

Child acquisition of complex consonants: The role of oral motor control in developing acoustic targets^{a)}

Anton Malmi^{1,2,b)}  and Claire Nance^{2,3} 

¹*Institute of Estonian and General Linguistics, University of Tartu, Tartu, Estonia*

²*Linguistics and English Language, Lancaster University, Lancaster, United Kingdom*

³*Plural Research Centre, Tampere University, Tampere, Finland*

ABSTRACT:

Our study examines the role of oral motor control in the acquisition of acoustic targets for lingually complex consonants in Estonian by speakers from four age groups. We focus on children's production of secondarily palatalised and non-palatalised consonants where anatomical development and acquisition of motor control skills make differing predictions about the timing and nature of speech production. Our aim is to disentangle the contribution of, on the one hand, speech motor control development and, on the other hand, anatomical growth and articulator change in proportion. We examined oral motor control using a single syllable alternating motion rate diadochokinetic (DDK) task and studied the difference between the second formant (F2) and first formant (F1) using functional principal component analysis. Our results show that palatalised and non-palatalised productions differ significantly across all age groups, indicating children can produce palatalisation contrasts from age 3 years, but consistency in production increases with age. Results also show that children and adults with greater oral motor control skills are better able to produce the palatalisation contrast. These results suggest that anatomical proportions pre-dispose children toward early acquisition of the acoustic properties of palatalisation. However, adult-like production of the phonological contrast also depends on the acquisition of oral motor control skills.

© 2026 Author(s). All article content, except where otherwise noted, is licensed under a Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>). <https://doi.org/10.1121/10.0046652>

(Received 23 February 2026; revised 23 August 2026; accepted 31 August 2026; published online 1 October 2026)

[Editor: Jalal Al-Tamimi]

Pages: 2859–2870

I. INTRODUCTION

In this paper, we examine the role of oral motor control in the acquisition of acoustic targets for lingually complex consonants. Child speech acquisition involves learning the cognitive categories for language but also the physical ability to produce them (Kent, 1992). During this process, a child must learn the skills necessary for precise movements of the speech articulators and adapt these movements responsively as their anatomy develops and changes in proportion. We focus on child production of consonants /l, t, n, s/, which can also be phonemically palatalised /l^j, t^j, n^j, s^j/ in Estonian. Research on anatomical development and acquisition of speech motor control skills make differing predictions about the timing and nature of speech production. Large cross-linguistic comparisons indicate that consonants with multiple lingual articulations are typically acquired later on in the developmental trajectory (McLeod and Crowe, 2018). This is hypothesised to be because of the development of speech motor control skills needed for precise tongue placement. However, in secondary palatalisation, a child's anatomical proportions might predispose them to produce palatal-sounding speech (Ciusa *et al.*, 2007; Vihman and Vihman, 2011).

Our aim is to disentangle the contribution of, on the one hand, motor control development and, on the other hand, anatomical growth and articulator change in proportion. In doing so, we test long-standing hypotheses about the order of acquisition for complex consonants in child speech. Additionally, we diversify child speech studies in two respects. First, we provide a detailed account of the acoustic acquisition of secondary palatalisation. Second, we respond to calls in the literature for more child language data from non-Indo-European languages (Kidd and Garcia, 2022) by analysing data from Estonian, a Finno-Ugric language, and providing normative data on oral motor control development.

In the remainder of this section, we discuss previous phonetic research which has considered the trajectory of consonant acquisition and the role of motor control skills (Sec. IA). Second, we discuss anatomical development and how this might predispose children to palatalise (Sec. IB). Third, we discuss secondary palatalisation in Estonian (Sec. IC), before outlining our research questions and hypotheses (Sec. ID).

A. Acquisition of oral motor control and complex consonants

It is generally observed that children acquire consonants in a hierarchical fashion of increasing difficulty and complexity (Shriberg, 1993). In addition to increasing

^{a)}This paper is part of a special issue on Data and Models in Speech Production: A Situated Perspective.

^{b)}Email: anton.malmi@ut.ee

complexity, the individual child's abilities and the environment in which a child grows up also play a role (Davis and Bedore, 2013). Our analysis focusses on the articulatory and physical factors in child speech development. We assess the role of individual differences in the acquisition of speech motor control and how this relates to consonant production at different ages. Cross-linguistic surveys have been instrumental in recognising consonant acquisition trends across the world and understanding the role of articulatory factors in general patterns of child acquisition. McLeod and Crowe (2018) assess normative data from 27 languages to identify that consonants with one primary articulation tend to be acquired earlier. For example /p, m, t/ are acquired by 2;11 with up to 100% accuracy. Consonants with multiple articulations or those requiring precise tongue placement and timing are usually acquired later, for example, /k^{wh}, t^ɕ/ are typically acquired by 4;11 (Table I in McLeod and Crowe, 2018). Previous detailed acoustic and articulatory work also indicates that segments requiring multiple lingual gestures are among the last sounds to be acquired, for example, coda lateral /l/ (Lin and Demuth, 2015) and rhotic /r/ (Klein *et al.*, 2013).

When children learn to produce speech sounds, they have to rely on somatosensory and auditory feedback to map and correlate articulatory movements to specific speech sounds considering the current context of the system (Guenther, 2006). As children's anatomy is constantly changing, they must adjust, refine and correct their sensory mappings based on the auditory (Ayala *et al.*, 2023) and articulatory feedback (Tremblay *et al.*, 2003). This feedback loop is important for the speech to be fluent (Guenther, 2006). With age, auditory feedback becomes less important, and sensorimotor representations reliably predict the motor commands needed to produce speech sounds (Bishai *et al.*, 2026). Our focus is on consonants requiring multiple lingual articulations. To produce such sounds, children need to gain control in particular of the superior longitudinal muscle through auditory and somatosensory feedback for intelligible production (Abakarova *et al.*, 2022).

To gain a window into the development of speech motor control and how this leads eventually to adult-like production, researchers have pursued two main strategies: first, investigating the overlap of segments through coarticulation, and second, by investigating tongue flexibility and timing. Coarticulatory research has revealed differences between adults and children, and suggests that child coarticulation patterns do not become adult-like until adolescence (Zharkova, 2018). This supports earlier work using other methods such as electromagnetic articulography (EMA), which indicates that motor control over the speech articulators does not stabilise until late adolescence (Cheng *et al.*, 2007b). In general, the research programme in child coarticulation finds that children coarticulate more than adults (Abakarova *et al.*, 2022; Noiray *et al.*, 2018; Rubertus and Noiray, 2024, 2018), although others find that children coarticulate less than adults (Barbier *et al.*, 2020; Zharkova, 2018). Child articulation is also more variable than adult articulation (Barbier *et al.*, 2020).

This body of research in child speech articulation shows clear differences in adult and child timing and consistency of gestures. One of the reasons cited for these differences is maturation of the fine motor control needed to differentiate and coordinate segments in an adult-like fashion (Abakarova *et al.*, 2022; Barbier *et al.*, 2020). Specifically, acquisition of the ability to manipulate the different parts of the tongue is referred to as "lingual differentiation" (Gibbon, 1999). Lingual differentiation is essential for the production of consonants with multiple lingual articulations. Articulatory work on lingual differentiation indicates that differentiation increases with age and that children with speech sound disorder might have difficulties in acquiring differentiation (Cleland and Scobbie, 2021; Dokovova *et al.*, 2023; Kabakoff *et al.*, 2021). Crucially, acquisition of lingual differentiation must be accompanied by acquisition of adult-like timing patterns for differentiated gestures to be perceived as appropriately coarticulated adult-like speech (Cychosz, 2021; Howson and Redford, 2021; Kpogo and Chang, 2025).

To measure the acquisition of oral motor control, we used a diadochokinetic (DDK) speech task, a tool commonly used by speech and language therapists and researchers to assess the extent of articulator agility, mobility, and oral motor coordination (Duffy, 2013; Kent *et al.*, 2022). In those DDK tasks participants are asked to clearly and rapidly repeat a variety of syllables, syllable combinations or words in a given amount of time. Although researchers have questioned the usefulness of studying non-speech oral movements as it uses a different mechanism from meaningful speech (Ziegler *et al.*, 2019), the method is still widely used in clinical assessment, multilingualism, and child research (e.g., Chu *et al.*, 2021a; Ha, 2023; Ong *et al.*, 2024). In child research concentrating on typically developing children, the findings indicate that overall the repetition rate is increasing with age (Ha, 2023; Murali *et al.*, 2024).

