ/ later, rather than earlier, than the adults), but only with respect to the time-course of coarticulation. Overall, the results suggest that children of 10-12 years of age exhibit similar precision of lingual control and broadly similar lingual coarticulatory behaviour to adults.

Tongue movement trajectories in the two vowel contexts

Visual observation of Figure 1 showed that there was a greater extent of tongue movement during the fricative in syllables with /i/ than those with /a/ in both age groups, and this pattern was particularly pronounced in /s/.[6] When using difference between contexts as the basis for measuring extent of coarticulation, it is important to bear in mind that the speech sound in question (the candidate for coarticulation, so to speak) may be inherently more compatible with one conditioning context than the other. Thus the amount of midsagittal tongue movement

[6] It is important to point out that our findings on /s/ apply only to the activity of the tongue in the midsagittal plane. At the sides of the tongue, the opposite scenario would apply for /s/ in these two vowel contexts. Because /s/ is characterised by a pronounced groove along the tongue midline, reflected in EPG data (e.g., Hardcastle et al., 1991), during the syllable /si/ the tongue sides stay in contact with the hard palate, while during the syllable /sa/ there is an abrupt change in the amount of lateral tongue-palate contact between the consonant and the vowel, as the sides of the tongue lower for /a/.

required for the passage from /s/ to /i/, for example, would appear to be much greater than that required in the passage from /s/ to /a/, either because the tongue position for /a/ is more easily accommodated while articulating /s/ or because the tongue posture for /a/ is close to some sort of default posture for /s/ (e.g., a low tongue tip for /s/, cf. Mooshammer et al., 2007). The possibility of perseverative coarticulation (an effect on the following vowel from consonant identity, which would reduce the difference between the target /a/ and /i/ vowels) must also be borne in mind. However, it is hard to think of a plausible role for perseverative coarticulation in explaining the subtle timing differences between the age groups with respect to one of the consonants, that emerged in the results.

The direction of tongue movement during the given sequence of phonemes is relevant to the outcomes of the analysis of coarticulation. For example, in /əsi/ and /əʃi/ sequences, the tongue has to move continuously upwards and forwards, and this movement can start before the fricative onset (cf. considerable tongue raising in /aʃi/ and /asi/ during the initial vowel in adult speakers, reported in Recasens et al., 1997: 555-556), while in /əʃa/ sequences, the tongue needs to first raise towards /ʃ/, and then lower again towards /a/. We will return to this point when discussing age-related differences observed in the study.

Cross-group comparison of variability

The hypothesis on greater variability in preadolescents than in adults (H1) was not supported, since the two groups were not significantly different from each other. Our finding that the preadolescents had adult-like variability does not accord with that of Lee et al. (1999) who found greater than adult variability up to the age of 14 years. However, it is possible that the 10-12 year

old period represents an interlude, in this respect, due to it being a period of stability in vocal tract development. Temple et al. (2002) claimed that the period between 10 and 12 years of age is characterised by a plateau in the growth of the vocal tract, after the period of rapid growth between seven and ten years old. On the other hand, this finding does accord well with the general pattern in the findings of the present study. The fact that the amount of random variability of tongue posture was similar in the two groups lends support to a conclusion that the preadolescents had a near-adult-like style of lingual motor control.

Development of a vowel effect over the time-course of the consonant

The hypotheses on age-related differences in the development of the vowel effect in /s/ and /ʃ/ (H2 and H3) were both rejected, since no age-related differences were reported in the spatial domain for either consonant or in the temporal domain for /s/, while for /ʃ/ the age-related difference in the temporal domain was in the opposite direction to that predicted (i.e., the adults demonstrated earlier onset of a vowel effect). The hypothesis on consonant-specific differences in adults (H4) was supported with respect to the results on the size of effect between NTPs 5 and 9, in that the effect was significantly greater in /s/ than in /ʃ/ (though no significant difference in the time of onset was found). The hypothesis on consonant-specific differences in preadolescents (H5) was rejected both with respect to the relative timing of onset in the two fricatives and to the relative size of effect.

The general picture that emerges from these results, then, is that there are consonant-identity-dependent patterns, but they are not the same as we predicted on the basis of the outcomes of our studies of younger children. The claim that /s/ has a lower coarticulation

resistance than /ʃ/ in adult speech, as proposed in Recasens & Espinosa (2009), was supported not only by the adult results, but also by the preadolescent results, in that there were greater AS distances in /s/ than in /ʃ/ for both groups of speakers. However the two fricative consonants did not stand in the same relation to each other in the speech of the preadolescents as in the speech of the adults. In the preadolescents, the onset of anticipatory coarticulation for /ʃ/ was manifested later in the consonant than for /s/, while the difference between /s/ and /ʃ/ in the onset of effect was not significant for the adults. Also, the relationship between the two consonants in size of coarticulatory effect was different between the two groups of speakers. The finding of a significant interaction of Age Group and Consonant at NTPs 4 to 7 indicates that the preadolescents differentiated the coarticulatory behaviours of the two fricatives more than the adults.

The results make good sense insofar as they suggest that preadolescents have more mature coarticulation patterns than younger children (as evidenced by the fact that the preadolescents in this study had an adult-like amount of coarticulation at mid-/s/ and at mid-/ʃ/, unlike the 6-9-year-olds in Zharkova et al., 2011; 2012) and slightly less mature coarticulation patterns than adults, with the immaturity of pronunciation concerning timing of onset of coarticulation. Temporal aspects of speech production have been shown to present a challenge for children (e.g., Smith & Goffman, 1998, for 4- and 7-year-olds, and Koenig et al., 2008, for 4-6 and 8-11-year-olds) and even for 14- and 16-year-old adolescents (Sadagopan & Smith, 2008). Below we consider how the characteristics of the target segment sequences might affect temporal coarticulation of the two consonants.

