JSLHR

Research Article

1861

How Do Children Organize Their Speech in the First Years of Life? Insight From Ultrasound Imaging

Aude Noiray,[a,b] Dzhuma Abakarova,[a] Elina Rubertus,[a]
Stella Krüger,[a] and Mark Tiede[b]

Purpose: This study reports on a cross-sectional investigation of lingual coarticulation in 57 typically developing German children (4 cohorts from 3.5 to 7 years of age) as compared with 12 adults. It examines whether the organization of lingual gestures for intrasyllabic coarticulation differs as a function of age and consonantal context.
Method: Using the technique of ultrasound imaging, we recorded movement of the tongue articulator during the production of pseudowords, including various vocalic and consonantal contexts.
Results: Results from linear mixed-effects models show greater lingual coarticulation in all groups of children as compared with adults with a significant decrease from the kindergarten years (at ages 3, 4, and 5 years) to the end of the 1st year into primary school (at age 7 years). Additional differences in coarticulation degree were found across and within age groups as a function of the onset consonant identity (/b/, /d/, and /g/).
Conclusions: Results support the view that, although coarticulation degree decreases with age, children do not organize consecutive articulatory gestures with a uniform organizational scheme (e.g., segmental or syllabic). Instead, results suggest that coarticulatory organization is sensitive to the underlying articulatory properties of the segments combined.

In the domain of spoken language acquisition, great attention has been focused on coarticulation, which concerns the overlapping of articulatory gestures for neighboring segments (for a review, see Hardcastle & Hewlett, 2006). Coarticulation is an important characteristic of fluent speech. It is an important mechanism to investigate as it taps into the phonetic instantiations of phonological units from various sizes, such as phonemes or syllables, and, therefore, offers a chance to reveal how unit organization matures over time as children learn to speak their native language. In addition, coarticulation engages multiple speech articulators (e.g., the lips, the tongue) whose actions must be coordinated in time and in the space of the vocal tract to produce intelligible phonetic outputs in the native language. Investigating the development of coarticulatory patterns therefore provides a unique opportunity to address both the maturation of the speech motor system and its attunement to the phonetic regularities of the language spoken.

In this study, we were specifically interested in examining how differences in lingual vowel-to-consonant coarticulation can shed light on the phonetic organization of speech in young German children, from 3 years of age (when they are in kindergarten) to 7 years of age when children are in primary school. In addition, we aimed to provide a first quantitative survey of anticipatory coarticulation in German learners. Much developmental work on lingual coarticulation has focused on English variants. However, studies in languages other than English are needed to possibly disentangle universal versus language-specific patterns of coarticulation. In German at least, most assessments of consonant acquisition have used measures of individual production accuracy (e.g., Fox-Boyer, 2006). Using the technique of ultrasound imaging, we examined the organization of gestures of the tongue, an organ whose control is essential to vowels' and consonants' acquisition (e.g., Barbier et al., 2015; Klein, Byun, Davidson, & Grigos, 2013; Ménard & Noiray, 2011; Noiray, Ménard, & Iskarous, 2013; Song, Demuth, Shattuck-Hufnagel, & Ménard, 2013; Zharkova, Hewlett, Hardcastle, & Lickley, 2014). In the

---

[a] Laboratory for Oral Language Acquisition, University of Potsdam, Germany
[b] Haskins Laboratories, New Haven, CT

Correspondence to Aude Noiray: anoiray@uni-potsdam.de

Editor-in-Chief: Julie Liss
Editor: Natalia Zharkova

Received April 19, 2017
Revision received September 11, 2017
Accepted February 7, 2018
https://doi.org/10.1044/2018_JSLHR-S-17-0148

Disclosure: The authors have declared that no competing interests existed at the time of publication.

past, articulatory tracking methods, such as electromagnetic articulography (EMA) and electropalatography (EPG) have been employed in school-aged children and adolescents (e.g., EPG: Cheng, Murdoch, Goozée, & Scott, 2007; EMA: Katz & Bharadwaj 2001; Terband, Maassen, Van Lieshout, & Nijland, 2011). More recently, the technique of ultrasound imaging has been adapted to the developmental field to make articulatory recordings of the tongue possible in young populations (e.g., Barbier et al., 2015; Ménard & Noiray, 2011; Noiray et al., 2013; Song et al., 2013; Zharkova, 2017). Compared to EMA and EPG, ultrasound is a more suitable technique to use with young children because it requires neither long preparation time prior to testing nor invasive procedures to track tongue movement during speech (e.g., gluing EMA pellets on young children's tongues or placing an artificial palate). Hence, with the technique of ultrasound imaging, it is now possible to revisit questions related to speech organization in young children while directly examining the articulatory mechanisms underlying speech production rather than inferring those mechanisms from the acoustic outputs.

