Development of lingual coarticulation and articulatory constraints between childhood and adolescence: an ultrasound study

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

Queen Margaret University, Edinburgh, UK

nzharkova@qmu.ac.uk, rlickley@qmu.ac.uk, wjhardcastle@gmail.com

Abstract

This study used ultrasound tongue imaging in order to compare lingual coarticulatory patterns in 13-year-old adolescents and 5-year-old children. Previous articulatory work has reported age-related differences in consonant-specific coarticulation. According to the Degree of Articulatory Constraint model, speech segments with more constrained tongue dorsum have more resistance to lingual coarticulation in adult speech. Bilabial and postalveolar consonants are at opposing ends of a spectrum, with the latter being strongly resistant to vowel-related lingual coarticulation. In this study, tongue shapes for the bilabial stop and the postalveolar fricative were compared across contrasting vowel contexts. For both groups of speakers, there were robust differences between the two consonants, with the bilabial adapting to the vowel influence more than the postalveolar. We conclude that the DAC model can predict some lingual coarticulatory patterns in children as well as in adolescents.

Keywords**:** lingual coarticulation, articulatory constraints, children, adolescents, development

1. Introduction

Evidence from a number of acoustic and articulatory studies has shown that lingual coarticulation is present in children from an early age (e.g., Sereno et al. 1987; Nittrouer et al. 1989; 1996; Katz et al. 1991; Sussman et al. 1992; Zharkova et al. 2011; 2012; Noiray et al. 2013; Song et al. 2013). An influential theory of coarticulation development (Nittrouer et al. 1989) claims that motor planning of speech in young children is carried out at the syllabic level. This idea is based on the observed reduction with age of the extent of anticipatory vowel-on-consonant coarticulation within a consonant-vowel (CV) syllable, as shown by acoustic analysis. This theoretical approach contrasts with the view that children coarticulate less than adults, due to the overly segmental style of speech planning in young children (Kent 1983). Acoustic studies on the topic have produced conflicting results (e.g., no significant differences in lingual coarticulation between children and adults were reported in Katz et al. 1991). There is some support for Nittrouer's theory from articulatory research (e.g., Zharkova et al. 2011), though previous articulatory work has also reported age-related differences in consonant-specific coarticulation (Katz & Bharadwaj 2001; Zharkova et al. 2012; in press). In the present study, the claim by Nittrouer et al. was explored in the context of an approach to lingual coarticulation based on articulatory constraints. According to the Degree of Articulatory Constraint (DAC) model (Recasens et al. 1997), speech segments with more constrained tongue dorsum have more resistance to lingual coarticulation. Bilabial and postalveolar consonants are at opposing ends of a spectrum, with the latter being strongly resistant to vowel-related lingual coarticulation. Our study used ultrasound tongue imaging to compare lingual coarticulatory patterns in 13-year-old adolescents and 5-year-old children, using two consonants which differ substantially in their DAC properties.

In order to analyse articulatory data from our 5-year-old speakers numerically, we needed a way of quantification that would not be based on a set of tongue curves located in the same coordinate space. The 13-year-olds in our study wore a custom-designed headset (Articulate Instruments Ltd 2008), to stabilise the transducer in relation to the head. This headset was not suitable for the 5-year-old participants. A way to obtain quantitative information in the absence of transducer stabilisation is to use measurements based on a single curve (e.g., Bressmann et al. 2005; Ménard et al. 2012; Stolar & Gick 2013). In order to have a reliable way of comparing between groups, our study required a measure that would provide the same results regardless of whether the head was stabilised or not. To quantify lingual coarticulation in this study, we used the Tongue Constraint Position Index (TCPI), an index of tongue shape developed in Zharkova (2013a). TCPI represents the location of the most "bunched" part of the tongue, with its negative or positive values occurring when the place of maximal excursion of the tongue is further backward or further forward along the tongue contour, respectively, in relation to the middle of the straight line between two ends of the curve. An index value can be obtained from a single tongue curve, which makes it possible to use on the data collected without head stabilisation. The TCPI has been shown to capture vowel-induced lingual coarticulation on consonants (Zharkova 2013b), and to be unaffected by whether participants wear the headset or not (Zharkova & Hewlett 2013).