Specifically in the production of monosyllables, children are slower than adults and three-year-olds slower than older children (Ha, 2023). Although the DDK rate increases with age, the relationship is not linear. Children from three to five have been shown to have similar rates but increase in accuracy and consistency during that time (Stackhouse, 2000). The DDK rate in adults has been shown to slow down with age, and older people are slower than middle-aged with no apparent gender differences in that trend (Chu *et al.*, 2021b). Interestingly, because of phonetic and phonological differences between languages, DDK rates also vary cross-linguistically, and the rates are not always directly comparable between languages (Chu *et al.*, 2021a). Estonian speech and language therapists have reported that they regularly use DDK tasks in their work to help to diagnose issues with oro-motor control (Lahtein-Kürsa *et al.*, 2025). However, despite the widespread use of this tool, there is currently only one master's thesis that presents normative data on DDK scores at different ages in Estonian (Kanarbik, 2026). An additional contribution of our study is to provide supplementary data to this study.

We assess the role of oral motor control in the acquisition of acoustic targets for consonants requiring multiple lingual articulations. We assume that lingual differentiation and timing will be essential in this process, and we expect to observe developmental differences. Our case study involves the context of secondary palatalisation, which challenges some of the previously held assumptions about motor control and speech acquisition, as we discuss in Sec. IB.

B. Acquiring secondary palatalisation

We define “complex consonants” as those requiring multiple independently controlled articulatory constrictions (Shaw *et al.*, 2021). Shaw *et al.*’s (2021) work indicates that complex segments, e.g., /p^j/, differ from segment sequences, e.g., /pj/, in their coordination patterns and gestural landmarks (Gafos, 2002). Within the grouping of complex consonants, we focus on phonemic secondary palatalisation. For these consonants in coda position (as in Estonian, Sec. IC), it is argued that phonemic palatalisation is timed in-phase with consonantal gestures, although timing varies with individual and contextual perceptual maximisation (Bennett *et al.*, 2025).

In secondary palatalisation, a consonant has a primary place of articulation (PoA), e.g., the alveolar ridge, and also a secondary constriction at the hard palate (Bateman, 2011). This additional palatalisation gesture can form the basis of distinguishing minimal pairs, e.g., *kann* /kan:/ “jug” and *kann* /kan^j:/ “toy” in Estonian. The literature discussed above would predict that secondary palatalisation should be acquired later on in child development because of the multiple lingual constrictions required. However, there are indications that secondarily palatalised consonants might be acquired earlier than other complex segments. For example, in Russian, a language with an extensive system of secondary palatalisation contrasts, children sometimes acquire palatalised /t^j/ before /t/ (Zharkova, 2004). Similarly, Timm (1977)’s diary study of Russian acquisition by one child notes that /g^j/ and /k^j/ are produced with fewer errors than /g/ and /k/. Russian child speech has long been noted as characterised by an “overall softness” (palatalised quality), which is attributed to the evolution of the vocal tract morphology (Zharkova, 2004).

Zharkova provides an observational and acoustic analysis of the acquisition of the palatalisation contrast in seven Russian first language (L1) children aged 1;3 to 3;2. She notes that production is quite variable, and children appear to be testing different strategies until the age of three years. The oldest child in the sample produces speech that perceptually approximates adult targets. In a diary study of acquisition of Estonian and American English bilingual child from 12–21 months, researchers found that she frequently produced forms with a “palatal template,” where her first words were produced with high front vowels, or front rising diphthongs (Vihman and Vihman, 2011). This showed that the child used palatal forms at a higher rate than the input received from her parents (Vihman and Vihman, 2011).

Existing studies of the acquisition of Russian and Estonian suggest anatomical reasons behind the early acquisition of secondary palatalisation and the overall palatal quality of child speech (Vihman and Vihman, 2011; Zharkova, 2004). When compared to three-year-old children, six-year-olds have greater palatal height, greater palatal slope, and an overall more domed palate shape (Ciusa *et al.*, 2007). Similarly, electropalatography (EPG) data indicate greater palatal contact in children aged six years compared to adolescents and adults (Cheng *et al.*, 2007a), and also more fronted tongue articulations in six-year-olds compared to adolescents (Fletcher, 1989). In the development of the vocal tract, many of the speech articulators undergo a period of very rapid evolution around 13–14 months and begin to take on adult-like proportions, for example, vocal tract length, tongue length, naso-oro-pharyngeal length, and laryngeal height. The hard palate, however, continues to evolve very rapidly until approximately 25 months (Vorperian *et al.*, 2005). This means that although the child is producing their first words and longer utterances, the ratio of tongue size to hard palate length is not in adult-like proportion. As such, they either need to make articulatory adaptations to achieve adult-like acoustic targets, or they will produce speech with a palatal quality.

To summarise, two contrasting factors might contribute to the acquisition of palatalisation in child speech: (1) the process of acquiring tongue motor control and timing patterns (predicting difficulty in acquiring lingually complex consonants) and (2) anatomical factors such as palate shape and the relative length of the hard palate in younger children (predicting earlier and increased palatalisation). Our study investigates the relationship between the acquisition of motor control and acoustic targets for phonemic secondary palatalisation. If children who have greater motor control skills have more adult-like production of the palatalisation contrast, then production of these complex consonants is related to the development of speech motor control. If there is no or a weaker relationship between tongue control and palatalisation, then production of the palatalisation contrast is more closely related to anatomical development than acquisition of tongue motor control.

C. Secondary palatalisation in Estonian

Estonian coronal consonants /l, t, n, s/ can be non-palatalised or palatalised /l^j, t^j, n^j, s^j/ (Asu and Teras, 2009). Palatalised consonants can only occur in syllable codas, either at the end of monosyllabic words or after the first stressed syllable in disyllabic words (Teras and Pajusalu, 2014). Typologically, Estonian is unusual in only having a palatalisation contrast in coda position. Typically, languages that have phonemic secondary palatalisation have it in onsets and then variably in codas. For example, contrastive palatalisation in codas has been neutralised in Bulgarian (Iskarous and Kavitskaya, 2018) and in Nenets (Kochetov, 2002).

Estonian palatalisation results in minimal pairs, e.g., *kott* /kot:/ “worn out shoe” and *kott* /kotʲ:/ “bag” and is not shown in contemporary orthography. For Estonian palatalised consonants, the primary PoA on the alveolar ridge is accompanied by a near-simultaneous constriction with the tongue body at the hard palate (Malmi and Lippus, 2019). Estonian palatalised consonants are prepalatalised, meaning that most of the perceptual and acoustic quality of palatalisation is carried by the preceding vowel. This leads to a front-raising diphthongal quality to vowels preceding palatalised consonants and increased duration (Malmi *et al.*, 2023; Teras and Pajusalu, 2014). The importance of the preceding vowel was highlighted in a perception experiment: Estonian listeners were not able to accurately identify palatalised consonants when the transition preceding a palatalised consonant was removed (Liiv, 1965).

Dynamic acoustic analysis of palatalised and non-palatalised vowels and consonants in adult speech is presented in Malmi *et al.* (2023). This study indicates that in adult speech, second formant (F2) values during the vowel preceding palatalisation reliably distinguish palatalised and non-palatalised consonants at all places of articulation and in all vowel contexts. In all cases, F2 is higher in palatalisation. The first formant (F1) tends to be lower in palatalisation, and third formant (F3) does not systematically vary according to palatalisation. Articulatorily those changes in the formant values reflect the higher and more fronted tongue position with palatalisation (Malmi and Lippus, 2019). These findings confirm other descriptions of secondary palatalisation cross-linguistically, where most often raised F2 and lowered F1 are reported in the vowels surrounding palatalised consonants (Kochetov, 2017; Nance and Kirkham, 2020). In Estonian, the timepoint where formant trajectories in vowels preceding palatalised consonants diverge from those preceding non-palatalised consonants is at vowel onset or in the first 50% of vowel duration for the majority of possible contexts and consonants (Malmi *et al.*, 2023). In contrast, acoustic cues for palatalisation in Estonian consonants themselves are less consistent and typically affect only the first part of the consonant (Malmi *et al.*, 2023). Together, these findings suggest that adult-like palatalisation in Estonian can robustly be identified through analysis of F1 and F2 in the vowels preceding palatalised consonants.

D. Research questions and hypotheses

Our study investigates the acquisition of complex consonants by children and adults focussing on secondary palatalisation. We examine the contrasting hypotheses that palatalisation should be acquired later on because of the slower acquisition of fine motor control skills and manipulation of different tongue muscles and that palatalisation should be acquired early because of anatomical biases in child physiology. We include measures of oral motor control ability to test for individual differences in this respect and

provide a first set that we know of normative DDK data for Estonian. We address the following research questions:

- (1) Are there age-related differences in oral motor control skills in Estonian?
- (2) Are there age-related differences in the acoustic realisation of secondary palatalisation?
- (3) Is there a relationship between the development of motor control skills and the acquisition of secondary palatalisation acoustics?