The adults coarticulated the two fricatives in much the same way during the first half of the consonant. The preadolescents, on the other hand, displayed different coarticulatory strategies for the two fricatives. The explanation for this pattern of results might lie in the actual tongue movement patterns for the fricatives in the two vowel contexts. It might have been that in sequences with /i/ produced by the adults, the tongue started the raising-advancing movement already during the preceding schwa (this would accord with the tongue movement pattern reported in Recasens et al., 1997). The movement during /əʃa/ could have also started relatively early, and included not only the dorsum raising towards the consonant, but also a concomitant retraction of the posterior tongue anticipating the low vowel, with little change during the fricative (Figure 2 illustrates a relatively retracted back of the tongue in an adult speaker). Thus in the adult productions, the coarticulatory effect from the following vowel would be observed relatively early in the consonant, for both fricatives. In the preadolescents, the tongue raising and advancing in sequences with /i/ could have started early – indeed, examining Figure 1 suggests that during the first half of the consonant there was relatively more movement in the preadolescent than in the adult, especially in the front of the tongue. However in /əʃa/ sequences, the preadolescents probably had a relatively less retracted tongue root early on in the consonant than the adults (such a pattern is illustrated in Figure 3), thus making /ʃ/ display less coarticulation than in the adults (and also less than the preadolescent /s/) in the early stages of the consonant. In terms of coarticulatory behaviour, then, the two consonants would pattern more or less together in the adults, as opposed to the preadolescents (as indeed shown by the Age Group x Consonant interaction described above). The reason for a delay in coarticulation for the preadolescent /ʃ/ could be that the sequence /əʃa/ differs from the other three sequences involved

in this study, in that the tongue dorsum has to move consecutively in two opposing directions (i.e., raising then lowering).

The implications of the foregoing explanation for the discussion of units of speech production may be worth considering. Whenever coarticulation is found, it follows that both (or all) segments involved must be contained within the same programming chunk. The converse does not follow. That is, where coarticulation is *not* found it does not follow that the relevant motor sequences must have been in different chunks. Certainly, the latter would seem most unlikely in the present case, which concerns time of onset within the duration of a single segment. It seems more likely that the observed differences between the preadolescents and the adults relate to the detailed content of the relevant motor programs, rather than the size of the chunks involved. This interpretation agrees with evidence from our previous findings of segment-specific coarticulatory patterns differing across age (Zharkova et al., 2011, 2012). Had the difference between children and adults been only in the size of units, then in those two studies both /s/ and /ʃ/ would have been coarticulated more in one group of speakers and less in the other group, i.e., in the same way within a group. In combination with our previous work, the results of the present study suggest that age-related differences in motor programming may depend not on whether the units are syllables or segments but on the nature of the task to be performed by the articulators in each case.

The pattern observed in the preadolescents suggests that there might be differences between preadolescents and adults at the motor programming stage of the utterance production, in that preadolescents differ from adults in programming complex sequences of tongue shapes. Previous research has found a relationship between complexity of articulatory movements and speech motor programming. For example, syllables with a more complex segmental structure

(such as “CCCV” vs “CV”) have been shown to take speakers longer to program (Wright et al., 2009). The study by Wright et al. (2009) was carried out within the paradigm where speech motor programming is investigated by measuring reaction times to stimuli (e.g., Klapp, 1996), rather than by quantifying coarticulation within an utterance. To our knowledge, there are no studies assessing whether complexity in direction of tongue movement within a sequence of segments affects the timing of lingual coarticulation during speech development. This suggestion would need to be verified using articulatory data on tongue displacement. We are currently exploring possibilities for such analysis.

Methodological issues and limitations

Comparative measures of coarticulation can be made in either the temporal (e.g., Katz & Bharadwaj, 2001) or the spatial domain (e.g., Nittrouer et al., 1989), or both (e.g., Farnetani & Recasens, 1993; Recasens et al., 1997). The higher frame rate that has now become available in ultrasound systems for speech analysis (see, e.g., Wrench & Scobbie, 2008; Stone, 2010; Miller & Finch, 2011; Lee et al., 2013) enabled us to monitor coarticulation in the temporal as well as the spatial domain: i.e., to determine time of onset of a vowel effect on the fricative as well as to measure the size of the effect during the fricative. The use of normalised time points, as opposed to absolute intervals, is unlikely to have made any difference to the outcomes of the study. There is evidence to suggest that the development of /s/ duration follows a U-shaped trajectory (shortest in the middle), with the turn likely occurring somewhere toward the end of preadolescent years (Lee et al., 1999). In our study, fricative durations turned out to be rather similar in the two groups (177 ms and 181 ms for /s/ and /ʃ/, respectively in the adults and 169 ms and 177 ms for /s/ and /ʃ/, respectively, in the preadolescents). In principle, however, we

would still claim that making comparisons on the basis of relative rather than absolute time is probably the more valid approach. Indeed, in a segment-based approach, it is more or less a necessity to apply a normalisation for time in order to create some sort of equivalence in the sampling points over the duration of the segment, since segment durations will obviously vary to some extent among tokens.

While it is known that the ultrasound transducer in speech studies may restrict jaw motion during speech (see Stone, 2005), there is evidence from other articulatory data, suggesting that our results are unlikely to have been affected by this. The greater movement required over the time-course of /s/ before /i/ as compared with /s/ before /a/ accords with the findings reported by Iskarous et al. (2011), who used x-ray microbeam data. The range of the tongue blade and dorsum motion along the along the *y* axis during the fricative /s/, in a range of following vowel contexts, was shown in their Figure 3 to be noticeably greater in the context of /i/ than in the context of open vowels. From the methodological point of view, in respect to the potential influence from the ultrasound transducer on jaw movement, it is particularly reassuring that the tongue movement data in that figure were not normalised relative to the jaw position.

