DEVELOPMENT OF LINGUAL MOTOR CONTROL IN CHILDREN AND ADOLESCENTS

Natalia Zharkova[1], William J. Hardcastle[1], Fiona E. Gibbon[2] & Robin J. Lickley[1]

[1]CASL Research Centre, Queen Margaret University, Edinburgh
[2]Department of Speech and Hearing Sciences, University College Cork
nzharkova@qmu.ac.uk; wjhardcastle@gmail.com; f.gibbon@ucc.ie; rlickley@qmu.ac.uk

ABSTRACT

An important insight into speech motor control development can be gained from analysing coarticulation. Despite a growing number of acoustic and articulatory studies of lingual coarticulation in children, there are conflicting opinions on how the extent of coarticulation changes during childhood. There is also increasing evidence that age-related patterns vary depending on speech sounds involved. The present study employed ultrasound tongue imaging to compare anticipatory V-on-C coarticulation in 13-year-old adolescents and 5-year-old children, using the consonants /p/ and /t/, which differ in the amount of lingual coarticulation in adult speech. For /p/, the two groups had a similar amount of coarticulation. For /t/, both groups had a vowel effect on the extent of tongue bunching, while only adolescents had an effect on the location of tongue bunching. Token-to-token variability in absolute tongue position was larger in the 5-year-olds. We discuss the findings in relation to previous studies and existing theories.

Keywords: coarticulation, motor control, children, adolescents, speech production

1. INTRODUCTION

Speech motor control development in children has been investigated over the years in acoustic and articulatory studies. A large area of research that informs theories of motor control is coarticulation. The question of how children reach adult-like control of articulators has been asked in a number of studies using different methodologies (e.g., [10, 20, 14, 9, 23, 15, 13, 6, 12, 29, 16, 22, 30, 31]). Perhaps partly due to this methodological diversity, there are conflicting opinions on how the extent of coarticulation changes during childhood.

One influential point of view is that children coarticulate less than adults, suggesting a segment-by-segment style planning ([10]). Another view is that children plan speech in syllable-sized units and, hence, coarticulate within the syllable more than adults ([14, 15]). There is also evidence from studies that have not reported measurable differences in coarticulation between adults and children (e.g, [9, 13]). Finally, age-related patterns of coarticulation have been shown to vary depending on the nature of speech segments ([23, 8, 29, 16, 30, 31]). For example, 4-to-5-year-old children have been reported to have less coarticulation for alveolar than for labial stops, like adults ([23, 16]). On the other hand, 6-to-9-year-old children, unlike adults, did not show evidence of vowel-on-/s/ coarticulation in [29].

In the present study, we employed ultrasound tongue imaging to compare vowel-on-consonant anticipatory lingual coarticulation in 13-year-olds and 5-year-olds, using the consonants /p/ and /t/. Our choice of the measurement time point was guided by previous studies. Measuring at the consonant offset, as in locus equation studies [23] and [16], would likely show adult-like patterns of coarticulation in both groups of speakers, while we were interested in documenting age-related differences. The time point of 30 ms before the consonant offset, used for fricative F2 analysis in [14] and [15], would not be practical for stops, as it would likely correspond to inconsistent acoustic events across speakers and consonants. Instead, we chose mid-closure for measurements because, based on [29] and [31], we could expect some age-related consonant-specific differences at this time point.

For the adolescents, coarticulatory effects were expected for both /p/ and /t/, with a larger effect on /p/ (see [30, 31]). The 5-year-olds were hypothesised to coarticulate /p/ (based on [31]), but not /t/ (based on [29]). More versus less coarticulation in the younger speaker group would support the “syllabic” [14] versus “segmental” [10] planning points of view, respectively. Finally, based on previous studies that have reported large variability as an indication of immature motor control in children (e.g., [23, 21, 29]), we expected to observe larger within-speaker variability in tongue position in the younger age group.