In both age groups, higher TCPI values were expected in the context of /i/ than in the context of /a/. In the adolescents, more coarticulation was expected in /p/ than in /ʃ/, consistent with the DAC model predictions for adult speech. If more coarticulation was observed in /p/ than in /ʃ/ for the children, then the findings would also be consistent with the DAC model. If the children had more coarticulation than the adolescents, this would provide support for the theory by Nittrouer et al. If the relative difference in amount of coarticulation between the two consonants did not pattern in the same way in the two age groups then we could conclude that the effect of articulatory constraints on lingual coarticulation is not uniform throughout the developmental process from childhood to adolescence.

2. Method

The study used ultrasound tongue movement data (sampling rate 100 Hz) synchronised with the acoustic signal. CV syllables with the consonants /p/ or /ʃ/ and the vowels /a/ or /i/ were produced in the carrier phrase "It's a ..., Pam" (each target repeated five times) by two groups of speakers of

Scottish Standard English. One group consisted of ten adolescents aged between 13 years 0 months and 13 years 11 months. The other group consisted of ten children aged between 5 years 0 months and 5 years 11 months. The adolescent participants wore a headset to stabilise the ultrasound transducer in relation to the head. For the child participants, the transducer was hand-held by the experimenter. Articulate Assistant Advanced software (Articulate Instruments Ltd 2012) was used to collect and analyse the data. All participants from the younger group were video recorded during the data collection using a separate channel of the multichannel ultrasound system. The video data were collected in two planes: *en face* and in profile. In the data analysis, all recordings from 5-year-olds were examined in order to ensure that during the target CV sequence the transducer was relatively stable under the chin, and that a midsagittal tongue image was present. The tokens which did not satisfy these conditions were excluded from the analysis (nine tokens in total). Additionally, the /ʃ/ tokens produced by one child (Child 7) were realised as unaspirated dental stops. As the measurements in this study were focused on tongue shape, and the dental stops may have a noticeably different tongue shape from the postalveolar fricative, all /ʃ/ tokens produced by this child were excluded from all analyses.

2.1. Data analysis

In each CV token, an annotation point was placed at mid-consonant (mid-closure for /p/), based on the acoustic data. Cubic splines were fitted to midsagittal tongue curves at each annotation point. The TCPI, which is a ratio, was calculated for each token. The denominator in the ratio equalled half the length of the straight line *n* between two ends of the tongue curve. The numerator was the distance between the longest perpendicular from *n* to the tongue curve and the perpendicular to the tongue curve traced from the midpoint of *n*.

Linear mixed models (LMMs) were performed in R (R Development Core Team 2012), with the lmer software package (Baayen 2008). In order to establish whether the contrasting vowels influenced the tongue shape at mid-consonant, a LMM was carried out for each age group, with vowel as a fixed factor. To find out which consonant was more affected by the following vowels, a LMM was run for each age group with vowel and consonant as fixed factors. In this analysis, a significant interaction between consonant and vowel would mean that the two consonants have different amounts of coarticulation. To establish whether the 5-year-olds coarticulate more than the 13-year-olds, the model included vowel and age as fixed factors. A significant interaction in this model, accompanied by a larger difference between the two vowel contexts in the 5-year-olds than in the 13-year-olds, would be taken as evidence of more coarticulation in the 5-year-olds. Finally, a LMM was run with vowel, consonant and age group as fixed factors. A significant interaction between the three factors would suggest that articulatory constraints on the tongue dorsum affect coarticulatory properties in young children differently to adolescents. In all statistical tests, Speaker was modelled with both random slopes and random intercepts. Following Zharkova et al. (in press), the outcome was deemed significant at the 0.01 level if the *F* value in the ANOVA exceeded 8.49.

3. Results

Figure 1 shows tongue contours from representative speakers from each age group. It is immediately noticeable that the tongue curves for the 5-year-old child are much more variable in absolute position across repetitions than in the adolescent. This difference was expected, given the recording setup for the 5-year-olds, with the handheld transducer. Despite this large variability in absolute tongue position, the tongue shapes for /p/ in the two vowel contexts are clearly different for the 5-year-old in Figure 1, with the bunched shapes in the context of /i/, and the “hump” located towards the anterior part of the tongue. In the context of /a/, the tongue shapes for the 5-year-old’s /p/ are noticeably flatter than in the context of /i/, reflecting the tongue dorsum lowering for producing the low vowel. Similar patterns are observed for the adolescent. The fact that all the curves for the adolescent are located within one coordinate plane, due to the head-to-transducer stabilisation, lets us see that the absolute tongue root and tongue blade positions for the two vowel contexts are quite different. The shape of the tongue, regardless of the absolute position of the tongue curves, is different across the two vowel contexts, with the “hump” location appearing more anterior in the context of /i/ and more posterior in the context of /a/.