We hypothesise that oral motor control skills increase with age (Ha, 2023; Murali *et al.*, 2024). This has been demonstrated with DDK data in other languages, but we need to establish this result first as a basis for research questions 2 and 3. As children get older, their ability to coordinate the timing and movement of different articulators improves. We predict there are developmental differences in the acoustics of palatalisation. Specifically, younger children might not differentiate minimal pairs to the extent that adults do (Malmi and Nance, 2025). We predict that anatomical biases might lead to greater palatalisation in younger children (Ciusa *et al.*, 2007; Zharkova, 2004). We hypothesise that children with greater motor control skills will be able to better manipulate the muscles of the tongue (Dokovova *et al.*, 2023; Gibbon, 1999) and thus differentiate minimal pairs for palatalisation to a greater extent. Because of the expected timing differences in adult and child speech (Howson and Redford, 2021; Noiray *et al.*, 2018), our analysis is designed to incorporate dynamic information about formant trajectories across the vowels preceding palatalised consonants.

II. MATERIALS AND METHODS

A. Participants

The data were collected in three Estonian preschools and schools in the autumn of 2025 and early winter 2026. We worked closely with speech language therapists to find and record typically developing monolingual Estonian children from three age groups: 3;2–4;6 [8 female (F), 10 male (M)], 5;6–7;1 (5 F, 13 M), and 10;0–11;0 (12 F, 12 M). Additionally, we recorded a comparison group consisting of 20 adults who are first-language Estonian speakers (18–35 years old, 10 M, 10 F). In total, we analysed data from 80 speakers in this study.

B. Recordings and materials

The children were recorded in a quiet, separate room at their preschools and schools during their school day. Adults were recorded in the phonetics lab at the University of Tartu. With the most cooperative participants, the experiment took around 30 min to complete. Once we had obtained the consent forms from the parents or adults, the participants were asked to complete four different tasks: (1) syllable production (DDK), (2) picture description, (3) perception of palatalisation, and (4) production of individual words. For this study, we focus on tasks 1 and 4. In the first

three tasks only audio were recorded using Gorilla (Anwyl-Irvine *et al.*, 2021). In the fourth task both the ultrasound and audio signals were recorded using ARTICULATE ASSISTANT ADVANCED (AAA) version 221.5.11 (Wrench, 2024). All throughout the experiment, the children and adults wore a DPA 4080 Cardioid Lavalier Microphone (DPA Microphones A/S, Kikkedal, Denmark). The acoustic data recorded in AAA were recorded at a sampling rate of 22,050 Hz (16 bit). To record the ultrasound data, the participants wore a lightweight headset (Spreafico *et al.*, 2018) that securely held the ultrasound probe under the speaker's chin. The current study is part of a larger project that examines the articulatory properties of Estonian children's productions. Here, we are looking only at the acoustic properties of their speech.

In our version of the DDK task, we asked each participant to rapidly and accurately repeat the syllable /pa/ as many times as possible within a 5 s time window. This was then followed with a trial with /ta/, and then /ka/. The time window was kept to 5 s to be comfortably within the maximum expected phonation time for younger children (Mendes Tavares *et al.*, 2012) and greatly exceeded their maximum capacity that they could produce in one breath. For practice, the participants were first asked to produce different syllables slowly to get the acquainted with the task. We then recorded them producing their best-effort attempt. A full DDK protocol can include many syllable combinations as well as non-speech movements such as lip opening and closing (Kent *et al.*, 2022). We reduced the DDK task to be mindful of the time spent with the children and avoid overburdening them.

The production experiment in AAA consisted of 28 words in isolation including eight minimal pairs for the palatalisation contrast (16 words total) and 4 one-syllable consonant-overlong vowel-consonant (CV::C) containing the vowels /a, e, u, i/ respectively for the purpose of vowel normalisation. We chose common and high frequency words, which were tested in our pilot study (Malmi and Nance, 2025). Also, most of the targets were disyllabic and in nominative case (excluding three) which made them structurally simple. As palatalisation is not represented in contemporary Estonian orthography, and some of the children were not fully literate, we elicited each word by asking participants to listen and repeat an audio recording while looking at a picture representing the target item similar to Lin and Demuth (2015) (see Malmi and Nance, 2026, for our pictures and recordings). It would not be possible to represent all words with pictures alone and without other cues. The audio recordings were made by an L1 Estonian-speaking 37-year old male speaker at the phonetics lab in the University of Tartu using an AKG C414 microphone (AKG Acoustics GmbH, Vienna, Austria) and Presonus StudioLive 16R soundcard (PreSonus Audio Electronics, Inc., Baton Rouge, LA). For adults each test word was repeated five times, for the 10;0–11;0 age group four times, and for the 3;2–4;6 and 5;6–7;1 age groups three times. Here, we analyse the eight minimal pairs for palatalisation,

TABLE I. Word list used in the experiment.

Word	IPA	English	Part of speech	Category
kas	/kas:/	<i>whether</i>	INT.PRON ^a	Minimal pair
kass	/kas ⁱ :/	<i>cat</i>	SG.NOM ^b	Minimal pair
kus	/kus:/	<i>where</i>	INT.PRON	Minimal pair
kuss	/kus ⁱ :/	<i>shush</i>	2SG.IMP ^c	Minimal pair
pats	/pats:/	<i>pat</i>	SG.NOM	Minimal pair
pats	/pat ⁱ :/	<i>braid</i>	SG.NOM	Minimal pair
nutt	/nut:/	<i>cry</i>	SG.NOM	Minimal pair
nutt	/nut ⁱ :/	<i>smarts</i>	SG.NOM	Minimal pair
kann	/kan:/	<i>jug</i>	SG.NOM	Minimal pair
kann	/kan ⁱ :/	<i>toy</i>	SG.NOM	Minimal pair
tunde	/tun:te/	<i>feeling</i>	SG.PART ^d	Minimal pair
tunde	/tun ⁱ :te/	<i>hours</i>	PL.PART ^e	Minimal pair
tall	/tal:/	<i>lamb</i>	SG.NOM	Minimal pair
tall	/tal ⁱ :/	<i>barn</i>	SG.NOM	Minimal pair
mul	/mul:/	<i>me</i>	1SG.GEN ^f	Minimal pair
mull	/mul ⁱ :/	<i>bubble</i>	SG.NOM	Minimal pair
kaal	/ka::l/	<i>weight</i>	SG.NOM	Normalisation
keel	/ke::l/	<i>tongue</i>	SG.NOM	Normalisation
kuul	/ku::l/	<i>moon</i>	SG.ADD ^g	Normalisation
kiil	/ki::l/	<i>dragonfly</i>	SG.NOM	Normalisation

^aInterrogative pronoun.

^bSingular nominative case.

^c2nd person singular imperative case.

^dSingular partitive case.

^ePlural partitive case.

^fFirst person singular genitive case.

^gSingular adessive case.

and we used the four CV::C words for the purposes of acoustic normalisation. This word list is shown in Table I.

The youngest children sometimes produced audible palatalisation in words in which the consonant was not phonologically palatalised. As this could be considered a high-stress situation for them, the experimenter asked the children to repeat the word to ensure children had heard it correctly and knew which target word they were aiming for. In most cases, they did not change their pronunciation, and we accepted their second attempt as their production of that word. As the software used for the recording ultrasound data does not start recording before a beep that occurs approximately a second before pressing record button, the youngest children very regularly imitated the heard stimuli immediately after hearing it, so it was very common to ask them to repeat it again in the same recording.

C. Data processing

We first extracted the syllable repetition recordings from Gorilla and the word list recordings from AAA. The recordings from Gorilla were in WEBM format, which we converted to MP3 using FFmpeg to be readable in PRAAT (Boersma and Weenik, 2025). The word list recordings were force-aligned using the method in Alumäe *et al.* (2018). We then manually checked the force-aligned boundaries in PRAAT. Eleven time-normalised F1 and F2 measurement points of the vowel that preceded each consonant were

extracted using a script that finds a unique formant ceiling for each of the vowels (see Malmi and Nance, 2026, for script). As discussed in Sec. IC, the acoustics of the vowel preceding each consonant is a robust measure of palatalisation in Estonian (Malmi *et al.*, 2023). For adult male speakers, the maximum ceiling was set at 5000 Hz and for female speakers, 5500 Hz. As we did not expect any sex differences for Estonian children under the age of 12 years (Meister and Meister, 2021), following the PRAAT manual, the maximum ceiling for both boys and girls was set at 8000 Hz. Further data analysis was carried out using R and RSTUDIO (r Core Team, 2025). To compare the F values of adults and children of differing ages, the data were normalised using the Lobanov method.