The findings did not support the predictions concerning differences between adults and preadolescents. Of course, there is always the possibility that a hypothesised effect was not in fact absent but merely insufficiently robust to emerge from our data or that the statistical power of the tests was insufficient to overcome inter-participant variability in either or both groups, possibly due to factors which were not controlled for, such as gender in preadolescents (cf. Fitch & Giedd, 1999) or age in adults (cf. Bennett et al., 2007). We did however observe statistically significant age-related differences in consonant-specific patterns of coarticulation, albeit in an unpredicted direction.

While some results from the analyses of temporal and spatial coarticulation patterned similarly (e.g., the later onset of coarticulation in preadolescent /ʃ/, and a relatively larger difference between coarticulatory patterns of the two fricatives in the preadolescents than in the adults), there were also results suggesting that time of onset and size of effect at a given time point are not inter-changeable measures of extent of coarticulation. The evidence from our results on /ʃ/ is that two groups of speakers can have similar amounts of coarticulation during the latter part of the segment in question, while the temporal onset of an effect, following the beginning of frication, is retarded in one group compared to the other.

Conclusion

The internal coherence of the results of this study confirms the usefulness of ultrasound as a means for investigating speech development and coarticulation. The results suggest an adult-like degree of lingual control on the part of preadolescent children, with respect to the level of random variability observed and the size of the vowel effect during the consonant in fricative-vowel sequences. However the preadolescents differed from the adults in the timing of the development of coarticulation during /ʃ/. The explanation we have suggested is that complexity of tongue movement patterns including a radical change of direction can result in a delayed onset of anticipatory vowel-on-consonant coarticulation in preadolescents.

Acknowledgements

We wish to thank Robert Rush for advice on statistical analysis, and Alan Wrench and Stephen Cowen for technical advice and assistance with instrumentation. The work was supported by ESRC research grants RES-000-22-4075 and ES/K002597/1.

References

Articulate Instruments Ltd (2008). *Ultrasound Stabilisation Headset Users Manual: Revision 1.4*. Edinburgh, UK: Articulate Instruments Ltd.

Articulate Instruments Ltd (2010). *Articulate Assistant Advanced User Guide: Version 2.12*. Edinburgh, UK: Articulate Instruments Ltd.

Baayen, R. H. (2008). *Analyzing Linguistic Data. A Practical Introduction to Statistics Using R*. 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.

Bennett, J. W., van Lieshout, P. H. H. M., & Steele, C. M. (2007). Tongue control for speech and swallowing in healthy younger and older subjects. *International Journal of Orofacial Myology, 33*, 5–18.

Beňuš, S. (2012). Phonetic variation in Slovak yer and non-yer vowels. *Journal of Phonetics*, In Press.

Bladon, R. A. W., & Al-Bamerni, A. (1976). Coarticulation resistance in English /l/. *Journal of Phonetics, 4,* 137–150.

Cheng, H., Murdoch, B., Goozée, J., & Scott, D. (2007a). 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.

Cheng, H., Murdoch, B., Goozée, J., & Scott, D. (2007b). Physiologic development of tongue-jaw coordination from childhood to adulthood. *Journal of Speech, Language, and Hearing Research, 50,* 352–360.

Davis, B. L., & MacNeilage, P. F. (1995). The articulatory basis of babbling. *Journal of Speech and Hearing Research, 38,* 1199-1211.

Farnetani, E., & Recasens, D. (1993). Anticipatory consonant-to-vowel coarticulation in the production of VCV sequences in Italian. *Language and Speech*, *36,* 279-302.

Fitch, W. T., & Giedd, J. (1999). Morphology and development of the human vocal tract: A study using magnetic resonance imaging. *Journal of the Acoustical Society of America, 106,* 1511–1522.

Gibbon, F. E. (1999). Undifferentiated lingual gestures in children with articulation/phonological disorders. *Journal of Speech, Language, and Hearing Research, 42*, 382–397.

Gick, B., Bacsfalvi, P., Bernhardt, B. M., Oh, S., Stolar, S., & Wilson, I. (2008). A motor differentiation model for liquid substitutions: English /r/ variants in normal and disordered acquisition. *Proceedings of Meetings on Acoustics 1, 060003*, 1–9.

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*, 1424–1437.

Gordeeva, O. (2005). *Language Interaction in the Bilingual Acquisition of Sound Structure: A Longitudinal Study of Vowel Quality, Duration and Vocal Effort in Pre-School Children Speaking Scottish English and Russian.* Unpublished PhD thesis, Queen Margaret University College, Edinburgh.

Green, J. R., Moore, C., 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*, 239–255.

Green, J. R., Moore, C. A., Reilly, K. J. (2002). The sequential development of jaw and lip control for speech. *Journal of Speech, Language, and Hearing Research, 45*, 66–79.

Grigos, M. I. (2009). Changes in articulator movement variability during phonemic development: A longitudinal study. *Journal of Speech, Language, and Hearing Research, 52*, 164–177.

Guenther, F. H., & Perkell, J. S. (2004). A neural model of speech production and its application to studies of the role of auditory feedback in speech. In B. Maassen, R. D. Kent, H. F. M. Peters, P. H. H. M. van Lieshout, & W. Hulstijn (Eds.), *Speech Motor Control in Normal and Disordered Speech* (pp. 29–49). Oxford: Oxford University Press.

Hardcastle, W. J., Gibbon, F. E., & Jones, W. (1991). Visual display of tongue-palate contact: Electropalatography in the assessment and remediation of speech disorders. *British Journal of Disorders of Communication, 26*, 41–74.