2. METHOD

2.1. Participants and data collection

The participants were speakers of Scottish Standard English, ten adolescents aged between 13;0

[years;months] and 13;11 (six girls), and ten children aged between 5;5 and 5;11 (five girls). The children were judged by a speech and language therapist to have typically developing speech.

Ultrasound tongue movement data were collected at 100 Hz, synchronised with the acoustic signal sampled at 22050 Hz. CV syllables with the consonants /p/ or /t/ and the vowels /a/ or /i/ were produced in the carrier phrase “It’s a ..., Pam” (each target repeated five times) by the two groups. In all recordings, the ultrasound transducer was hand-held by the experimenter (cf. [31], where child hand-held data were compared with adolescent head-to-transducer stabilised data). Because of the age of the younger children, and based on previous research, we were unable to use the typical setup including head-to-transducer stabilisation that we would have used with older speakers, as it involves the need to wear a headset that is generally too heavy and uncomfortable for 5-year-olds. We feel justified in recording the speakers without head stabilisation because the study used recently identified measures of coarticulation that produce the same results for non-stabilised data and for stabilised data ([27]).

The data were collected using Articulate Assistant Advanced software ([1]), and this software was also used to trace tongue curves. All participants were video recorded *en face* and in profile during the data collection, using a separate channel of the multichannel ultrasound system (see Fig. 1). All recordings were examined in order to ensure that the transducer was relatively stable under the chin during the production of the target CV sequence, and that a midsagittal tongue image was present, along with the shadow of the chin and the shadow of the hyoid bone. These conditions were not satisfied in five tokens of /p/ produced by four children from the younger group: two tokens of /p/ from /pa/ (5yo3 and 5yo7) and three tokens of /p/ from /pi/ (two by 5yo4 and one by 5yo5). These five tokens were excluded from the analysis.

Figure 1: A video frame from the recording software, with combined views of a 5-year-old participant in two planes during the data collection.

2.2. Data analysis

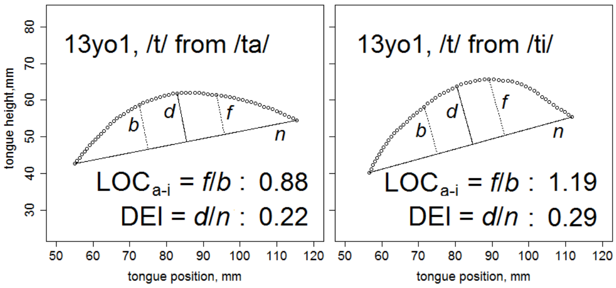
For both stops, tongue shapes were traced at mid-closure automatically, with some manual correction. Text files with *xy* coordinates of the tongue curves were exported from Articulate Assistant Advanced, for further analyses in R ([17]).

2.2.1. Measurements of tongue shape

$LOC_{a\text{-}i}$ index ([27]) was used to quantify differences in tongue shapes depending on whether the following vowel was /a/ or /i/. This index was chosen for across-group comparisons in this study based on previous work. In [27], adolescent speakers were recorded producing the same stimuli with and without head stabilisation, and $LOC_{a\text{-}i}$ had a consistent performance across conditions for both V-on-/p/ and V-on-/t/ anticipatory coarticulation, even though for the former consonant different amounts of the tongue curve were imaged in the two stabilisation conditions.

The calculations of $LOC_{a\text{-}i}$ are illustrated in Fig. 2, which shows tongue shapes for /t/ in the two vowel contexts. $LOC_{a\text{-}i}$ is a ratio of the straight line *f* (a perpendicular from one third of line *n*, starting from the front, to the tongue curve) to line *b* (a perpendicular from two thirds of line *n* to the tongue curve). In the context of /i/, for both /p/ and /t/, $LOC_{a\text{-}i}$ has higher values than in the context of /a/ in adults and adolescents without speech disorders ([26, 27]). This pattern is demonstrated in Fig. 2 for an adolescent from the present study.

