The production of articulatorily complex consonants in Estonian by children and adults

Anton Malmi[1,2], *Claire Nance*[2]
[1] *Institute of Estonian and General Linguistics, University of Tartu, Estonia*
[2] *Department of Linguistics, Lancaster University, United Kingdom*
anton.malmi@ut.ee, c.nance@lancaster.ac.uk

Abstract

Our ongoing research programme aims to understand how children acquire the motor control skills needed for speech. In this pilot study, we looked at how Estonian children and adults produce consonants that require complex tongue gestures. Ultrasound and acoustic data were collected from 8 adults and 8 children aged 4-5 and 6-8. The participants produced palatalised and non-palatalised consonants. Analysis of tongue shapes using Number of Inflections and Modified Curvature Index and showed that palatalised productions consistently resulted in more complex and differentiated tongue configurations (more inflections, higher curvature). Children produced a smaller difference in tongue shape between palatalised and non-palatalised consonants, indicating ongoing maturation of articulatory control, vocal tract, and tongue shape.

Introduction

The aim of this paper is to compare the tongue shapes of Estonian adults and children producing consonants with multiple articulations using ultrasound tongue imaging.

Previous research indicates that children take time to acquire ‘lingual differentiation’, i.e. control over different parts of the tongue (Cleland and Scobbie 2021). For example, consonants that require a higher degree of lingual differentiation, e.g. affricates or trills, are acquired at age 5 years or later (McLeod and Crowe 2018). The reasons for this are cognitive-developmental (Vihman and Croft 2007) and also physiological (Kent 1992). Specifically, younger children do not have full control of the superior longitudinal muscle, meaning multiple lingual constrictions are physiologically challenging for children younger than 6 years (Abakarova, Fuchs, and Noiray 2022). At the same time, children’s tongues are proportionally larger than adults leading to a palatal quality in child speech (Vorperian et al. 2005; Vihman and Vihman 2011).

While previous child articulatory and acoustic research has considered the development of motor control and timing *between* segments (e.g. (Abakarova, Fuchs, and Noiray 2022; Howson and Redford 2021), here we investigate how children acquire tongue shape *within* segments via a case study of phonemic secondary palatalisation in Estonian. In doing so, we aim to build an understanding of lingual differentiation as a window into the acquisition of speech motor control in complex consonants.

In Estonian, the coronal consonants /l, n, s, t/ can undergo secondary palatalisation in non-initial position. In palatalised consonants, the primary place of articulation on the alveolar ridge is accompanied by a secondary place of articulation at the hard palate (Bateman 2011; Malmi 2022). With secondary palatalisation, the body of the tongue is fronted and raised, and the tongue is constricted at two points.

Such consonants require the near-simultaneous coordination of multiple parts of the tongue and are presumably harder to acquire in adult-like fashion for children than for adults, according to typologies of consonant acquisition, development of the superior longitudinal muscle, and timing patterns (McLeod and Crowe 2018; Abakarova, Fuchs, and Noiray 2022).

The aim of this article is also to test out the recording procedure and analysis methods before expanding our research to fully investigate the acquisition of multiple lingual gestures and their coordination. The research questions are linguistic and development as follows:

1. Do palatalised and non-palatalised consonants differ in the extent of lingual differentiation? We hypothesise that palatalised consonant productions have a higher degree of lingual differentiation compared to non-palatalised consonants.

2. Do adults and children differ in the extent of lingual differentiation in consonant production? We hypothesise that children exhibit a lesser degree of lingual differentiation than adults due to the development of the superior longitudinal muscle. Lingual differentiation is expected to increase with age.

Material and methods

The experiment was approved by the Research Ethics Committee of the University of Tartu, Estonia, approval number 397/T-7.

The acoustic and ultrasound data for the experiment were collected in Estonia from 16 speakers. Four aged 3–4 years, four aged 6–8, and eight first-language Estonian-speaking adults aged 32–46. The recordings took place in participants' homes, with low background noise and no distractors.