Iskarous, K., Shadle, C. H., & Proctor, M. I. (2011). Articulatory-acoustic kinematics: The production of American English /s/. *Journal of the Acoustical Society of America, 129*, 944–954.

Katz, W. F., & Bharadwaj, S. (2001). Coarticulation in fricative-vowel syllables produced by children and adults: A preliminary report. *Clinical Linguistics and Phonetics, 15*, 139–143.

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. (1983). The segmental organization of speech. In P.F. MacNeilage (Ed.), *The Production of Speech*. New York: Springer-Verlag. Pp. 57–89.

Kent, R. D. (1992). 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.

Kent, R. D., & Forner, L. L. (1980). Speech segment durations in sentence recitations by children and adults. *Journal of Phonetics, 8*, 157–168.

Klapp, S. T. (1996). Reaction time analysis of central motor control. In H. N. Zelaznik (Ed.), *Advances in Motor Learning and Control*. (pp. 13–35). Champaign, IL: Human Kinetics.

Klapp, S. T. (2003). Reaction time analysis of two types of motor preparation for speech articulation: Action as a sequence of chunks. *Journal of Motor Behavior, 35*, 135-150.

Koenig, L. L., Lucero, J. C., & Perlman, E. (2008). Speech production variability in fricatives of children and adults: Results of functional data analysis. *Journal of the Acoustical Society of America, 124*, 3158–3170.

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

Lee, A., Zharkova, N., & Gibbon, F. (2013). Vowel imaging. In M. Ball & F. Gibbon (Eds), *Handbook of Vowels and Vowel Disorders* (2nd edition) (pp. 138-159). Hove: Psychology Press.

Locke, J. (1983). *Phonological Acquisition and Change.* New York: Academic Press.

Lutz, M. (2008). *Learning Python* (3rd edn.) Beijing, Cambridge, Farnham, Köln, Sebastopol, Taipei, Tokyo: O'Reilly Media.

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

Mooshammer, C., Hoole, P., & Geumann, A. (2007). Jaw and order. *Language and Speech, 50*, 145–176.

Murdoch, B. E., Cheng, H.-Y., & Goozée, J. V. (2011). Developmental changes in the variability of tongue and lip movements during speech from childhood to adulthood: An EMA study. *Clinical Linguistics & Phonetics, Early Online*, 1–16.

Nam, H., Goldstein, L. M., Giulivi, S., Levitt, A. G., & Whalen, D. H. (2013). Computational simulation of CV combination preferences in babbling. *Journal of Phonetics, 41*, 63-77.

Niebuhr, O., & Meunier, C. (2011). The phonetic manifestation of French /s#ʃ/ and /ʃ#s/ sequences in different vowel contexts: On the occurrence and the domain of sibilant assimilation. *Phonetica, 68*, 133–160.

Niebuhr, O., Clayards, M., Meunier, C., & Lancia, L. (2011). On place assimilation in sibilant sequences – comparing French and English. *Journal of Phonetics, 39*, 429–451.

Nijland, L., Maassen, B., Van der Meulen, S., Gabrieëls, F., Kraaimaat, F. W., & Schreuder, R. (2002). Coarticulation patterns in children with developmental apraxia of speech. *Clinical Linguistics and Phonetics, 16*, 461–483.

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

Nittrouer, S., Estee, S., Lowenstein, J. H., & Smith, J. (2005). The emergence of mature gestural patterns in the production of voiceless and voiced word-final stops. *Journal of the Acoustical Society of America, 97*, 351–364.

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 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 and Hearing Research, 39*, 379–389.

Noiray, A., Ménard, L. & Iskarous, K. (2013). The development of motor synergies in children: Ultrasound and Acoustic Measurements. *Journal of the Acoustical Society of America, 133,* 444–452.

Pouplier, M., Hoole, P., & Scobbie, J. M. (2011). Investigating the asymmetry of English sibilant assimilation: Acoustic and EPG data. *Journal of Laboratory Phonology, 2*, 1–33.

Recasens, D., & Espinosa, A. (2009). An articulatory investigation of lingual coarticulatory resistance and aggressiveness for consonants and vowels in Catalan. *Journal of the Acoustical Society of America, 125*, 2288–2298.

Recasens, D., Pallarès, M. D., & Fontdevila, J. (1997). A model of lingual coarticulation based on articulatory constraints. *Journal of the Acoustical Society of America, 102*, 544–561.

Reubold, U., Harrington, J., & Kleber, F. (2010). Vocal aging effects on F0 and the first formant: a longitudinal analysis in adult speakers. *Speech Communication, 52,* 638–651.

Sadagopan, N., & Smith, A. (2008). Developmental changes in the effects of utterance length and complexity on speech movement variability. *Journal of Speech, Language, and Hearing Research, 51,* 1138–1151.

Scobbie, J. M., Gordeeva, O. G., & Matthews, B. (2007). Scottish English speech acquisition. In S. McLeod (Ed.), *The International Guide to Speech Acquisition* (pp. 221–240). Clifton Park, NY: Thomson Delmar Learning.

Scobbie, J., Wrench, A., van der Linden, M. (2008). Head-probe stabilisation in ultrasound tongue imaging using a headset to permit natural head movement. In R. Sock, S. Fuchs & Y. Laprie (Eds.), *Proceedings of the 8th International Seminar on Speech Production 2008, Strasbourg, France, 8-12 December 2008,* 373–376.

Serkhane, J. E., Schwartz, J. L., Boë, L. J., Davis, B. L., Matyear, C. L. (2007). Infants' vocalizations analyzed with an articulatory model: a preliminary report. *Journal of Phonetics, 35,* 321-340.

Sharkey, S., & Folkins, J. (1985). Variability of lip and jaw movements in children and adults: Implications for the development of speech motor control. *Journal of Speech and Hearing Research, 28,* 3–15.