Units of Speech Production in Children?

Finding the units of speech organization in the first years of life has been one of the most challenging endeavors for developmental psycholinguists, but the quest is important for advancing both theories of language acquisition and clinical assessment of disordered speech.

Over the past two decades, research examining intrasyllabic coarticulatory patterning in typically developing (TD) children has provided conflicting results and hypotheses regarding the nature of these units. A number of studies have reported less coarticulation in children compared with adults with limited influence of the vowel on the preceding consonant. Such findings lead to the hypothesis that spoken language organization is initially segmentally driven (e.g., Gibson & Ohde, 2007; Green, Moore, & Reilly, 2002; Katz, Kripke, & Tallal, 1991; Kent, 1983). In this view, children are supposed to proceed through a sequential maturation process by which articulatory controls for individual segments progressively develop into more complex interarticulator organizations for larger units, with increasing intrasyllabic coarticulation as a result. An opposite view holds that children initially display greater consonant–vowel (CV) coarticulation than adults, suggesting a broader planning unit of their speech than the segmental unit (e.g., Goodell, & Studdert-Kennedy, 1993; Nijland et al., 2002; Nittrouer, Studdert-Kennedy, & Neely, 1996; Nittrouer & Whalen, 1989; Rubertus, Abakarova, Tiede, & Noiray, 2015). In this more holistic perspective, maturation of coarticulatory patterns would consist in decreasing encroachment between consonantal and vocalic components and development of increasingly differentiated controls over individual articulators for a more segmental organization of articulatory movements. Other studies have found equivalent patterns of coarticulatory degree in children and adults but reported greater variability in children's patterns (e.g., Katz et al., 1991; Munson, 2004; Repp, 1986; Sereno, Baum, Marean, & Lieberman, 1987). To date, organizational units of speech production are still discussed.

Additional research is evidently needed not only to disentangle the origin(s) of current theoretical discrepancies but also because a detailed understanding of lingual coarticulatory development over age in TD children would provide useful information for advancing detection of atypical trajectories (Maas & Mailend, 2017). Indeed, a series of experimental studies conducted by Nijland et al. revealed inconsistent coarticulatory organization in children with childhood apraxia of speech (CAS) compared with TD controls (e.g., Nijland, Maassen, & van der Meulen, 2003). Whereas some children with CAS seem to exhibit greater coarticulation than TD children, others show the opposite patterns. More recently, Terband (2017) examined coarticulatory patterns from 16 children with CAS aged between 5.5 and 7.5 years producing /bi, di, bu, du/ and reported greater coarticulation than the eight age-matched TD children tested for comparison. However, deviant coarticulatory patterns in children with CAS were only observed in some phonetic contexts but not all. This suggests that the deficit in anticipatory coarticulation observed in children with CAS is not uniform as assumed in American Speech-Language-Hearing Association descriptions (American Speech-Language-Hearing Association, 2007) but specific to certain phoneme combinations that may involve more complex articulatory coordinations compared with others. Difficulty in coarticulatory organization has also been noticed in children who stutter (e.g., Soo-Eun, Ohde, & Conture, 2002) who seem to exhibit smaller coarticulatory differences across consonantal contexts than TD age-matched children. Taken together, these results have important implications as to the links between the articulatory properties of the speech material investigated, speech motor control, and the breadth of coarticulatory organization in atypical development. To provide reference data in German, this study focuses on lingual vowel-to-consonant coarticulation in TD children.