Figure 2: Tongue curves for /t/ from /ta/ (left) and /t/ from /ti/ (right), produced by adolescent speaker 13yo1, illustrating the two indices. The front of the tongue is on the right in this figure and in Fig. 3.



For the alveolar consonant, we also quantified the extent of excursion of the middle portion of the tongue, using the Dorsum Excursion Index (DEI, [25]). In [27], DEI was shown to produce the same results on coarticulation in adolescents across different stabilisation conditions for /t/, but not for /p/. This suggests that comparing tongue shape during bilabials produced by two groups of speakers without head stabilisation may not yield reliable results for DEI, due to a possible influence from the length of the imaged curve. Therefore in the present study we used DEI for analysing /t/, but not /p/. The calculations are illustrated in Fig. 2. DEI is a ratio of line *d* (a perpendicular to the tongue curve from mid-*n*) to line *n*. DEI has higher values for /t/ in the context of /i/ than in the context of /a/ for typical adults and adolescents ([26, 27]).

2.2.2. Token-to-token variability

For each target consonant in each vowel context, within-set (WS) nearest neighbour distances ([28]) were calculated between curves for the five tokens produced by each 13-year-old speaker (see [27]) and each 5-year-old speaker. These WS distance values were used to compare token-to-token variability in absolute tongue position across age groups. In the event that WS distances were significantly greater in adolescents (due to developmental differences in vocal tract length, [3]), normalisation across speakers for vocal tract size, based of relative length of tongue contour, would be carried out before across-group comparisons.

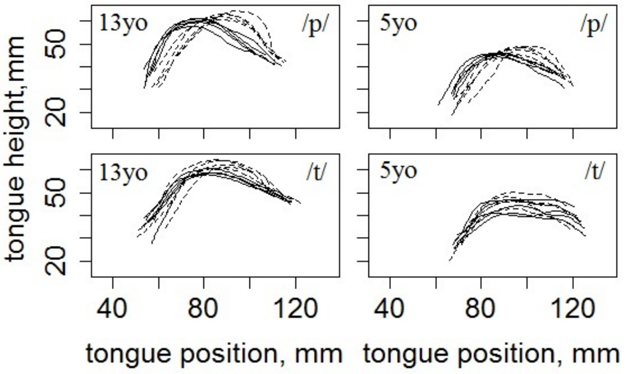
2.2.3. Statistical analyses

Linear Mixed Models (LMMs, [2]) were used for inferential statistical analyses, including random intercept and slope for speaker. Following [30], the *F* value in the ANOVA needed to exceed 8.49 for an effect to be deemed significant at the 0.01 level. To establish the presence of a coarticulatory effect, LMMs on tongue shape indices were carried out within consonant and age group, with vowel as a fixed factor. In the event of a significant effect on $\mathrm{LOC}_{a\text{-}i}$ for both consonants in the adolescents, a LMM was carried out with consonant and vowel as fixed factors, and a significant interaction would indicate a cross-consonant difference in the extent of coarticulation. In the event of a significant effect for a given consonant in both age groups, interaction of age group and vowel in a larger model was used to establish any age-related difference.

3. RESULTS

Fig. 3 has curves from one representative participant per age group. Visual inspection shows that in both speakers, tongue curves for /p/ in the contrasting vowel contexts are, not surprisingly, more differentiated in shape than tongue curves for /t/. For /p/, the most bunched part of the tongue in the context of /i/ is clearly further forward along the tongue curve than in the context of /a/. For /t/ in the adolescent, the shape of the tongue appears to differ both in terms of the extent of bunching (larger in the context of /i/) and in terms of the location of the most bunched part of the tongue in relation to the ends of the curve (further forward in the context of /i/). In the younger child, the curves in the context of /i/ appear to be generally somewhat more bunched than in the context of /a/.