Siren, K. A., & Wilcox, K. A. (1995). Effects of lexical meaning and practiced productions on coarticulation in children's and adults' speech. *Journal of Speech and Hearing Research, 38*, 351-359.

Smith, A., & Goffman, L. (1998). Stability and patterning of speech movement sequences in children and adults. *Journal of Speech and Hearing Research, 41*, 18–30.

Smith, A., & Goffman, L. (2004). Interaction of motor and language factors in the development of speech production. In B. Maassen, R. D. Kent, H. F. M. Peters, P. H. H. M. van Lieshout, & W. Hulstijn (Eds.), *Speech Motor Control in Normal and Disordered Speech* (pp. 227–252). Oxford: Oxford University Press.

Smith, A., & Zelaznik, H. N. (2004). Development of functional synergies for speech motor coordination in childhood and adolescence. *Developmental Psychobiology, 45*, 22-33.

Stathopoulos, E. T., Huber, J. E., & Sussman, J. E. (2011). Changes in acoustic characteristics of the voice across the life span: Measures from individuals 4–93 years of age. *Journal of Speech, Language, and Hearing Research, 54*, 1011–1021.

Stone, M. (2005). A guide to analyzing tongue motion from ultrasound images. *Clinical Linguistics and Phonetics*, **19**, 455-502.

Stone, M. (2010). Laboratory techniques for investigating speech articulation. In Hardcastle, W. J., Laver, J. & Gibbon, F. E. (Eds.), *The Handbook of Phonetic Sciences* (2nd ed.) (pp. 9–38). Chichester, West Sussex, U.K.; Malden, MA: Wiley-Blackwell.

Sussman, H. M., Hoemeke, K. A., & McCaffrey, H. A. (1992). Locus equations as an index of coarticulation for place of articulation distinctions in children. *Journal of Speech and Hearing Research, 35*, 769–781.

Temple, E. C., Hutchinson, I., Laing, D. G., & Jinks, A. L. (2002). Taste development: Differential growth rates of tongue regions in humans. *Developmental Brain Research, 135*, 65–70.

Terband, H., Maassen, B., van Lieshout, P., & Nijland, L. (2011). Stability and composition of functional synergies for speech movements in children with developmental speech disorders. *Journal of Communication Disorders, 44*, 59–74.

Van der Merwe, A. (2009). A theoretical framework for the characterization of pathological speech sensorimotor control. In M. R. McNeil (Ed.), *Clinical Management of Sensorimotor Speech Disorders* (2nd edition). (pp. 3-18). New York: Thieme Medical Publishers.

Vihman, M. M. (1996). Phonological Development: The Origins of Language in the Child. Oxford: Blackwell.

Vihman, M. M., DePaolis, R. A., & Keren-Portnoy, T. (2009). A Dynamic Systems approach to babbling and words. In E. Bavin (Ed.), *Handbook of Child Language* (pp. 163-182). Cambridge: Cambridge University Press.

Vorperian, H. K., Wang, S., Chung, M. K., Schimek, E. M., Durtschi, R. B., Kent, R. D., Ziegert, A. J., & Gentry, L. R. (2009). Anatomic development of the oral and pharyngeal portions of the vocal tract: An imaging study. *Journal of the Acoustical Society of America, 125*, 1666–1678.

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

Wrench, A., & Scobbie, J. (2008). High-speed cineloop ultrasound vs. video ultrasound tongue imaging: Comparison of front and back lingual gesture location and relative timing. In R. Sock, S. Fuchs & Y. Laprie (Eds.), *Proceedings of the 8th International Seminar on Speech Production 2008, Strasbourg, France, 8-12 December 2008*, 57–60.

Wright, D. L., Robin, D. A., Rhee, J., Vaculin, A., Jacks, A., Guenther, F. H., & Fox, P. T. (2009). Using the self-select paradigm to delineate the nature of speech motor programming. *Journal of Speech, Language, and Hearing Research, 52*, 755–765.

Zharkova, N., & Hewlett, N. (2009). Measuring lingual coarticulation from midsagittal tongue contours: Description and example calculations using English /t/ and /ɑ/. *Journal of Phonetics, 37*, 248–256.

Zharkova, N., & Lickley, R. (2010). An acoustic and articulatory investigation of high vowels in Scottish English children and adults. Oral paper at the British Association of Academic Phoneticians Colloquium, London, 29-31 March 2010.

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*, 193-208.

Ziegler, W. & Maassen, B. (2004). The role of the syllable in disorders of spoken language production. In B. Maassen, R. D. Kent, H. F. M. Peters, P. H. H. M. van Lieshout, & W. Hulstijn (Eds.), *Speech Motor Control in Normal and Disordered Speech* (pp. 415–447). Oxford: Oxford University Press.

Figure legends

Figure 1. Consecutive tongue curves (every 10 ms) during one repetition of each consonant in each vowel context, in two representative speakers. The curves during the first half of the consonant are in red lines; the curves during the second half of the consonant are in black dotted lines.

Figure 2. Tongue curves for /ʃ/ in the two vowel contexts (black dotted curves – the context of /i/; red curves – the context of /a/), for the nine NTPs, for Adult 5. On each plot, there are curves for six repetitions in each vowel context.

Figure 3. Tongue curves for /ʃ/ in the two vowel contexts (black dotted curves – the context of /i/; red curves – the context of /a/), for the nine NTPs, for Preadolescent 6. On each plot, there are curves for six repetitions in each vowel context.

Figure 4. AS distances (solid lines), WS_a distances (dotted lines) and WS_i distances (dashed lines), in mm, at each of the nine NTPs during /s/ (top) and /ʃ/ (bottom), in adults (left panels) and preadolescents (right panels).