For /ʃ/, for both speakers and in both vowel contexts the “hump” is located towards the anterior part of the tongue, so the tongue shapes for /ʃ/ are resembling those for /p/ in the context of /i/ more than of those for /p/ in the context of /a/.

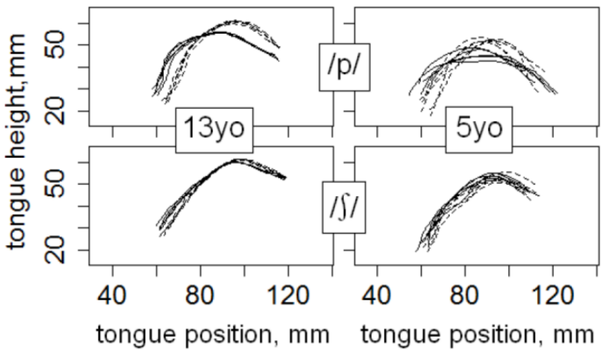

Figure 1: *Tongue curves in the two vowel contexts (solid lines for the context of /a/; dashed lines for the context of /i/).*

Mean values of the TCPI for individual speakers are presented in Table 1. For /p/, the difference in tongue shape across the two vowel contexts, noticeable in Figure 1, was observed across speakers and age groups. All adolescent and child speakers had a larger TCPI value for /p/ in the context of /i/ than in the context of /a/. For /ʃ/, consistent with the observation from Figure 1 of rather similar tongue shapes in the two vowel contexts, only seven adolescents and five 5-year-olds had a larger TCPI value in the context of /i/.

Table 1: *TCPI values in the two vowel contexts for each participant, rounded to two decimal places.*

Speaker	Cons.	13-year-olds		5-year-olds	
		/a/	/i/	/a/	/i/
1	/p/	-0.24	0.12	-0.02	0.23
	/ʃ/	0.25	0.27	0.34	0.33
2	/p/	-0.17	0.21	-0.14	0.13
	/ʃ/	0.43	0.44	0.37	0.31
3	/p/	-0.36	0.18	0.09	0.14
	/ʃ/	0.30	0.35	0.11	0.11
4	/p/	0.20	0.31	-0.07	0.11
	/ʃ/	0.44	0.47	0.17	0.22
5	/p/	0.15	0.37	-0.19	0.12
	/ʃ/	0.32	0.39	0.31	0.36

6	/p/ /ʃ/	-0.05 0.33	0.14 0.28	0.17 0.33	0.20 0.29
7	/p/ /ʃ/	-0.16 0.48	0.22 0.47	0.21 0.23	0.42 0.23
8	/p/ /ʃ/	-0.31 0.18	0.07 0.27	0.10 0.18	0.25 0.31
9	/p/ /ʃ/	-0.02 0.05	0.11 0.00	-0.12 0.22	0.24 0.29
10	/p/ /ʃ/	-0.36 0.04	-0.10 0.10	0.01 0.42	0.17 0.26
Mean	**/p/** **/ʃ/**	**-0.10** **0.28**	**0.16** **0.30**	**0.01** **0.27**	**0.20** **0.27**

3.1. Within-group and across-group comparisons

Figure 2 presents TCPI values for each age group, consonant and vowel context. The TCPI values for /p/ in the context of /a/ stand out, for both groups, in that they are visibly smaller than all the other values. This is consistent with the difference in tongue shapes that can be observed in Figure 1.

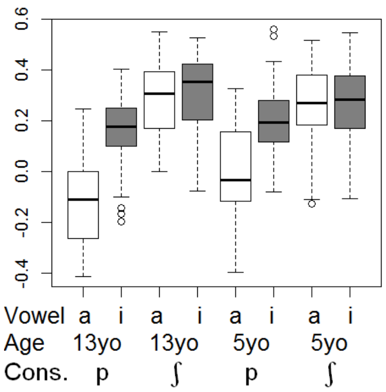

Figure 2: *TCPI values (the boxes for /i/ are shaded in grey). The bottom, central and top lines of each box represent the 25th, 50th and 75th percentiles, respectively; the small circles are outliers.*