The data were recorded, and prompts were presented to the participants using Articulate Assistant Advanced (AAA) (Wrench 2024). Midsagittal ultrasound recordings were made with a Telemed MicrUs scanner, a 20mm convex probe (3 MHz, 90 Hz frame rate), and a lightweight headset (Spreafico, Pucher, and Matosova 2018) which held the probe in place under the chin. Acoustic recordings were made with a DPA 4080 Cardioid Lavalier Microphone.

The participants were asked to produce words that they heard via an audio prompt and saw orthographically on the laptop screen with an accompanying picture. Test words consisted of eight minimal pairs for the palatalisation contrast (16 words total) and a further eight words with a palatalised consonant where no corresponding non-palatalised minimal pair exists. The 24 words were repeated five times, where possible. The experiment lasted around 20 minutes. In this paper, we consider the minimal pairs only (see Table 1).

Table 1. Word list used in the experiment.

Word	IPA	English
kas	[kɑsː]	*int.pron*
kass	[kɑsʲː]	*cat*
kus	[kusː]	*inter.pron*
kuss	[kusʲː]	*sush*
pats	[pɑtːs]	*slap*
pats	[pɑtʲːs]	*braid*
nutt	[nutː]	*cry*
nutt	[nutʲː]	*smarts*
kann	[kɑnː]	*jug*
kann	[kɑnʲː]	*toy*
punu	[punu]	*to braid*
punu	[punʲu]	*tummy*
tall	[tɑlː]	*lamb*
tall	[tɑlʲː]	*barn*
sulg	[sulːk]	*bracket*
sulg	[sulʲːk]	*feather*

The acoustic data were force-aligned (Alumäe, Tilk, and Asadullah 2018) and manually checked in Praat (Boersma and Weenink 2025). Tongue splines were fitted with DeepLabCut (DLC) (Mathis et al. 2018) in AAA. Tongue coordinates were rotated to each speaker's occlusal plane using a bite plate recording (Scobbie et al. 2011) made at the start of each speaker's data collection.

Our analysis of tongue shape includes two measures which aim to capture the extent of lingual differentiation. Firstly: Number of Inflection Points (NINFL), which shows the number of times a tongue shape changes from convex to concave (Preston et al. 2019). A greater NINFL should indicate greater lingual differentiation. Secondly: Maximum Curvature Index (MCI) (Dawson, Tiede, and Whalen 2016). MCI measures the extent of tongue curvature relative to the arc length, indicating how curled up or stretched out the tongue is. A higher value indicates a more curled-up tongue, while a lower value indicates a more stretched-out tongue shape.

NINFL was calculated in AAA and exported with tongue spline coordinates. Following Dokovova et al. (2023), we added +1 to the NINFL values and filtered out the values above 6. MCI was calculated using a Python script from Dawson, Tiede, and Whalen (2016).

In this paper, we analysed NINFL and MCI at a static timepoint on the Vowel-Consonant boundary. This is because previous work on Estonian palatalisation has shown that this is where the effect of palatalisation on the quality of the consonant is the strongest (Malmi 2022). After manual data checking and filtering, we were left with 1222 tokens.

For the data analysis, we used RStudio (RStudio Team 2020). For the NINFL data using the *ordinal* package (Christensen 2023), and for the MCI data, we fitted a linear mixed effects model using the *lme4* package (Bates et al. 2015). The ordinal and linear models had NINFL and MCI as the dependent variables, respectively. Both models contained participant age in three groups (adult, 6-8 years, 4-5 years), consonant (/l/, /t/, /n/, /s/), preceding vowel (/ɑ/, /u/) and palatalisation ("yes" vs. "no") and all the possible two-way interactions between them as independent variables. The test word and speaker were included as random intercepts. Significance testing was carried out by first assessing the output of the full model. Non-significant interactions were removed step-wise until an optimal model was achieved. The code and data are available: https://osf.io/ydt6g/.

Results

Number of Inflections

The NINFL data are plotted in Figure 1, and the results of the statistical modelling are shown in Table 2.