Figure 5. AS distances (solid lines), WS_a distances (dotted lines) and WS_i distances (dashed lines), in mm, at each of the nine NTPs during /s/ (top) and /ʃ/ (bottom), in one adult (left panels) and one preadolescent (right panels).

Tables

Table 1. Ages of participants.

Participant number	Adults (years)	Preadolescents (years, months)
1	30	10;10
2	52	10;0
3	46	11;3
4	58	11;3
5	48	11;6
6	46	11;1
7	48	12;4
8	26	10;11
9	26	10;11
10	43	11;10
11	42	11;10
12	29	11;0
13	18	11;10
14	22	10;6
15	21	10;6

Table 2. *F* values from linear mixed models testing for the presence of a significant vowel-related difference in tongue curves at the nine NTPs. An asterisk next to an *F* value means that a significant vowel-related difference was present after the Bonferroni adjustment. In cases where at least one of the two relevant *p* values (for the two pairwise comparisons) was greater than 0.00139, both *p* values are given in the relevant cell. In all other cases, $p < 0.001$.

NTP	/s/		/ʃ/	
	Adults	Preadolescents	Adults	Preadolescents
1	$F = 150.08$ *	$F = 14.31$ AS-WS_a: $p < 0.001$ AS-WS_i: $p = 0.002$	$F = 119.90$ *	$F = 2.93$ AS-WS_a: $p = 0.080$ AS-WS_i: $p = 0.200$
2	$F = 203.59$ *	$F = 27.07$ *	$F = 146.57$ *	$F = 5.32$ AS-WS_a: $p = 0.026$ AS-WS_i: $p = 0.021$
3	$F = 242.21$ *	$F = 48.34$ *	$F = 253.22$ *	$F = 9.98$ AS-WS_a: $p = 0.005$ AS-WS_i: $p < 0.001$
4	$F = 329.46$ *	$F = 96.58$ *	$F = 273.81$ *	$F = 18.53$ *
5	$F = 433.11$ *	$F = 217.46$ *	$F = 404.07$ *	$F = 41.12$ *
6	$F = 697.01$ *	$F = 455.43$ *	$F = 462.73$ *	$F = 99.14$ *
7	$F = 1272.70$ *	$F = 936.53$ *	$F = 667.54$ *	$F = 266.18$ *
8	$F = 2474.40$ *	$F = 1928.20$ *	$F = 1099.30$ *	$F = 667.04$ *
9	$F = 3922.80$ *	$F = 2931.10$ *	$F = 1566.80$ *	$F = 1301.30$ *

Table 3. Normalised AS distances for /s/ and /ʃ/ in adults and preadolescents, for those NTPs where a significant vowel-related difference between consonant tongue contours was observed, in both consonants and both age groups.

NTP	Adults		Preadolescents	
	/s/	/ʃ/	/s/	/ʃ/
4	2.27	2.15	2.33	1.73
5	2.64	2.24	2.73	1.87
6	3.33	2.55	3.48	2.12
7	4.57	2.94	4.58	2.65
8	6.25	3.78	6.17	3.57
9	7.97	4.90	7.51	4.62

Table 4. Results of significance testing for size of coarticulatory effect (*F* values are rounded to two decimal places). An asterisk next to an *F* value means that there was a significant effect at $p < 0.01$.

NTP	Adults vs preadolescents		/s/ vs /ʃ/		Age Group x Consonant
	/s/	/ʃ/	Adults	Preadolescents	
4	*F* = 0.05	*F* = 3.97	*F* = 6.56	*F* = 123.70 *	*F* = 44.02 *
5	*F* = 0.08	*F* = 2.60	*F* = 72.64 *	*F* = 247.06 *	*F* = 40.54 *
6	*F* = 0.19	*F* = 2.62	*F* = 259.46 *	*F* = 557.97 *	*F* = 58.49 *
7	*F* = 0.00	*F* = 0.82	*F* = 976.42 *	*F* = 950.86 *	*F* = 13.26 *
8	*F* = 0.03	*F* = 0.28	*F* = 1666.00 *	*F* = 1465.70 *	*F* = 2.13
9	*F* = 0.64	*F* = 0.32	*F* = 1936.90 *	*F* = 1614.20 *	*F* = 3.52