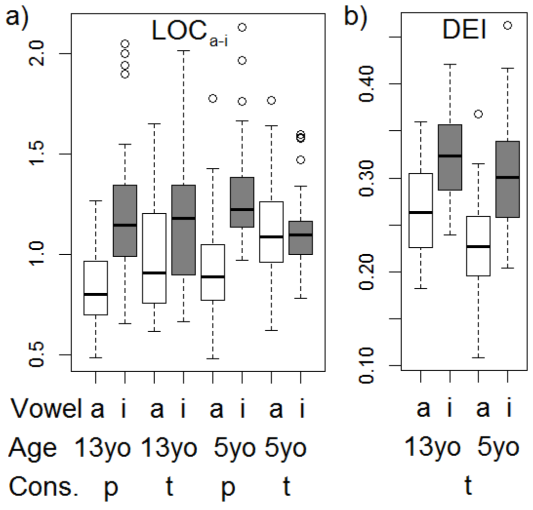
Figure 3: Tongue curves for five repetitions of /p/ and /t/ in the two vowel contexts, produced by one representative participant from each age group. Solid lines – /a/ context, dashed lines – /i/ context.



3.1. Presence of coarticulatory effects

Fig. 4 has $\mathrm{LOC}_{a\text{-}i}$ values for both consonants and DEI values for /t/, in the two age groups and vowel contexts. The figure confirms visual observations from Fig. 3, showing noticeable vowel-related differences in tongue shape at mid-/p/ in both age groups ($\mathrm{LOC}_{a\text{-}i}$ in Fig. 4a). For /t/, Fig. 4b shows vowel-related differences in DEI for both age groups. $\mathrm{LOC}_{a\text{-}i}$ results for /t/ in Fig. 4a show no vowel-related difference for the 5-year-olds, while for 13-year-olds any vowel-related difference is considerably smaller than for the bilabial stop.

Figure 4: Group averages for (a) $\mathrm{LOC}_{a\text{-}i}$ and (b) DEI values. The boxes for /i/ are shaded.



Results from LMMs testing for the presence of the vowel effect can be found in Table 1. For the adolescents there was a significant vowel effect on LOC_{a-i} for both consonants, and also on DEI for /t/. For the 5-year-olds, there was a significant effect on LOC_{a-i} for /p/ and on DEI for /t/.

Table 1: *F* values from LMMs testing for the presence of the vowel effect on the consonant, within age group. The values that were significant at the 0.01 level are accompanied by an asterisk.

	/p/	/t/	
	LOC_{a-i}	LOC_{a-i}	DEI
13yo	32.83 *	16.03 *	23.79 *
5yo	27.15 *	0.02	19.93 *

3.2. Size of coarticulatory effects

For the adolescents, in a LMM on LOC_{a-i} there was a significant interaction between consonant and vowel ($F = 32.40$). The direction of the difference conforms to our prediction. It is illustrated by the ratios of the mean group value in the context of /i/ to that in the context of /a/: for /p/ the ratio was 1.46, and for /t/ it was 1.17. In LMMs on LOC_{a-i} for /p/ and on DEI for /t/, there was no significant interaction of vowel and group in either case.

3.3. Token-to-token variability

Table 2 shows mean WS distance values. There was a significant effect of age group on WS distances ($F = 30.99$), with larger values for the younger group, so we can conclude without across-speaker normalisation that 5-year-olds were more variable in absolute tongue position than adolescents.

Table 2: Mean WS distance values for the two age groups, vowels and consonants.

	/p/		/t/	
	/a/	/i/	/a/	/i/
13yo	2.08	2.22	1.86	2.00
5yo	2.77	3.50	3.03	3.37

4. DISCUSSION

The results from this study provide new information on consonant-specific coarticulation in children. We have shown that 5-year-olds can anticipate the tongue position of the upcoming vowel not only at mid-/p/ (cf. [31]), but also, to some extent, at mid-/t/ (cf. [23, 16, 29]). The finding that the two age groups demonstrated a similar amount of vowel-related coarticulation for /p/ (see also [31]) does not support either of the two opposing theoretical views on developmental changes in amount of coarticulation, “segmental” ([10]) or “syllabic” ([14]). Some support for the former view comes from the finding that at mid-/t/, 5-year-old children showed certain developmental immaturities in adapting the tongue shape to the following vowel, discussed below.