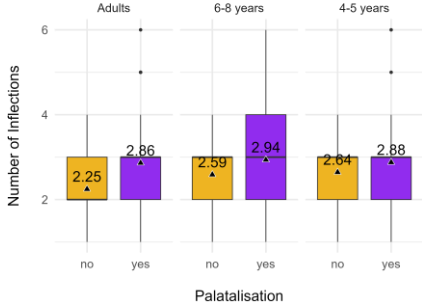

Figure 1. The NINFL values according to palatalisation by different age groups. The black triangles denote the mean values.

NINFL values indicated that palatalised productions showed a higher number of inflections than non-palatalised ones. Adult speakers produced non-palatalised tokens with lower NINFL values than children. But this result needs to be interpreted via the significant interaction between age group and palatalisation (see Table 2).

Table 2. Summarised output of the ordinal linear mixed model of NINFL beta (*β*), standard error (std.err) and *p* values.

Fixed effects	*β*	Std.err	*p*
6-8 years	0.86	0.37	0.012
4-5 years	0.95	0.34	0.014
Palatalisation	1.62	0.27	<0.001
/n/	0.63	0.22	0.004
Preceding /u/	1.13	0.16	<0.001
6-8 years: palatalisation	-0.68	0.28	0.014
4-5 years: palatalisation	-0.86	0.29	0.003
Palatalisation: /s/	0.88	0.32	0.006
Palatalisation: /u/	-1.53	0.23	<0.001

This indicates that the children produced a smaller distinction between minimal pairs than the adults. Indeed, the medians for palatalised and non-palatalised consonants are the same for the child groups (Figure 1). When the preceding vowel was /u/, NINFL values were lower. If the consonant was /s/, NINFL values were higher in palatalised productions, and higher overall in /n/.

Modified Curvature Index

The MCI data are plotted in Figure 3, and the results of the statistical modelling are shown in Table 3.

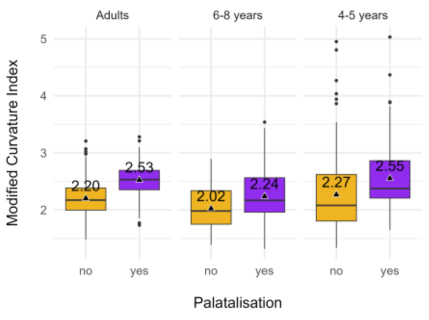

Figure 2. The MCI values of non-palatalised (no - yellow) and palatalised (yes - purple) tokens by different age groups. The black triangles denote the mean values.

MCI values were higher for palatalised productions. There were no overall significant differences between the groups. However, the interaction between palatalisation and the 6–8-year-old group was significant, suggesting that older children showed lower MCI values in palatalised productions compared to adults (see Table 3).

Table 3. Summarised output of the linear mixed model of MCI beta (*β*), standard error (std.err) and *p* values.

Fixed effects	*β*	Std.err	*p*
Palatalisation	0.38	0.03	<0.001
/s/	0.23	0.06	0.006
6-8 years: palatalisation	-0.10	0.04	0.008
4-5 years: palatalisation	-0.04	0.04	0.30
6-8 years: /s/	-0.17	0.05	0.002
4-5 years: preceding /u/	-0.11	0.04	0.003
Palatalisation: /t/	-0.13	0.05	0.004

MCI values were significantly higher when the consonant was /s/, but in the interaction between 6–8-year-old group and /s/, the MCI values were lower compared to the adult productions of /l/. The vowel that preceded the consonant had an effect on MCI values in the interaction with 4–5-year-old group, where the values were lower. MCI values were also lower in the interaction between palatalisation and /t/.

Discussion

To study the lingual differentiation of non-palatalised and palatalised consonants by children and adults, we looked at the Modified Curvature Index (MCI) and the Number of Inflections (NINFL).

Firstly, we wanted to study whether palatalised and non-palatalised consonants differ in the extent of lingual differentiation. We confirmed the hypothesis that palatalised consonant productions have a higher degree of lingual differentiation compared to non-palatalised consonants. Both the NINFL and MCI values were higher, indicating that the tongue was more curled and had more inflections with palatalisation. This is mostly in line with the findings by Nance and Kirkham (2024), who compared velarised and palatalised consonants and found a more complex tongue shape for onset palatalised consonants.