Figure 1

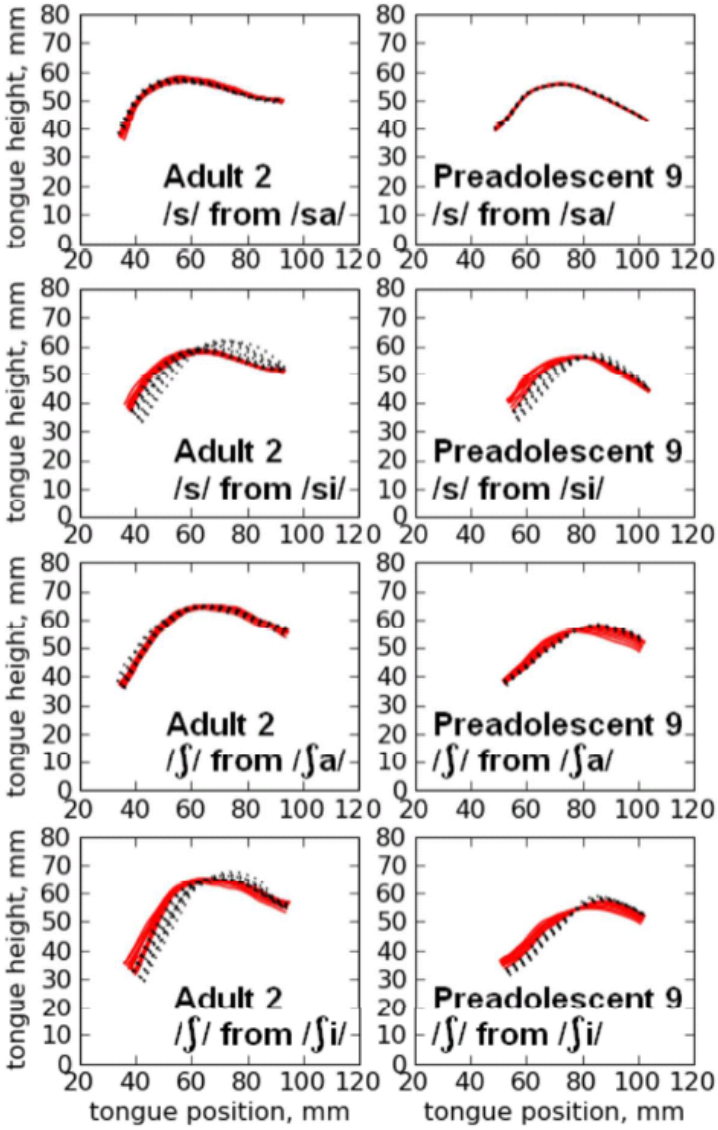

tongue height, mm
tongue position, mm
Adult 2
/s/ from /sa/
Preadolescent 9
/s/ from /sa/
Adult 2
/s/ from /si/
Preadolescent 9
/s/ from /si/
Adult 2
/ʃ/ from /ʃa/
Preadolescent 9
/ʃ/ from /ʃa/
Adult 2
/ʃ/ from /ʃi/
Preadolescent 9
/ʃ/ from /ʃi/

Figure 2

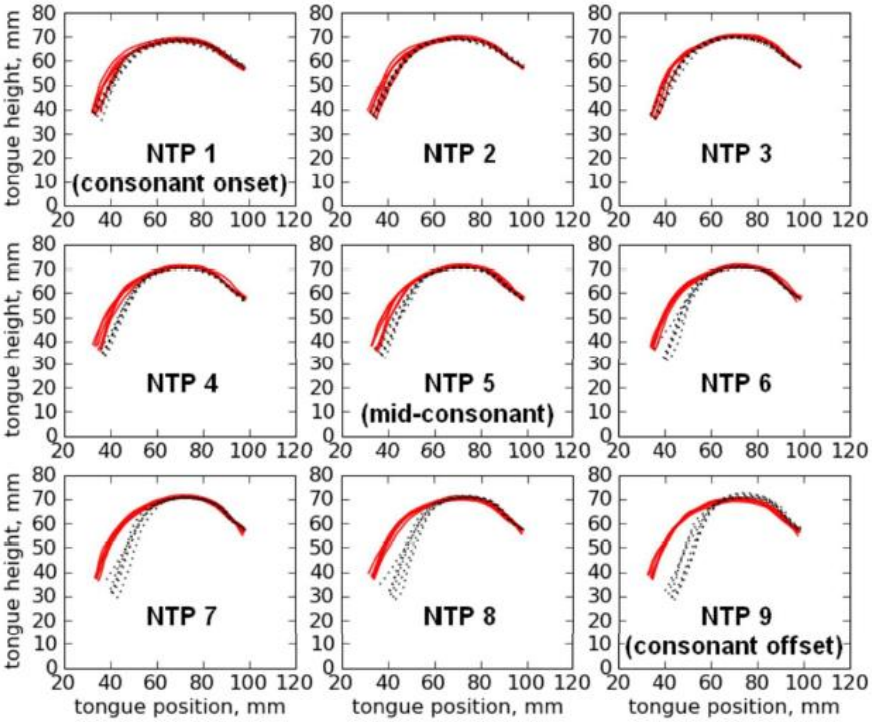

NTP 1
(consonant onset)
NTP 2
NTP 3
NTP 4
NTP 5
(mid-consonant)
NTP 6
NTP 7
NTP 8
NTP 9
(consonant offset)
tongue height, mm
tongue position, mm

Figure 3

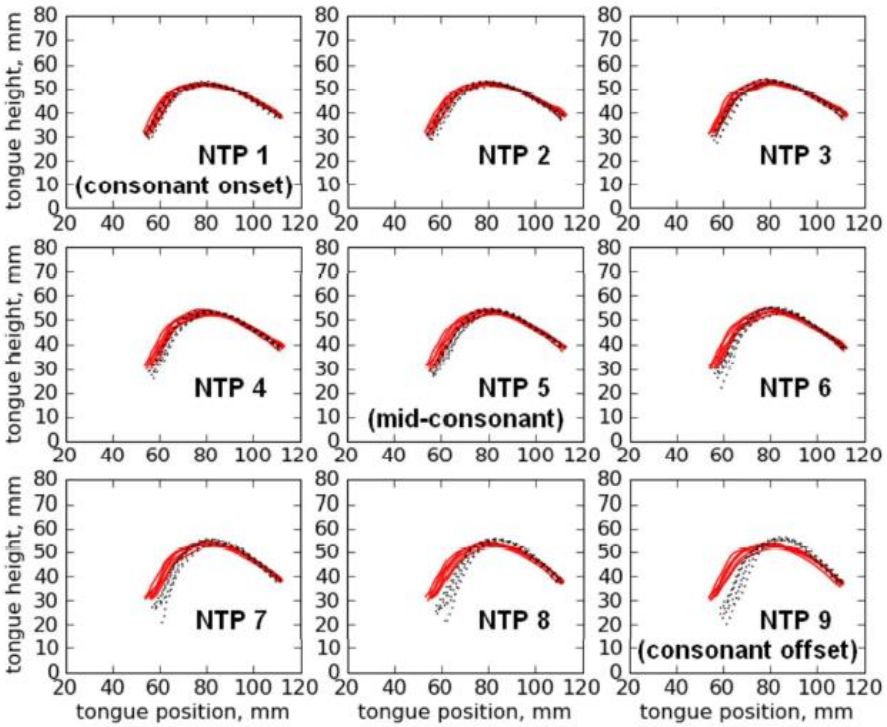

NTP 1
(consonant onset)
NTP 2
NTP 3
NTP 4
NTP 5
(mid-consonant)
NTP 6
NTP 7
NTP 8
NTP 9
(consonant offset)
tongue height, mm
tongue position, mm