The alveolar stop has more complex production requirements than the bilabial, as the tip/blade-to-palate closure needs to coincide in time with adjustments to the rest of the tongue in anticipation of the following vowel (see [18]). Our explanation of the fact that the 5-year-olds did better on the bunching extent measure than on the bunching location measure relates to the different aspects of controlling tongue movements captured by these two measures. Bunching extent is mostly modified by dorsum raising (see [26], where /k/ had the largest DEI value of several consonants), while bunching location seems to be modified by more complex adjustments, which can involve coordinated actions of tongue body and tongue tip (cf. [7]). Thus, with LOC_{a-i} reflecting a more complex lingual articulatory trajectory, the lack of tongue adaptation on this measure in younger children would reflect less mature control than in adolescents, relating to the lack of functional differentiation between the front and the back of the tongue (see [4, 5]).

We have shown evidence that the 5-year-old group had more within-speaker token-to-token variability in absolute tongue position than the 13-year-old group. This finding confirms our predictions and agrees with previous studies (e.g., [23, 21, 29]). Interpretation of this result, however, requires some caution, because we cannot rule out the possibility that some of this larger variation in the younger group may have been due to more head movement in relation to the ultrasound transducer in younger speakers, rather than to across-group differences in motor control.

Finally, our results demonstrate that timing is an important factor when studying coarticulation development (cf. [24, 11, 19]). While previous studies have reported adult-like ([16]) or more-than-adult ([15]) coarticulation in children at later time points in alveolar consonants, our findings suggest that at mid-closure for the alveolar stop, 5-year-olds do not adapt the tongue shape to the upcoming vowel as much as adolescents. In future studies, we are planning to combine the methods used in [30] and [27], in order to compare temporal lingual coarticulation across age groups.

5. ACKNOWLEDGEMENTS

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

6. REFERENCES

[1] Articulate Instruments Ltd. 2012. *Articulate Assistant Advanced Ultrasound Module User Guide: Version 2.14*. Edinburgh, UK: Articulate Instruments Ltd.

[2] Baayen, R. H. (2008). *Analyzing Linguistic Data. A Practical Introduction to Statistics Using R*. Cambridge: Cambridge University Press.

[3] Fitch, W. T., Giedd, J. 1999. Morphology and development of the human vocal tract: A study using magnetic resonance imaging. *J. Acoust. Soc. Am.* 106, 1511–1522.

[4] Gibbon, F. E. 1999. Undifferentiated lingual gestures in children with articulation / phonological disorders. *J. Speech Lang. Hear. Res.* 42, 382–397.

[5] 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. *Proc. Meetings on Acoustics* 1, 060003, 1–9.

[6] Goffman, L., Smith, A., Heisler, L., Ho, M. 2008. The breadth of coarticulatory units in children and adults. *J. Speech Lang. Hear. Res.* 51, 1424–1437.

[7] Iskarous, K., Fowler, C. A., Whalen, D. H. 2010. Locus equations are an acoustic expression of articulator synergy. *J. Acoust. Soc. Am.* 128, 2021–2032.

[8] Katz, W. F., Bharadwaj, S. 2001. Coarticulation in fricative-vowel syllables produced by children and adults: A preliminary report. *Clin. Ling. Phon.* 15, 139–143.

[9] Katz, W. F., Kripke, C., Tallal, P. 1991. Anticipatory coarticulation in the speech of adults and young children: Acoustic, perceptual, and video data. *J. Speech Lang. Hear. Res.* 34, 1222–1232.

[10] Kent, R. D. 1983. The segmental organization of speech. In: MacNeilage, P. F. (ed.), *The Production of Speech*. New York: Springer-Verlag, 57–89.