In the analysis of the adolescent data, there was a significant effect of vowel on TCPI (*F* = 31.45), showing evidence of lingual coarticulation at mid-consonant. In the model including both vowel and consonant as fixed factors, there was a significant interaction (*F* = 74.72), with the difference in TCPI values between the context of /a/ and the context of /i/ being larger for the stop (see Figure 2). This interaction provides evidence of the adolescents having adult-like coarticulatory patterns, with the more constrained postalveolar fricative showing less coarticulation than the less constrained bilabial stop. In individual models for the two consonants separately, the effect of the vowel context on TCPI at mid-consonant was significant for /p/ (*F* = 61.41), but not for /ʃ/ (*F* = 2.11).

For the five-year-olds, the vowel context had a significant effect on TCPI at mid-consonant (*F* = 15.98). When both vowel and consonant were included as fixed factors, there was a significant interaction (*F* = 28.64), with the difference in TCPI values between the two vowel contexts being larger for /p/, similarly to the adolescents (see Figure 2). In individual models for the two consonants, again similarly to the adolescents, the effect of the vowel context on TCPI was significant for /p/ (*F* = 32.96), but not for /ʃ/ (*F* = 0.07).

There was no significant interaction in the model including vowel and age group as fixed factors, for both consonants pooled together (*F* = 0.95), or for individual consonants (/p/: *F* = 1.85; /ʃ/: *F* = 0.26). In the model including vowel, consonant and age group as fixed factors, there was no three-way interaction (*F* = 1.41).

4. Discussion

This study showed that for both 13-year-old adolescents and 5-year-old children, contrasting vowels /a/ and /i/ have an effect on the tongue shape at mid-consonant. These results agree with previous findings of lingual coarticulation in young children (Katz et al. 1991; Sussman et al. 1992; Noiray et al. 2013). We have also reported robust differences in coarticulatory patterns across consonants for the adolescents and for the younger children. The similarity of tongue shapes for /ʃ/ across vowel contexts was clear in both groups of speakers, while for /p/, the difference in tongue shapes across vowel contexts was equally clear for both groups.

The difference for the adolescents between the two consonants in the amount of tongue shape adaptation to the following vowel was consistent with the DAC model predictions, with the bilabial stop being significantly more affected than the postalveolar fricative. For the 5-year-olds, the difference between the two consonants in the amount of tongue shape adaptation to the vowel was in the same direction, and also significant. Thus we can conclude that the DAC model, as far as concerning the difference in articulatory constraint on the tongue between bilabials and postalveolars, can predict lingual coarticulatory behavior in 5-year-old children.

One possible reason for the lack of any significant differences between the two age groups is that there was much more variation in absolute tongue positions for the 5-year-olds than for the 13-year-olds, because of the hand-held transducer. In order to find out whether the difference in the recording setup could have been a contributing factor to the across-group results, TCPI was compared across age group and vowel context for /p/, using the data for the 5-year-olds and the data for the same 13-year-old speakers collected without the headset. Those data were collected from the 13-year-olds within a larger project in order to investigate how the presence or absence of head-to-transducer stabilisation may affect quantitative measures of tongue shape and position (see Zharkova & Hewlett 2013). The results of the across age group comparison were the same as the results reported in Section 3.1, with no significant interaction between vowel and age group (*F* = 0.33).

In addition to the lack of significant across-group differences in the extent of lingual coarticulation, there was no significant effect of the vowel context on tongue shape for the postalveolar consonant. This does not necessarily contradict previous studies reporting a decrease of the extent of coarticulation with increasing age, such as an acoustic study by Nittrouer et al. (1996) and an ultrasound study by Zharkova et al. (2011). In the former study, the measurement point was located later in the consonant than in the present study, so at mid-consonant any age-related difference may not have developed yet. The latter study used a different method to ascertain the presence of coarticulation, the nearest neighbour distance method of comparing two sets of tongue curves (Zharkova & Hewlett 2009). It is possible that the TCPI index is not sufficiently robust to detect some coarticulation-related differences that can be captured using measurements based on the data from whole curves and/or measurements comparing absolute positions of different curves. For example, a difference for the preadolescent /ʃ/ in the tongue root area between tongue curves in the two vowel contexts (as indicated in Figure 1) could be quantified by a measure specifically targeting the tongue root. This difference would also be likely to show in an analysis based on comparing sets of tongue curves which uses information from the whole of the visible tongue contour (e.g. the nearest neighbor distance method, or SSAnova, Davidson 2006). Indeed, Zharkova (2013c) has compared /ʃ/ curves from these preadolescent speakers using the nearest neighbour method, and reported a significant difference between the curve sets for /ʃ/ in the two vowel contexts. While such measures have a number of advantages over single curve based measures (they are not based on the visible ends of the tongue curve, and they make use of the absolute location of a set of curves), they are not applicable to the data from multiple tongue curves which are not located within the same coordinate space, hence in the present study we used a method that could provide us with robust results on tongue shape in the absence of head-to-transducer stabilisation. We are currently investigating whether other single curve based measures, which can potentially capture more relevant aspects of the tongue shape, can be reliably used to analyse lingual articulation regardless of whether the speaker is wearing a transducer stabilising headset.