Differences in Lingual Coarticulation Degree and Resistance Across Consonantal Contexts

An important variable to consider when investigating variance in lingual coarticulatory patterns within a sample of participants or across populations regards the articulatory properties of the sequences produced. In adults, differences in intrasyllabic coarticulation degree (CD) within individuals reflect differences in consonants' place of articulation with labial-V syllables showing a high CD contrary to alveolar or alveopalatal stop-V syllables, which show a lower degree of coarticulation between consonantal and vocalic lingual gestures. Interestingly, in adults, these patterns have been consistently reported across various languages (e.g., in American English: Fowler, 1994; Fowler & Brancazio, 2000; Iskarous, Shadle, & Proctor, 2011; Australian languages: Graetzer, 2006; Canadian French: Noiray et al., 2013; Catalan: Recasens, 1985; Recasens &

Espinosa, 2009; German: Abakarova, Iskarous, & Noiray, 2017; Iskarous, Fowler, & Whalen, 2010; Swedish: Lindblom & Sussman, 2012; Thai, Cairene Arabic, and Urdu: Sussman, Hoemeke, & Ahmed, 1993).

A main hypothesis is that differences in CD are related to the degree of coarticulatory resistance of the consonant (e.g., Bladon & Al-Bamerni, 1976; Fowler, 1994; Fowler & Brancazio, 2000; Recasens, 1985; Recasens & Espinosa, 2006). In this view, resistance varies across consonants' place and manner of articulation as a result of differences in articulatory demands on target articulators, which affect their degree of temporal and spatial overlap with adjacent segments (e.g., Fowler & Saltzman, 1993). The more constraints on an articulator are involved in the production of a consonant, the more resistant the consonant may be to large coarticulatory overlap with contiguous vowels. For example, in adult speakers, alveolar stops /t, d/ resist lingual coarticulation with adjacent vowels more than labial stops /p, b/ do (e.g., Iskarous et al., 2010; Recasens, 1985; Sussman et al., 1993). Various studies across languages have outlined that palatal consonants, such as [ɲ], exert more coarticulatory resistance than alveolars, such as [n] (in Catalan: Recasens & Rodríguez, 2016; in English: Fowler & Brancazio, 2000). These findings corroborate the predictions from the degree of articulatory constraints model of coarticulation (Recasens, 1999), which directly relates the degree and direction of coarticulation to the demands imposed on the tongue body for consecutive articulatory gestures. Taken together, results suggest that CD varies along a continuum depending on consonant identity and its degree of resistance to coarticulation with adjacent segments (for a detailed discussion, see Iskarous et al., 2013). Given that adults' coarticulatory organization depends on the interaction of articulatory gestures, any study looking at the ontogenetic development of coarticulation would gain in explanatory power by considering articulatory gestures' intrinsic properties from which coarticulatory overlap versus resistance originates.

There is converging evidence from acoustic studies (e.g., Nittrouer, 1993, 1995; Reidy, 2015; Sussman, Duder, Dalston, & Cacciatore, 1999) and articulatory studies (e.g., in 7- and 5-year-old American English children: Katz & Bharadwaj, 2001; in 5- and 13-year-old Scottish children: Zharkova, Hardcastle, Gibbon, and Lickley, 2015; in 4–5-year-old Canadian French children: Noiray et al., 2013) that children's coarticulatory patterns differ across phonetic contexts. Greater lingual anticipatory coarticulation is observed for heterorganic sequences (e.g., labial CV syllables) compared with homorganic sequences for which the same target organ is recruited for both the consonant and vowel (e.g., alveolar CV). Studies employing the locus equation approach in child speech (mostly in English) have reported a decreasing degree of coarticulation from labial to velar to alveolar stops (e.g., Goodell & Studdert-Kennedy, 1993; Sussman et al., 1999), with labial and velar stops sometimes yielding similar CD depending on speakers (e.g., Noiray et al., 2013; review in Gibson & Ohde, 2007). Interestingly, alveolar stops considered as more resistant to coarticulation than labials show a decrease in CD with age. One hypothesis is that children progressively develop synergistic relationships among muscles and functional subparts of the tongue (e.g., tongue body and tongue tip) to achieve lingual constrictions (e.g., the tongue body moving front to support the tongue tip in achieving the alveolar constriction; Noiray et al., 2013). This point will be further addressed in the Discussion.