Figure 4

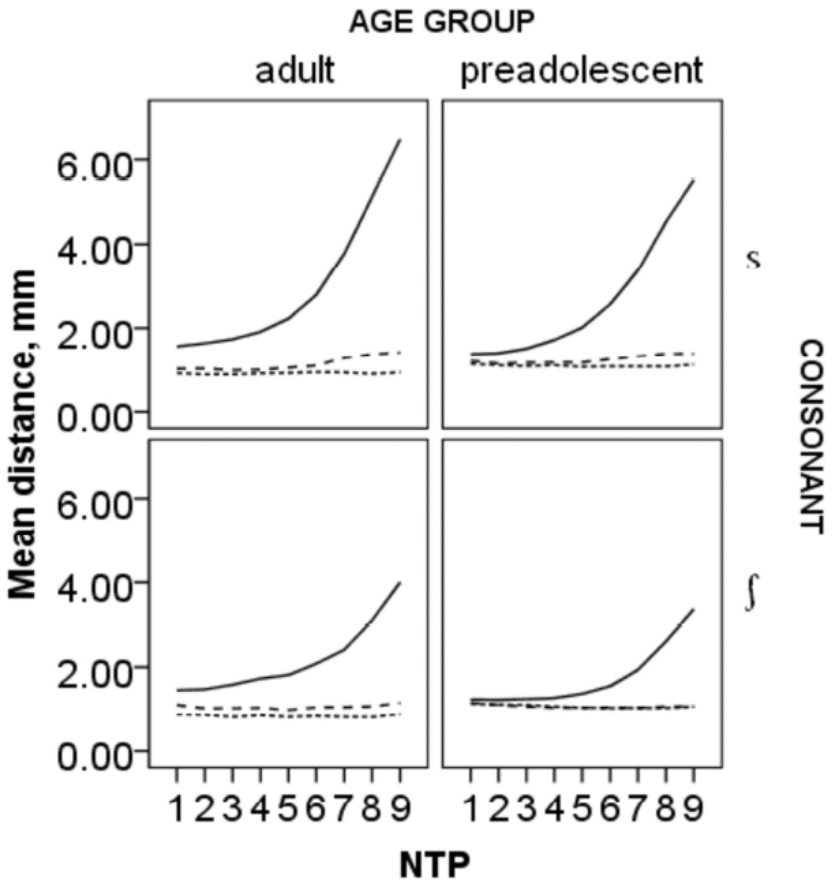
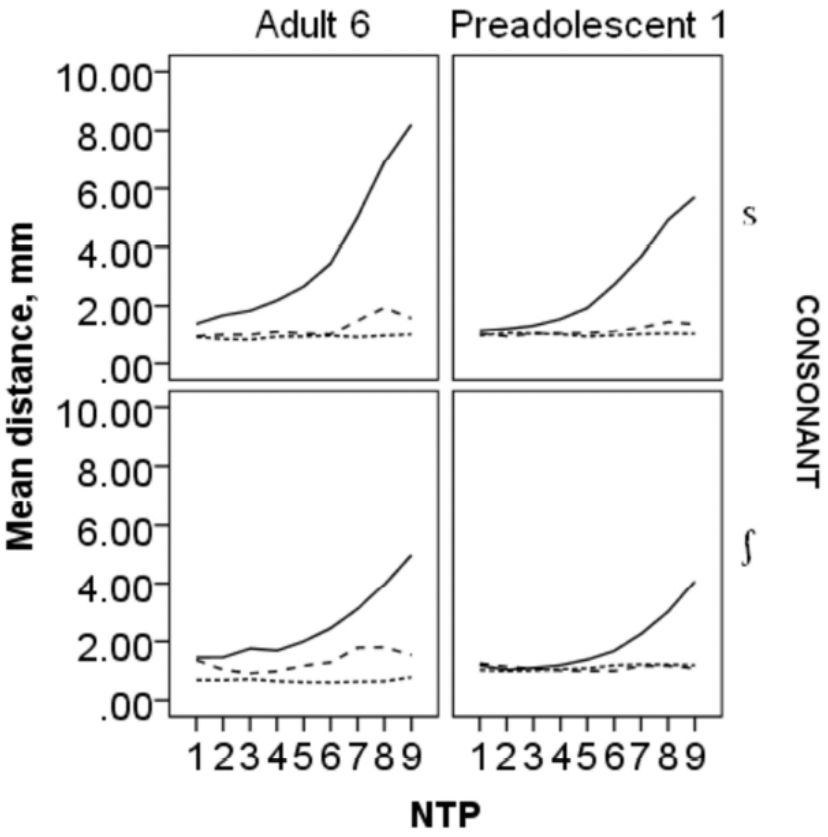

AGE GROUP
adult
preadolescent
Mean distance, mm
6.00
4.00
2.00
0.00
1 2 3 4 5 6 7 8 9
NTP
s
ʃ
CONSONANT

Figure 5


Adult 6
Preadolescent 1
Mean distance, mm
10.00
8.00
6.00
4.00
2.00
.00
s
ʃ
CONSONANT
1 2 3 4 5 6 7 8 9
NTP