The preceding vowel also affected the degree of lingual differentiation. When the palatalised consonant was preceded by a high back unrounded vowel /u/ vs low back unrounded vowel /ɑ/, the number of inflections were smaller. This is due to coarticulation: when the preceding vowel is low, there is a higher magnitude of tongue displacement (Malmi 2022). With higher vowels, the tongue has a smaller degree of freedom as it is already closer to the palate.

Secondly, we wanted to compare the production of palatalisation in three different age groups to see whether adults and children differ in the extent of lingual differentiation in consonant production. Few previous studies have directly compared adults and children using these metrics (Kabakoff et al. 2023; Nance and Kirkham 2024). Our developmental findings interact with the palatalisation contrast: adult and child NINFL and MCI results are similar for palatalised consonants, but children produce less of a distinction between palatalised and non-palatalised consonants than adults. Their tongue shapes for non-palatalised consonants are more palatal (greater NINFL, higher MCI) than adult non-palatalised consonants. This finding supports previous observations that child speech has a palatal quality (Vihman and Vihman 2011) but does not indicate predicted age-related differences in lingual differentiation for the articulatorily complex palatalised consonants.

Our data suggest that the inflections and curvature needed to produce palatalisation are acquired by age 4 years, but producing the tongue shapes needed for phonemic contrast takes longer, into later childhood and support the findings of a previous investigation of laterals where the allophonic distinction for secondary articulations was acquired up to age 7;11 (Lin and Demuth 2015).

Given the small-scale nature of our pilot study, more data will be needed before we can confirm these findings. Our current working hypothesis is that the palatalisation contrast is easy to perceive for even young children, but this will be fully tested in future data collection. We will also continue to assess the utility of NINFL and MCI for investigating lingual differentiation (Palo and Lulich 2024).

Conclusions

The results of our pilot study show that palatalisation leads to greater lingual differentiation across measures, with developmental differences. Specifically, children are less able to produce differences in lingual differentiation needed for phonemically contrastive palatalisation. These patterns suggest ongoing maturation of speech-motor control during early childhood for phonemic contrasts involving secondary articulations such as palatalisation.

In our future work, we will expand to a larger sample size, including children aged 3–11 years. We will also consider the temporal nature of secondary articulation contrasts and how this develops across childhood.

Acknowledgements

The authors would like to thank the participants of the study for their trust and cooperation. This research has been funded by the Estonian Research Council (PUTJD1246).

References

Abakarova, Dzhuma, Susanne Fuchs, and Aude Noiray. 2022. "Developmental Changes in Coarticulation Degree Relate to Differences in Articulatory Patterns: An Empirically Grounded Modeling Approach." Journal of Speech, Language, and Hearing Research 65(9):3276–99.

Alumäe, Tanel, Ottokar Tilk, and Asadullah. 2018. "Advanced Rich Transcription System for Estonian Speech." Pp. 1–8 in *Human Language Technologies - the Baltic Perspective : Proceedings of the Eighth International Conference*, edited by K. Muischnek and K. Müürisep. Amsterdam: IOS Press.

Bateman, Nicoleta. 2011. "On the Typology of Palatalization." *Language and Linguistics Compass* 5(8):588–602.

Bates, Douglas, Martin Mächler, Ben Bolker, and Stewen Walker. 2015. "Fitting Linear Mixed-Effects Models Using Lme4." *Journal of Statistical Software* 67(1):1–48.

Boersma, Paul, and David Weenink. 2025. "Praat: Doing Phonetics by Computer."

Christensen, R. H. B. 2023. "Ordinal—Regression Models for Ordinal Data."

Cleland, Joanne, and James M. Scobbie. 2021. "The Dorsal Differentiation of Velar From Alveolar Stops in Typically Developing Children and Children With Persistent Velar Fronting." *Journal of Speech, Language, and Hearing Research* 64(6S):2347–62.

Dawson, Katherine M., Mark K. Tiede, and D. H. Whalen. 2016. "Methods for Quantifying Tongue Shape and Complexity Using Ultrasound Imaging." *Clinical Linguistics & Phonetics* 30(3–5):328–44.