5. Acknowledgements

This work was supported by a grant from the Economic and Social Research Council (ES/K002597/1).

6. References

Articulate Instruments Ltd (2008). *Ultrasound Stabilisation Headset Users Manual: Revision 1.4*. Edinburgh, UK: Articulate Instruments Ltd, 21 p.

Articulate Instruments Ltd (2012). *Articulate Assistant Advanced Ultrasound Module User Guide: Version 2.14*. Edinburgh, UK: Articulate Instruments Ltd, 69 p.

Baayen, R. H. (2008). *Analyzing Linguistic Data. A Practical Introduction to Statistics Using R*. Cambridge: Cambridge University Press, pp. 241–302.

Bressmann, T., Thind, P., Uy1, C., Bollig, C., Gilbert, R.W., and Irish, J.C. (2005). "Quantitative three-dimensional ultrasound analysis of tongue protrusion, grooving and symmetry: Data from 12 normal speakers and a partial glossectomee". *Clinical Linguistics and Phonetics*, **19**, 573–588.

Davidson, L. (2006). "Comparing tongue shapes from ultrasound imaging using smoothing spline analysis of variance". *Journal of the Acoustical Society of America* , **120**, 407–415.

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

Kent. R. D. (1983). "The segmental organization of speech". In: *The Production of Speech*, edited by P. F. MacNeilage. New York: Springer-Verlag, pp. 57–89.

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

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

Nittrouer, S., Studdert-Kennedy, M., and 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., and Iskarous, K. (2013). "The development of motor synergies in children: Ultrasound and acoustic measurements". *Journal of the Acoustical Society of America*, **133**, 444–452.

R Development Core Team. (2012). *R: a language and environment for statistical computing.* Vienna: R Foundation for Statistical Computing. http://www.R-project.org.

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

Sereno, J. A., Baum, A. R., Cameron Marean, G., and 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.

Song, J. Y., Demuth, K., Shattuck-Hufnagel, S., and Ménard, L. (2013). "The effects of coarticulation and morphological complexity on the production of English coda clusters: Acoustic and articulatory evidence from 2-year-olds and adults using ultrasound". *Journal of Phonetics*, **41**, 281–295.

Stolar, S. and Gick, B. (2013). "An index for quantifying tongue curvature". *Canadian Acoustics*, **41**, 11-15.

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

Zharkova, N. (2013a). "Using ultrasound to quantify tongue shape and movement characteristics". *The Cleft Palate-Craniofacial Journal*, **50**, 76–81.

Zharkova, N. (2013b). "A normative-speaker validation study of two indices developed to quantify tongue dorsum activity from midsagittal tongue shapes". *Clinical Linguistics & Phonetics*, **27**, 484–496.

Zharkova, N. (2013c). "Lingual coarticulation and articulatory constraints in adolescents". Oral presentation at Ultrafest VI (Ultrasound meeting), Edinburgh University, Edinburgh UK, 6–8 November 2013.

Zharkova, N. and 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. and Hewlett, N. (2013). "Quantitative analysis of ultrasound data collected with and without head stabilisation". Poster at Ultrafest VI (Ultrasound meeting), Edinburgh University, Edinburgh UK, 6–8 November 2013.

Zharkova, N., Hewlett, N., and 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., and 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.

Zharkova, N., Hewlett, N., Hardcastle, W. J., and Lickley, R. (in press). "Spatial and temporal lingual coarticulation and motor control in preadolescents". *Journal of Speech, Language and Hearing Research*. Retrieved from http://jslhr.pubs.asha.org/Article.aspx?articleid=1832512.