D. Oral motor control analysis

To analyse the DDK data, we manually counted the syllables in PRAAT. We chose this approach over an automated method to provide a qualitative judgment on what counted as a repeated syllable and what did not. We looked for clear breaks in the pitch and intensity traces of the speaker and excluded repetitions in which the separate productions were coarticulated with each other (see “Supplementary figures” in Malmi and Nance, 2026). We then removed the silences from the recording that was strictly limited to 5 s and calculated how many syllables per second were produced by dividing the repetition rates from the time they spent producing syllables (Gadesmann and Miller, 2008).

For the statistical analysis of DDK rates, we used the lme4 package to fit a linear mixed model where the dependent variable was syllables per second. The independent variables were age group, PoA (/pa/, /ta/, /ka/), and the interaction between them. Age group and PoA were sum coded. Speaker was included as a random intercept. Following the scoping review in (Kent *et al.*, 2022), typically sex does not affect DDK rates or are not reported in most of the literature. As such, we have pooled data from male and female speakers in this analysis. In addition, this strategy makes the best use of our statistical power for modelling. For significance testing, we compared the full model with a nested model not containing the predictor of interest *via* likelihood ratio testing in analysis of variance (ANOVA) (Winter, 2020).

E. Acoustics of palatalisation analysis

To capture the time-varying dynamics of the F movements preceding phonological palatalisation, we first calculated the difference between F2 and F1 (Kochetov, 2017; Nance and Kirkham, 2020). Then, using F2–F1, we fitted a functional principal component analysis to those values at the 11 timepoints from the start of the vowel to the start of the consonant. Principal component analysis (PCA) allows dimensionality reduction of complex data into a smaller number of uncorrelated variables (linear functions), which can then be used to generate scores for each token and subject to further analysis (e.g., Jolliffe and Cadima, 2016). Functional principal component analysis (FPCA) allows inclusion of time-series data into a PCA by taking into

account that subsequent timepoints are related to one another (Coretta and Sakr, 2026). As the FPCA output is a list of numbers representing Fs over time, this can be submitted to analysis by linear mixed effects models allowing us to fit a relatively complex structure that would be less feasible and less interpretable than with other methods of functional data analysis (FDA) such as generalised additive mixed models. This method has been most extensively applied to fundamental frequency (f0) contours (Gubian *et al.*, 2015), but we now apply it to the trajectory of F2–F1.

Following Coretta and Sakr (2026), we fitted an FPCA to the distance between F2–F1 trajectories using the fda, funData, multivariate functional principal component analysis (MFPCA), and landmarkregUtils packages (Gubian, 2025; Happ and Greven, 2018; Ramsay, 2003). In our data, functional principal component (FPC)1 captures 74% of the variation in the data, FPC2 captures only 16%, and the subsequent components capture less and less. Because of the very high proportion of variation captured by FPC1 and big drop-off in variation explained between FPC1 and FPC2, we focus on analysis of the scores for FPC1 only (see Turton, 2017, for a similar strategy with “standard” PCA). After that, we looked at the interquartile range and removed 90 outlier FPC1 values that exceeded ± 3 standard deviation (SD) from the mean. We were left with 4460 tokens for the analysis. To check that FPC1 captures the variation in F2–F1 trajectory as intended, Fig. 1 shows the mean of the measured values and also the mean measured values plus or minus fractions of the SD of FPC1. The figure indicates a rising F trajectory, which is expected for vowels preceding a combination of palatalised and non-palatalised consonants. When we initially plotted figures of the FPC scores, we observed that FPC1 correlates with *lowness* of F2–F1, i.e., *higher* FPC1 scores meant *lower* F2–F1. For ease of interpretation, we multiplied the FPC1 scores by -1 such that higher FPC1 scores can be interpreted as *greater* distance between F2 and F1 (i.e., more palatalisation).

The FPC1 scores for the difference between F2 and F1 trajectories were modelled as the dependent variable in a linear mixed model. All the discrete variables were sum coded (age group, vowel, palatalisation), and DDK was standardised. As DDK varied by PoA for all ages (Table II), we used the mean from three different productions as a

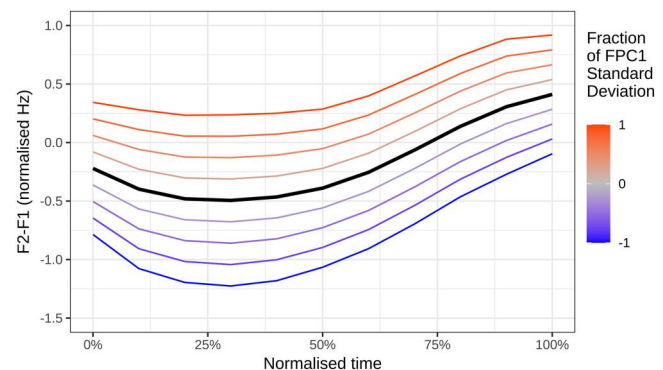


FIG. 1. Interpretation of FPC1 scores. The thick black line shows the mean measured F2. Red and blue lines are the mean F2 $\pm$ the SD of FPC1.

TABLE II. DDK modelling.

Full model intercept	$\hat{\beta}$	SE($\hat{\beta}$)	t	$p(t)$
	-1	0.19	-5.04	<0.001
Fixed effects		df	χ^2	$p(\chi^2)$
Age group		9	43.86	<0.001
PoA		8	24.78	0.002
Age group; PoA		6	6.37	0.38

representative value of their overall performance. Speaker and word were included as random intercepts. The FPC1 model included two-way interactions for palatalisation and vowel, and palatalisation and age group. We also included a three-way interaction for age group, DDK, and palatalisation to test for differing relationships between oral motor control and age in the realisation of the palatalisation contrast.

Where two-way interactions between predictors were significant, we carried out *post hoc* testing using the emmeans package using the Tukey method and effects plots using the effects and broom.mixed packages. The significant three-way interaction is interpreted *via* an effects plot. Our FPC1 analysis does not include sex because the data are normalised with the aim of making adult and child data comparable and also normalising for physical differences between participants. In adult Estonian, there are almost no differences according to social performance of gender in the realisation of palatalisation contrasts (Malmi *et al.*, 2023), so there is little reason to expect a gender difference among children. Additionally, we wished to avoid excessive model complexity where possible.

III. RESULTS

A. Oral motor control

The raw DDK values of syllables per second by age group and PoA are plotted in Fig. 2, the DDK modelling is presented in Table II, pairwise comparison between PoAs is presented in Table IV, and comparison between age groups is in Table III.

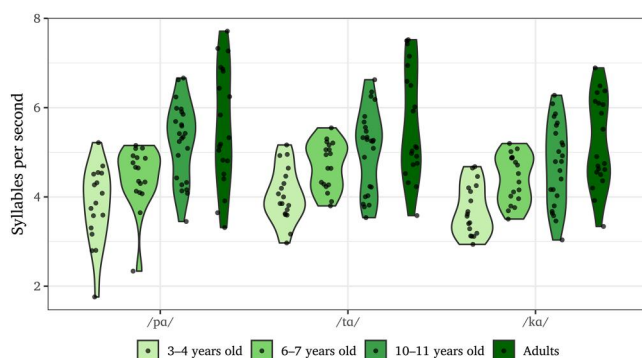


FIG. 2. DDK by PoA and age group.

TABLE III. DDK pairwise comparisons for age groups.

Age group	$\hat{\beta}$	SE($\hat{\beta}$)	t	p
3-4 years to 6-7 years	-0.61	0.24	-2.54	0.062
3-4 years to 10-11 years	-1.1	0.23	-4.72	<0.001
3-4 years to adults	-1.52	0.24	-6.45	<0.001
6-7 years to 10-11 years	-0.45	0.26	-2.01	0.195
6-7 years to adults	-0.90	0.24	-3.84	0.001
10-11 years to adults	-0.45	0.22	-2.10	0.176

Our analysis showed that age group is a significant predictor of the rate at which the participants could produce repetitions. The PoA was significant as a main effect but not in the interaction with age group indicating differences for PoA at all ages.

The youngest speakers (3-4 years) produced similar number of syllables per second (3.6-5.2, mean 4.2, SD 0.5) to 6- to 7-year-olds (3.8-5.6, mean 4.8, SD 0.5) but significantly fewer than 10- to 11-year-olds (3.5-6.7, mean 5.4, SD 0.8). The youngest speakers and 6- to 7-year-olds both produced significantly fewer syllables per second than adults (3.6-7.7, mean 6, SD 1.2) (Table III). Overall the repetition rates were slower with /ka/ than /pa/ or /ta/ (Table IV).