[11] Koenig, L. L., Lucero, J. C., Perlman, E. 2008. Speech production variability in fricatives of children and adults: Results of functional data analysis. *J. Acoust. Soc. Am.* 124, 3158–3170.

[12] Li, F., Edwards, J., Beckman, M. E. 2009. Contrast and covert contrast: the phonetic development of voiceless sibilant fricatives in English and Japanese toddlers. *J. Phon.* 37, 111–124.

[13] Munson, B. 2004. Variability in /s/ production in children and adults: Evidence from dynamic measures of spectral mean. *J. Speech Lang. Hear. Res*. 47, 58–69.

[14] 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. *J. Speech Lang. Hear. Res.* 32, 120–132.

[15] Nittrouer, S., Studdert-Kennedy, M., Neely, S. T. 1996. How children learn to organize their speech gestures: Further evidence from fricative-vowel syllables. *J. Speech Hear. Res.* 39, 379–389.

[16] Noiray, A., Ménard, L., Iskarous, K. 2013. The development of motor synergies in children: Ultrasound and acoustic measurements. *J. Acoust. Soc. Am*. 133, 444–452.

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

[18] Recasens, D., Pallarès, M. D., Fontdevila, J. 1997. A model of lingual coarticulation based on articulatory constraints. *J. Acoust. Soc. Am*. 102, 544–561.

[19] Romeo, R., Hazan, V., Pettinato, M. 2013. Developmental and gender-related trends of intra-talker variability in consonant production. *J. Acoust. Soc. Am.* 134, 3781–3792.

[20] Sereno, J.A., Lieberman, P. (1987). Developmental aspects of lingual coarticulation. *J. Phon.* 15, 247–257.

[21] Smith, A., Zelaznik, H. N. 2004. Development of functional synergies for speech motor coordination in childhood and adolescence. *Developmental Psychobiology* 45, 22–33.

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

[23] Sussman, H. M., Hoemeke, K. A., McCaffrey, H. A. 1992. Locus equations as an index of coarticulation for place of articulation distinctions in children. *J. Speech Lang. Hear. Res.* 35, 769–781.

[24] Walsh, B., Smith, A. 2002. Articulatory movements in adolescents: Evidence for protracted development of speech motor control processes. *J. Speech Lang. Hear. Res.* 45, 1119–1133.

[25] Zharkova, N. 2013. Using ultrasound to quantify tongue shape and movement characteristics. *The Cleft Palate-Craniofacial Journal* 50, 76–81.

[26] Zharkova, N. 2013. A normative-speaker validation study of two indices developed to quantify tongue dorsum activity from midsagittal tongue shapes. *Clin. Ling. Phon.* 27, 484–496.

[27] Zharkova, N., Gibbon, F. E. & Hardcastle, W. J. (2015). Quantifying lingual coarticulation using ultrasound imaging data collected with and without head stabilisation. *Clin. Ling. Phon.* 29, 249–265.

[28] Zharkova, N., Hewlett, N. 2009. Measuring lingual coarticulation from midsagittal tongue contours: description and example calculations using English /t/ and /ɑ/. *J. Phon* 37, 248–256.

[29] Zharkova, N., Hewlett, N., Hardcastle, W. J. 2012. An ultrasound study of lingual coarticulation in /sV/ syllables produced by adults and typically developing children. *J. Int. Phon. Assoc.* 42, 193–208.

[30] Zharkova, N., Hewlett, N., Hardcastle, W. J., Lickley, R. 2014. Spatial and temporal lingual coarticulation and motor control in preadolescents. *J. Speech Lang. Hear. Res.* 57, 374–388.

[31] Zharkova, N., Lickley, R. J., Hardcastle, W. J. 2014. Development of lingual coarticulation and articulatory constraints between childhood and adolescence: An ultrasound study. *Proc. 10th ISSP* Cologne, 472–475.