Dokovova, Marie, Ellie Sugden, Gemma Cartney, Sonja Schaeffler, and Joanne Cleland. 2023. "Tongue Shape Complexity in Children With and Without Speech Sound Disorders." *Journal of Speech, Language, and Hearing Research* 66(7):2161–83.

Howson, Phil J., and Melissa A. Redford. 2021. "The Acquisition of Articulatory Timing for Liquids: Evidence from Child and Adult Speech." *Journal of Speech, Language, and Hearing Research* 64(3):734–53.

Kabakoff, Heather, Sam Pearl Beames, Mark Tiede, D. H. Whalen, Jonathan L. Preston, and Tara McAllister. 2023. "Comparing Metrics for Quantification of Children's Tongue Shape Complexity Using Ultrasound Imaging." *Clinical Linguistics & Phonetics* 37(2):169–95.

Kent, Ray D. 1992. "The Biology of Phonological Development." Pp. 65–90 in *Phonological Development: Models, Research, Implications*. Baltimore: York Press.

Lin, Susan, and Katherine Demuth. 2015. "Children's Acquisition of English Onset and Coda /l/: Articulatory Evidence." *Journal of Speech, Language, and Hearing Research* 58(1):13–27.

Malmi, Anton. 2022. "The Production of Estonian Palatalization by Estonian and Russian Speakers." University of Tartu.

Mathis, Alexander, Pranav Mamidanna, Kevin M. Cury, Taiga Abe, Venkatesh N. Murthy, Mackenzie Weygandt Mathis, and Matthias Bethge. 2018. "DeepLabCut: Markerless Pose Estimation of User-Defined Body Parts with Deep Learning." *Nature Neuroscience* 21(9):1281–89.

McLeod, Sharynne, and Kathryn Crowe. 2018. "Children's Consonant Acquisition in 27 Languages: A Cross-Linguistic Review." *American Journal of Speech-Language Pathology* 27(4):1546–71.

Nance, Claire, and Sam Kirkham. 2024. "Acquiring Tongue Shape Complexity in Consonants with Multiple Articulations." in *Proceedings of the 5th International Symposium on Applied Phonetics*. Tartu, Estonia.

Palo, Pertti, and Steven M. Lulich. 2024. "Toward Assessing Stability and Sensitivity of Tongue Contour Metrics." in *Proceedings of Ultrafest 2024*. Aizu-Wakamatsu, Japan.

Preston, Jonathan L., Patricia McCabe, Mark Tiede, and Douglas H. Whalen. 2019. "Tongue Shapes for Rhotics in School-Age Children with and without Residual Speech Errors." *Clinical Linguistics & Phonetics* 33(4):334–48.

RStudio Team. 2020. *RStudio: Integrated Development Environment for R*. Boston, MA: RStudio, PBC.

Scobbie, James M., Eleanor Lawson, Steve Cowen, Joanne Cleland, and Alan A. Wrench. 2011. "A Common Co-Ordinate System for Mid-Sagittal Articulatory Measurement." QMU CASL Working Papers WP-20.

Spreafico, Lorenzo, Michael Pucher, and Anna Matosova. 2018. "UltraFit: A Speaker-Friendly Headset for Ultrasound Recordings in Speech Science." in *Proceedings of Interspeech*.

Vihman, Marilyn, and William Croft. 2007. "Phonological Development: Toward a 'Radical' Templatic Phonology." *Linguistics* 45(4):683–725.

Vihman, Marilyn, and Virve-Anneli Vihman. 2011. "From First Words to Segments: A Case Study in Phonological Development." in *Experience, Variation, and Generalization: Learning a first language, Trends in Language Acquisition Research*, edited by E. V. Clark and I. Arnon. Amsterdam: John Benjamins Publishing Company.

Vorperian, Houri K., Ray D. Kent, Mary J. Lindstrom, Cliff M. Kalina, Lindell R. Gentry, and Brian S. Yandell. 2005. "Development of Vocal Tract Length during Early Childhood: A Magnetic Resonance Imaging Study." *The Journal of the Acoustical Society of America* 117(1):338–50.

Wrench, Alan. 2024. "Articulate Assistant Advanced (version 221.5.4)."