B. Acoustics of palatalisation

The FPC1 modelling is presented in Table V, pairwise comparisons for palatalisation and age group are in Table VI, and the pairwise comparisons of palatalisation, age group, and DDK are in Table VII. Figure 3 shows the effect of vowel on FPC1 values, Fig. 4 shows the model estimated values of FPC1 by age and palatalisation, and Fig. 5 shows the three-way interaction among age, palatalisation, and DDK.

The FPC1 model indicates there are significant main effects for age group, palatalisation, vowel and DDK. Also, age group, palatalisation, and DDK are significant in interactions with each other. Vowel is significant as a main effect but not in the interaction with palatalisation. The significant differences for vowel, but lack of significant interaction between vowel and palatalisation indicate that FPC1 values are different for /a/ and /u/ vowels and are consistently different preceding palatalised and non-palatalised consonants (Fig. 3). Figure 3 shows that FPC1 is higher in /a/ context, which corresponds to higher F2 and a more fronter PoA /a/ compared to /u/ in Estonian (see Fig. 1, lower panel, in Asu and Teras, 2009).

We conducted pairwise comparisons for the interaction of palatalisation and age group (Table VI), which indicates that all the age groups make a contrast between palatalised

TABLE IV. DDK pairwise comparisons for PoA.

PoA	$\hat{\beta}$	SE($\hat{\beta}$)	t	p
/ka/ - /pa/	-0.26	0.08	-3.10	<0.001
/ka/ - /ta/	-0.33	0.08	-4.11	<0.001
/pa/ - /ta/	-0.11	0.08	-0.92	0.630

TABLE V. FPC1 modelling.

Full model intercept	$\hat{\beta}$	SE($\hat{\beta}$)	t	$p(t)$
	-0.3	0.1	-3.79	<0.001
Fixed effects		df	χ^2	$p(\chi^2)$
Age group		12	191.06	<0.001
Palatalisation		9	346.08	<0.001
Vowel		2	28.11	<0.001
DDK		8	38.62	<0.001
DDK; age group		7	37.54	<0.001
Palatalisation; vowel		1	1.33	0.248
Palatalisation; age group		3	132.72	<0.001
Palatalisation; age group; DDK		4	31.81	<0.001

and non-palatalised consonants. The relevant model estimates are shown in Fig. 4. The model also shows that the magnitude and consistency of the contrast made increase with age. From the figure we see that for the youngest (three–four years), the FPC1 values for palatalised and non-palatalised productions show a higher difference than for the six- to seven-year-old group, but with less consistency. After this, the magnitude of the contrast starts gradually getting bigger.

To interpret the role of DDK in the contrast for palatalisation by age group, we plotted the model estimates from three-way interaction in Fig. 5. The figure indicates there is a relationship between DDK scores and palatalisation for the youngest children and adults in which age groups with higher DDK rates also made a bigger difference between the phonological contrast (Table VII). To contextualise the DDK data in this figure, we overlaid the mean DDK syllables per second values onto these model estimates. This highlights that the 3- to 4-year-olds and 6- to 7-year-olds have low DDK scores, but DDK scores from the 10- to 11-year-olds and adults occupy a greater range.

IV. DISCUSSION

We first discuss the acquisition of oral motor control in Estonian by age (Sec. IV A), then the acquisition of the acoustic targets for palatalisation (Sec. IV B), and finally the interaction of anatomical factors and oral motor control skills in the acquisition of complex consonants (Sec. IV C).

A. Oral motor control in Estonian

To answer the first research question whether there are age-related differences in the oral motor control, we used a

TABLE VI. FPC1 pairwise comparisons for palatalisation and age group.

Age group	$\hat{\beta}$	SE($\hat{\beta}$)	z	$p(z)$
3–4 years	-01.55	0.27	-5.85	<0.001
6–7 years	-0.93	0.17	-5.50	<0.001
10–11 years	-1.25	0.13	-9.82	<0.001
Adults	-2.37	0.13	-17.95	<0.001

TABLE VII. FPC1 pairwise comparisons for palatalisation, age group, and DDK.

Age group	$\hat{\beta}$	SE($\hat{\beta}$)	z	$p(z)$
3–4 years	-0.65	0.21	-2.72	0.001
6–7 years	0.42	0.24	1.74	0.082
10–11 years	0.03	0.09	0.28	0.777
Adults	-0.31	0.07	-4.62	<0.001

DDK speech task, which is widely used by speech and language therapists and researchers alongside other metrics to assess the speech motor coordination of typically and atypically developing children and adults (Chu *et al.*, 2021a; Ong *et al.*, 2024). We hypothesised, based on Ha (2023) and Murali *et al.* (2024), that child oral motor control skills would increase with age. This hypothesis was confirmed. We found that the number of repetitions (measured in syllables per second) of different syllables produced by children and adults in our experiment was directly linked to their age. Overall, the youngest speakers produced the fewest number of repetitions, and the oldest speakers produced the highest. This is in line with the results of the scoping review in Kent *et al.* (2022), which analysed many studies across different ages and language backgrounds. In line with research by Devadiga and Bhat (2012), we also found that there was a reduction in the repetition rates when the PoA was more posterior. There were significantly fewer repetitions with /ka/ than /pa/ or /ta/. This can be explained by the need to move a larger mass in velar consonants compared to bilabials or apicals.

B. Acquisition of secondary palatalisation

For the second research question, we asked whether there are any age-related differences in the acoustic realisation of secondary palatalisation. We hypothesised that consonants requiring multiple articulations or very precise articulator placement are among the last sounds to be acquired (McLeod and Crowe, 2018). Cross-linguistic surveys largely support earlier work predicting a hierarchical acquisition trajectory based on articulatory complexity (Shriberg, 1993). Later acquisition of complex consonants is thought to be because of the time taken to acquire auditory and somatosensory mappings on which the motor and muscle control gesture execution relies (Ayala *et al.*, 2023; Cheng *et al.*, 2007b; Klein *et al.*, 2013; Lin and Demuth, 2015). Specifically, through gradual increasing effectiveness of auditory and somatosensory feedback and anatomical development, children acquire better control of the superior longitudinal muscle (Abakarova *et al.*, 2022) and thus lingual differentiation (Cleland and Scobbie, 2021; Dokovova *et al.*, 2023; Gibbon, 1999) and sophisticated timing patterns (Kpogo and Chang, 2025; Noiray *et al.*, 2018). However, because of a generally flatter palate shape in younger children (Ciusa *et al.*, 2007), and a delayed change in proportion of the hard palate compared to the tongue (Vorperian *et al.*, 2005), a child's anatomy might naturally predispose them to

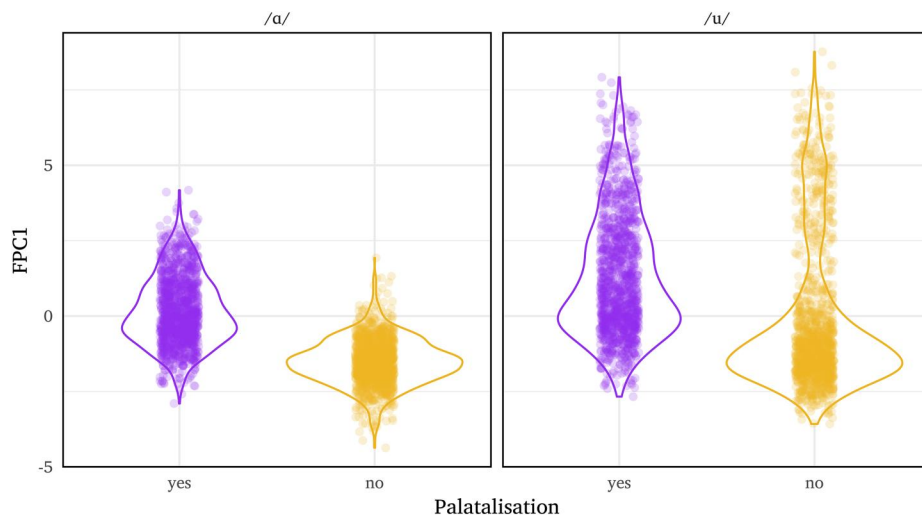


FIG. 3. FPC1 values for the different vowels.

palatalise. Previous research has indicated an overall palatalised quality to child speech in Estonian and Russian (Vihman and Vihman, 2011; Zharkova, 2004), but no detailed acoustic analysis of children's palatalisation acquisition at different ages had yet been carried out (see initial analysis in Zharkova, 2004).

Our results show that all children in our sample produce the acoustic differences between palatalised and non-palatalised consonants (Figs. 4 and 5, Table VI). Figure 4 indicates that the contrast develops across childhood as the distance between non-palatalised and palatalised FPC1 values is greater, but more variable, for three- to four-year-olds, narrower for six- to seven-year-olds and then greater again for older children. This curve is reflective of a learning pattern, where because of the constant adjustments to the feedback systems, the performance temporarily drops and then levels out again. Table VI shows there is a developmental increase in consistency for producing the contrast and in the magnitude of the differences between palatalised and non-palatalised. The z scores and standard errors from these comparisons show a clear developmental trajectory of

increasing acoustic distance and consistency for the contrast. Figure 4 also shows an increasing magnitude of contrast with chronological age based on the model output.

Figure 4 shows estimated values for FPC1 and demonstrates there are developmental differences for *non*-palatalised consonants as well as palatalised i.e., FPC1 values are higher for children than those of adults. These results support previous impressions that child speech has a palatal quality (Vihman and Vihman, 2011; Zharkova, 2004) and that palatalisation is acquired early. We suggest that because of anatomical reasons and a tendency toward greater palatal contact in child speech (Cheng *et al.*, 2007a; Ciusa *et al.*, 2007), children also have to learn to *not* palatalise in their acquisition of phonological contrast. We return to the role of anatomy *in combination* with motor control in Sec. IV C.

C. Interaction of motor control and anatomical development in complex consonants

The third research question about the relationship between motor control skills and palatalisation was designed

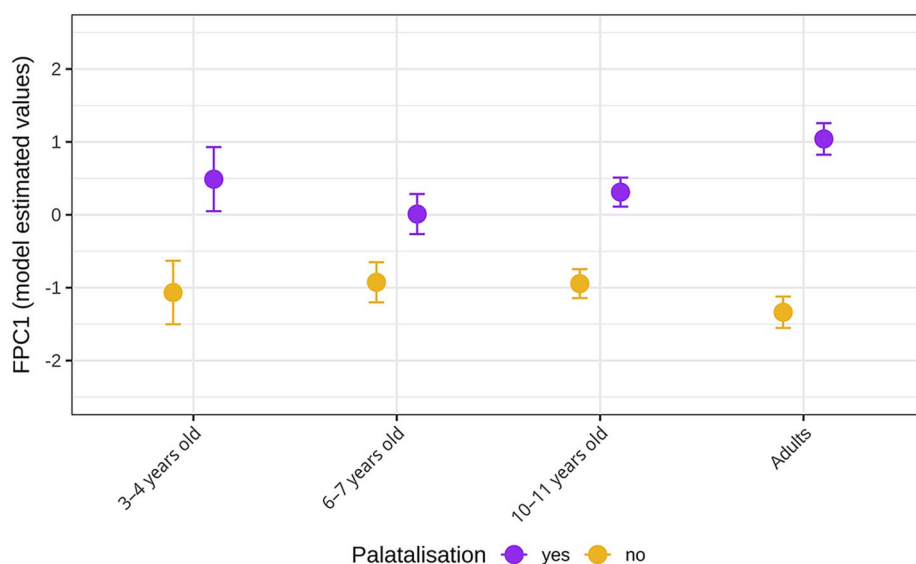


FIG. 4. Model estimates of the interaction between age group and palatalisation. Error bars show the 95% confidence intervals of the estimated values.

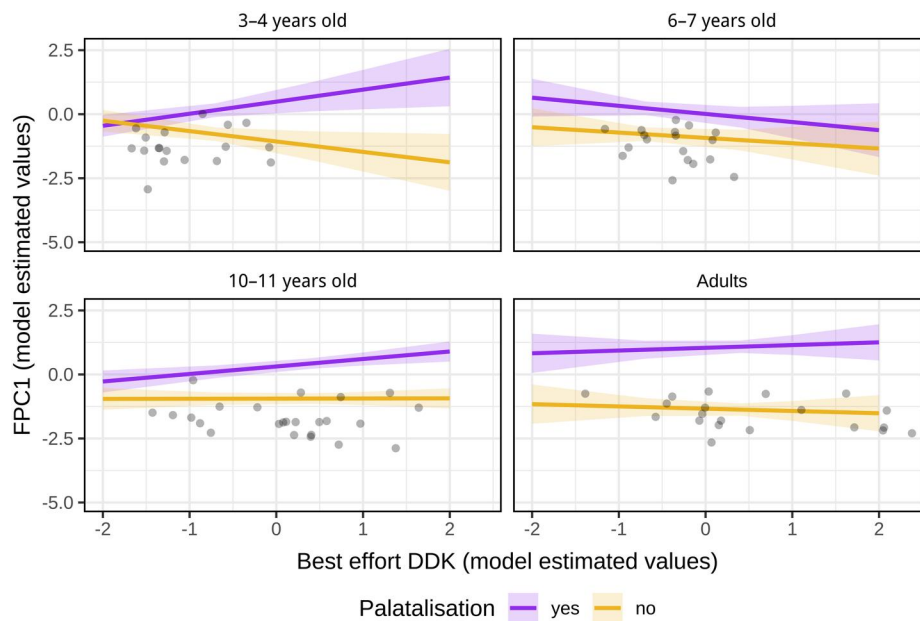


FIG. 5. Model estimates of the interaction among age group, palatalisation, and DDK with overlaid mean DDK values and 95% confidence intervals shown on the estimate values.

to test an apparent contradiction in the literature on child acquisition of consonants: on the one hand, complex articulations are typically acquired late because of developing lingual motor control, but on the other hand, palatalisation is reportedly acquired earlier because of anatomy. Figure 5 allows us to explore the relationship between the acquisition of the palatalisation contrast and oral motor control at different ages.

By interpreting the interaction between DDK, group, and palatalisation, we found there is a relationship for the youngest speakers and adults between the acquisition of oral motor control and the extent to which palatalisation is phonologically contrasted (Fig. 5, Table VII). This means that children with higher oral motor control needed for lingual differentiation in complex consonants realise the contrast to a greater degree. The result for adult speakers could be explained by their attempt in performing well in the tasks that were designed to be as easy for children as possible. This, in many cases, resulted in hyperarticulation of the palatalisation contrast. The results thus reflect how the adults, as our control group, behaved when tasked with an easy experiment with some participants being more diligent than others.

Our findings advance understanding of the acquisition of complex consonants and the roles of anatomy and motor control in cross-linguistic consonant acquisition. The results show that although an acoustic impression of palatalisation is acquired early (Kochetov and Alderete, 2011; Zharkova, 2004), the realisation of the phonological contrast between palatalised and non-palatalised takes time to acquire in adult-like fashion and proceeds in tandem with acquisition of oral motor control in young children.

D. Limitations and future work

Our study considers only the acoustic targets for palatalisation. We anticipate that the acoustic values for child

palatalised consonants are articulated differently from adult realisations and that children adapt their articulatory strategy as based on anatomy (Brunner *et al.*, 2009). Our future work will analyse ultrasound tongue imaging data from the same speakers to assess lingual differentiation and consistency of articulation directly. In addition, previous work on coarticulation indicates substantial timing differences between child and adult articulations (Noiray *et al.*, 2018; Rubertus and Noiray, 2024). Our future work will also explore the acquisition of timing of the palatalisation gesture. One limitation of our study is that we use chronological age as a proxy for predicted anatomical differences based on large-scale imaging studies such as Vorperian *et al.* (2005) or palatal studies (Ciusa *et al.*, 2007). Future work could also measure hard palate length as a more direct measurement of anatomical development. Another limitation was using a DDK task with monosyllables only. A task with trisyllabic utterances would have given us a more complete understanding of the development of the oral motor system. As we did not want to overburden children with too many tasks, we opted for using a simpler DDK protocol. In the production experiment with ultrasound we used audio recordings to elicit words from children. A more naturalistic approach might have yielded different results, but given our lengthy experiment protocol, elicitation by imitation was the most straightforward way to elicit reliable data.

V. CONCLUSION

Our analysis contributes to a long-standing discussion in child acquisition of phonology about the order of acquisition for different phonemes (Shriberg, 1993) and the role of physiological factors in speech production acquisition (Kent, 1992). Our aim was to disentangle the contribution of, on the one hand, motor control development and, on the other hand, anatomical growth and articulator change in proportion. We found that oral motor control, measured through a DDK task,

increases with age, and there is a relationship between greater oral motor control and realisation of the phonemic contrast for palatalisation for the youngest children. We suggest that explaining the developmental trajectory for palatalisation acquisition is possible through an interaction of motor control and physiological development: younger children are able to produce a phonetic impression of palatalisation from the age of three years. It is likely that the ratio of tongue to hard palate length predisposes them to produce acoustically palatal-sounding speech. However, at the same time, those who are most skilled in oral motor control are best able to differentiate palatalised and non-palatalised realisations. In this way, we argue that although anatomical development leads straightforwardly to the production of palatalised consonants, a high degree of motor control skills is also needed for adult-like realisation of the phonological contrast. In addition, we provide initial normative data for DDK in Estonian L1-speaking children and contribute a detailed acoustic account of child acquisition of a lesser-studied consonant contrast in a lesser-studied language family.

ACKNOWLEDGMENTS

The authors would like to thank all the participants, from youngest to oldest; the dedicated speech and language therapists who contributed their time and expertise to help us find suitable children; the schools and preschools that welcomed us to conduct the experiments; and the parents who trusted us to conduct research with their children. This study is a part of a project, “The articulatory properties of Estonian children’s palatalisation,” and was funded by the Estonian Research Council (Grant No. PUTJD1246).

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflicts to disclose.

Ethics Approval

The experiment was approved by the Research and Ethics Committee of the University of Tartu (397/T-7). We obtained written informed consent from all participants or from their parents or legal guardians.

DATA AVAILABILITY

The data that support the findings of this study are openly available at <https://osf.io/egmdj>.

- Abakarova, D., Fuchs, S., and Noiray, A. (2022). “Developmental changes in coarticulation degree relate to differences in articulatory patterns: An empirically grounded modeling approach,” *J. Speech. Lang. Hear. Res.* **65**(9), 3276–3299.
- Alumäe, T., Tilk, O., and Asadullah (2018). “Advanced rich transcription system for Estonian speech,” *arXiv:1901.0360*.
- Anwyl-Irvine, A., Dalmaijer, E. S., Hodges, N., and Evershed, J. K. (2021). “Realistic precision and accuracy of online experiment platforms, web browsers, and devices,” *Behav. Res.* **53**(4), 1407–1425.
- Asu, E. L., and Teras, P. (2009). “Estonian,” *J. Int. Phon. Assoc.* **39**(3), 367–372.

- Ayala, S. A., Eads, A., Kabakoff, H., Swartz, M. T., Shiller, D. M., Hill, J., Hitchcock, E. R., Preston, J. L., and McAllister, T. (2023). “Auditory and somatosensory development for speech in later childhood,” *J. Speech. Lang. Hear. Res.* **66**(4), 1252–1273.
- Barbier, G., Perrier, P., Payan, Y., Tiede, M. K., Gerber, S., Perkell, J. S., and Ménard, L. (2020). “What anticipatory coarticulation in children tells us about speech motor control maturity,” *PLoS One* **15**(4), e0231484.
- Bateman, N. (2011). “On the typology of palatalization,” *Lang. Linguist. Compass* **5**(8), 588–602.
- Bennett, R., Padgett, J., McGuire, G., Ní Chiosáin, M., and Bellik, J. (2025). “Gestural timing and contrast: An Irish case study,” in *Foundational Approaches to Celtic Linguistics*, edited by A. Carnie, D. Ohala, D. Hunter, S. Prins, M. Hammond, and L. Irizarry (Language Science Press, Berlin), pp. 283–341.
- Bishai, R., Molenaar, A. M., and Scheerer, N. E. (2026). “Exploring the impact of attention and sensory processing abilities on vocal motor control in children,” *J. Acoust. Soc. Am.* **159**(3), 2902–2916.
- Boersma, P., and Weenik, D. (2025). “Praat: Doing phonetics by computer (version 6.4.27) [computer program],” <http://www.praat.org/> (Last viewed 16 September 2026).
- Brunner, J., Fuchs, S., and Perrier, P. (2009). “On the relationship between palate shape and articulatory behavior,” *J. Acoust. Soc. Am.* **125**(6), 3936–3949.
- Cheng, H. Y., Murdoch, B., Goozée, J., and Scott, D. (2007a). “Electropalatographic assessment of tongue-to-palate contact patterns and variability in children, adolescents, and adults,” *J. Speech. Lang. Hear. Res.* **50**, 375–392.
- Cheng, H. Y., Murdoch, B., Goozée, J., and Scott, D. (2007b). “Physiological development of tongue-jaw coordination from childhood to adulthood,” *J. Speech Lang. Hear. Res.* **50**, 352–360.
- Chu, S. Y., Foong, J. H., Lee, J., Ben-David, B. M., Barlow, S. M., and Hsu, C. (2021a). “Oral diadochokinetic rates across languages: Multilingual speakers comparison,” *Intl. J. Lang. Comm. Disord.* **56**(5), 1026–1036.
- Chu, S. Y., Lee, J., Barlow, S. M., Ben-David, B., Lim, K. X., and Foong, J. H. (2021b). “Oral-diadochokinetic rates among healthy Malaysian-Mandarin speakers: A cross linguistic comparison,” *Int. J. Speech Lang. Pathol.* **23**(4), 419–429.
- Ciusa, V., Dimaggio, F. R., Sforza, C., and Ferrario, V. F. (2007). “Three-dimensional palatal development between 3 and 6 years,” *Angle Orthod.* **77**(4), 602–606.
- Cleland, J., and Scobbie, J. (2021). “The dorsal differentiation of velar from alveolar stops in typically developing children and children with persistent velar fronting,” *J. Speech. Lang. Hear. Res.* **64**(6S), 2347–2362.
- Coretta, S., and Sakr, G. (2026). “Multivariate analyses of tongue contours from ultrasound tongue imaging,” *Lang. Speech* (published online).
- Cychosz, M. (2021). “The coarticulation-duration relationship in early Quechua speech,” *J. Phon.* **87**, 101052.
- Davis, B., and Bedore, L. (2013). *An Emergence Approach to Speech Acquisition: Doing and Knowing* (Psychology Press, New York).
- Devadiga, D. N., and Bhat, J. (2012). “Oral diadochokinetic rate—an insight into speech motor control,” *Int. J. Adv. Speech Hearing Res.* **1**(1), 10–14.
- Dokovova, M., Sugden, E., Cartney, G., Schaeffler, S., and Cleland, J. (2023). “Tongue shape complexity in children with and without speech sound disorders,” *J. Speech. Lang. Hear. Res.* **66**(7), 2164–2183.
- Duffy, J. R. (2013). *Motor Speech Disorders: Substrates, Differential Diagnosis, and Management*, 3rd ed. (Elsevier, St. Louis, MO).
- Fletcher, S. G. (1989). “Palatometric specification of stop, affricate, and sibilant sounds,” *J. Speech. Lang. Hear. Res.* **32**(4), 736–748.
- Gadesmann, M., and Miller, N. (2008). “Reliability of speech diadochokinetic test measurement,” *Intl. J. Lang. Comm. Disord.* **43**(1), 41–54.
- Gafos, A. I. (2002). “A grammar of gestural coordination,” *Nat. Lang. Linguist. Theory* **20**(2), 269–337.
- Gibbon, F. E. (1999). “Undifferentiated lingual gestures in children with articulation/phonological disorders,” *J. Speech. Lang. Hear. Res.* **42**(2), 382–397.
- Gubian, M. (2025). “landmarkregUtils,” <https://github.com/uasolo/landmarkregUtils> (Last viewed 16 September 2026).
- Gubian, M., Torreira, F., and Boves, L. (2015). “Using functional data analysis for investigating multidimensional dynamic phonetic contrasts,” *J. Phon.* **49**, 16–40.
- Guenther, F. H. (2006). “Cortical interactions underlying the production of speech sounds,” *J. Commun. Disord.* **39**(5), 350–365.

- Ha, S. (2023). "Oral diadochokinetic production in children with typical speech development and speech-sound disorders," *Intl. J. Lang. Comm. Disord.* **58**(5), 1783–1798.
- Happ, C., and Greven, S. (2018). "Multivariate functional principal component analysis for data observed on different (dimensional) domains," *J. Am. Stat. Assoc.* **113**(522), 649–659.
- Howson, P., and Redford, M. A. (2021). "The acquisition of articulatory timing for liquids: Evidence from child and adult speech," *J. Speech. Lang. Hear. Res.* **64**(3), 734–753.
- Iskarous, K., and Kavitskaya, D. (2018). "Sound change and the structure of synchronic variability: Phonetic and phonological factors in Slavic palatalization," *Language* **94**(1), 43–83.
- Jolliffe, I. T., and Cadima, J. (2016). "Principal component analysis: A review and recent developments," *Phil. Trans. R Soc. A* **374**(2065), 20150202.
- Kabakoff, H., Harel, D., Tiede, M., Whalen, D. H., and McAllister, T. (2021). "Extending ultrasound tongue shape complexity measures to speech development and disorders," *J. Speech. Lang. Hear. Res.* **64**(7), 2557–2574.
- Kanarbik, T. (2026). "Diadochokineesi sooritus tavaarenguga ükskeelsetel eesti keelt emakeelena kõnelevatel lastel" ("Oral diadochokinetic rates in typically developing Estonian-speaking children"), Master's thesis, University of Tartu, Tartu, Estonia.
- Kent, R. (1992). "The biology of phonological development," in *Phonological Development: Models, Research, Implications*, edited by C. Ferguson, L. Menn, and C. Stoel-Gammon (York Press, Timonium, MD), pp. 65–90.
- Kent, R. D., Kim, Y., and Chen, L.-m (2022). "Oral and laryngeal diadochokinesis across the life span: A scoping review of methods, reference data, and clinical applications," *J. Speech. Lang. Hear. Res.* **65**(2), 574–623.
- Kidd, E., and Garcia, R. (2022). "How diverse is child language acquisition research?" *First Lang.* **42**(6), 703–735.
- Klein, H. B., McAllister Byun, T., Davidson, L., and Grigos, M. I. (2013). "A multidimensional investigation of children's /r/ productions: Perceptual, ultrasound, and acoustic measures," *Am. J. Speech. Lang. Pathol.* **22**(3), 540–553.
- Kochetov, A. (2002). "Production, perception, and emergent phonotactic patterns: A case of contrastive palatalization," Ph.D. thesis, University of Toronto, Toronto, Ontario, Canada.
- Kochetov, A. (2017). "Acoustics of Russian voiceless sibilant fricatives," *J. Int. Phon. Assoc.* **47**(3), 321–348.
- Kochetov, A., and Alderete, J. (2011). "Patterns and scales of expressive palatalization: Experimental evidence from Japanese," *Can. J. Linguist.* **56**(3), 345–376.
- Kpogo, F., and Chang, C. B. (2025). "Coarticulation and coordination in phonological development: Insights from children's and adults' production of complex-simple stop contrasts in Gã," *J. Phon.* **108**, 101378.
- Lahtein-Kürsa, M., Padrik, M., Daniute, S., Kairiene, D., Martikainen, A.-L., Vanhala-Haukijärvi, M., and Mailend, M.-L. (2025). "Diagnostic features of childhood apraxia of speech: A survey study of Estonian, Finnish, and Lithuanian speech-language pathologists," *Am. J. Speech. Lang. Pathol.* **34**(1), 97–117.
- Liiv, G. (1965). "Preliminary remarks on the acoustic cues for palatalization in Estonian," *Phonetica* **13**(1–2), 59–64.
- Lin, S., and Demuth, K. (2015). "Children's acquisition of English onset and coda /l/: Articulatory evidence," *J. Speech. Lang. Hear. Res.* **58**(1), 13–27.
- Malmi, A., and Lippus, P. (2019). "Keele asend eesti palatalisatsioonis" ("The position of the tongue in Estonian palatalization"), *ESUKA-JEFUL* **10**(1), 105–128.
- Malmi, A., Lippus, P., and Meister, E. (2023). "Spectral and temporal properties of Estonian palatalization," *J. Int. Phon. Assoc.* **53**(3), 748–773.
- Malmi, A., and Nance, C. (2025). "The production of articulatorily complex consonants in Estonian by children and adults," in *Proceedings of Fonetik 2025, May 21-23, Linnaeus University, Växjö, Sweden*, edited by G. Ambrazaitis, F. Raschella, and N. J. Young (Linnaeus University Press, Växjö, Sweden), pp. 33–38.
- Malmi, A., and Nance, C. (2026). "The role of oral motor control in developing acoustic targets [database]," <https://osf.io/egmdj/overview> (Last viewed 16 September 2026).
- McLeod, S., and Crowe, K. (2018). "Children's consonant acquisition in 27 languages: A cross-linguistic review," *Am. J. Speech. Lang. Pathol.* **27**(4), 1546–1571.
- Meister, E., and Meister, L. (2021). "Developmental changes of vowel acoustics in adolescents," *Proc. Interspeech 2021*, pp. 4014–4018.
- Mendes Tavares, E. L., Brasolotto, A. G., Rodrigues, S. A., Benito Pessin, A. B., and Garcia Martins, R. H. (2012). "Maximum phonation time and s/z ratio in a large child cohort," *J. Voice* **26**(5), 675.e1–675.e4.
- Murali, M., Cleland, J., Young, D., Stuart-Smith, J., Taylor, L., Smith, A., and Kuschmann, A. (2024). "Comparing DDK rate variability in typically developing children in central Scotland," in *Proceedings of Conference on Motor Speech*, Madonna Rehabilitation Hospitals, February 21–24, San Diego, CA.
- Nance, C., and Kirkham, S. (2020). "The acoustics of three-way lateral and nasal palatalisation contrasts in Scottish Gaelic," *J. Acoust. Soc. Am.* **147**(4), 2858–2872.
- Noiray, A., Abakarova, D., Rubertus, E., Krüger, S., and Tiede, M. (2018). "How do children organize their speech in the first years of life? Insight from ultrasound imaging," *J. Speech. Lang. Hear. Res.* **61**(6), 1355–1368.
- Ong, Y. Q., Lee, J., Chu, S. Y., Chai, S. C., Gan, K. B., Ibrahim, N. M., and Barlow, S. M. (2024). "Oral-diadochokinesis between Parkinson's disease and neurotypical elderly among Malaysian-Malay speakers," *Int. J. Lang. Commun. Disord.* **59**(5), 1701–1714.
- Ramsay, J. (2003). "FDA: Functional data analysis," <https://cran.r-project.org/web/packages/fda/index.html> (Last viewed 16 September 2026).
- R Core Team (2025). "The R project for statistical computing[computer program]," www.R-project.org (Last viewed 16 September 2026).
- Rubertus, E., and Noiray, A. (2018). "On the development of gestural organization: A cross-sectional study of vowel-to-vowel anticipatory coarticulation," *PLoS One* **13**(9), e0203562.
- Rubertus, E., and Noiray, A. (2024). "Children's coarticulation patterns as a window to the phonology-phonetics interface," in *Proceedings of the International Seminar on Speech Production*, 13–17 May, Autrans, France, pp. 43–46.
- Shaw, J., Sejin, O., Durvasula, K., and Kochetov, A. (2021). "Articulatory coordination distinguishes complex segments from segment sequences," *Phonology* **38**, 437–477.
- Shriberg, L. D. (1993). "Four new speech and prosody-voice measures for genetics research and other studies in developmental phonological disorders," *J. Speech. Lang. Hear. Res.* **36**(1), 105–140.
- Spreafico, L., Pucher, M., and Matosova, A. (2018). "UltraFit: A speaker-friendly headset for ultrasound recordings in speech science," *Proc. Interspeech 2018*, 1517–1520.
- Stackhouse, J. P. W. (2000). "Rate, accuracy and consistency: Diadochokinetic performance of young, normally developing children," *Clin. Linguist. Phon.* **14**(4), 267–293.
- Teras, P., and Pajusalu, K. (2014). "Palatalisatsioonist ja prepalatalisatsioonist spontaanses eesti keeles" ("Palatalization and prepalatalization in Estonian spontaneous speech"), *Keel ja Kirjandus* **4**, 257–269.
- Timm, L. (1977). "A child's acquisition of Russian phonology," *J. Child Lang.* **4**(3), 329–339.
- Tremblay, S., Shiller, D. M., and Ostry, D. J. (2003). "Somatosensory basis of speech production," *Nature* **423**(6942), 866–869.
- Turton, D. (2017). "Categorical or gradient? An ultrasound investigation of /l/-darkening and vocalization in varieties of English," *J. Assoc. Lab. Phonol.* **8**(1), 13.
- Vihman, M., and Vihman, V.-A. (2011). "From first words to segments: A case study in phonological development," in *Experience, Variation and Generalization: Learning a First Language*, edited by I. Aron and E. Clark (John Benjamins, Amsterdam), pp. 109–133.
- Vorperian, H. K., Kent, R. D., Lindstrom, M. J., Kalina, C. M., Gentry, L. R., and Yandell, B. S. (2005). "Development of vocal tract length during early childhood: A magnetic resonance imaging study," *J. Acoust. Soc. Am.* **117**(1), 338–350.
- Winter, B. (2020). *Statistics for Linguists: An Introduction Using R* (Routledge, New York).
- Wrench, A. (2024). "Articulate Assistant Advanced v221.5.11 [computer program]," <https://articulateinstruments.com/> (Last viewed 16 September 2026).
- Zharkova, N. (2004). "Strategies in the acquisition of segments and syllables in Russian-speaking children," *Leiden Papers Linguistics* **2**(1), 189–213.
- Zharkova, N. (2018). "An ultrasound study of the development of lingual coarticulation during childhood," *Phonetica* **75**(3), 245–271.
- Ziegler, W., Schölderle, T., Brendel, B., Amsellem, J., and Et, A. (2019). "Higher-faster-farther: Maximum performance tests in the assessment of neurogenic speech impairment," *Folia Phoniatr. Logop.* **71**, 261.