With respect to fricatives, results diverge across studies. Some report age-related differences in coarticulation (e.g., Maas & Mailend, 2017; Nittrouer, Studdert-Kennedy, & McGowan, 1989) with greater CD in children than in adults (e.g., in English: Nittrouer et al., 1996), whereas others do not (e.g., Katz et al., 1991). In German, a recent acoustic study investigating coarticulation between fricatives /s/ or /ʃ/ and vowels in children ages 4 to 6 years reported greater vocalic influence over fricatives in preschoolers than in adults (Kleber, 2015). In general, fricatives are complex consonants, whose productions stabilize later than stops' (Nittrouer, 1995), especially labials that are present in the early babbling repertoire (in German: Fox-Boyer, 2006, 2009; in English: Prather, Hendrick, & Kern, 1975; Stoel-Gammon & Dunn, 1985).

Overall, available evidence suggests that children's coarticulatory patterns exhibit sensitivity to contextual effects early in age. Hence, the question of early coarticulatory organization may be framed like in adults as gradient distinctions along a continuum rather than supporting a binary organization as often suggested in the developmental literature (segmental vs. syllabic).

Research Questions

Given the theoretical findings outlined above, our study asked the following questions: First, does lingual CD overall differ in German children as compared with German adults? Given preschoolers' immature phonological and speech motor systems (e.g., Smith, 2010), we expected children's coarticulatory patterns to differ significantly from adults. The discrepancies found as to whether children generally coarticulate equally more or less than adults made it difficult to formulate specific expectation regarding the direction of the difference. Second, we tested whether CD systematically differs between children and adults regardless of the onset consonant identity. Given that most developmental studies have investigated age-related differences in lingual coarticulation in one or two phonetic contrasts (e.g., /ti, ta/; Zharkova, 2017; /si, su, ʃi, ʃu/; Nittrouer et al., 1996), we expanded this research to consonants varying in places and manners of articulation.

Third, we examined whether CD in children varies as a function of the consonant's identity as it does in adults across languages. Taking preliminary results in Canadian French preschoolers (Noiray et al., 2013), we expected German children to show modulations in CD according to consonants' place and manner of articulation with a lower CD in consonantal contexts that have been shown in adults to resist coarticulatory overlap and greater CD in velar and labial contexts supporting large coarticulatory overlap.

Method

Participants

The production task was administered to four cohorts of children (total: 57) and one adult cohort (total: 12). The groups consisted of seventeen 3-year-old children (age range = 3;05 [years;months] to 3;07, *M* = 3;06), fourteen 4-year-old children (age range = 4;05 to 4;07, *M* = 4;06), thirteen 5-year-old children (age range = 5;05 to 5;07, *M* = 5;06), and 13 children at the end of the first school year (last month and a half) or beginning of second year (first month and a half; age range = 7;00 to 7;05, *M* = 7;02). The latter will be referred to as Grade 1 children. All five cohorts were monolingual German speakers. Parental questionnaires ensured that none of the participants had any language-related, hearing, or visual impairment. The 12 German adults (age range = 19 to 34 years, *M* = 25;08) also presented no history in language or hearing impairments.

Stimulus Material

A German female model speaker recorded production material. The stimuli were presented auditorily in a repetition task to all participants. They consisted of disyllabic trochaic C1VC2 pseudowords embedded in a carrier phrase with the German female article /aɪnə/ (e.g., "eine bide"). There were three stop consonants: /b/, /d/, and /g/. For the 4-year-old group, the school-aged children, and adults, the additional consonant /z/ was included. We also used the tense and long vowels /iː/, /yː/, /eː/, /aː/, /uː/, and /oː/. C1Vs were designed as a fully crossed set of Cs and Vs, whereas the second C2ə syllable was added in a way that C1 was not the same consonant as C2. Intrasyllabic coarticulation was measured in the first CV syllable, between C1 and V. We aimed for all children to repeat the CV syllables six times, whereas adults produced nine repetitions. Pseudowords were presented in randomized blocks to prevent habituation effects. To make the recording playful and meaningful for our young participants, stimuli were presented as a new language that they would use with aliens during a space journey. This scenario fitted the experimental procedure developed in our lab very well.

Experimental Procedure

Recordings took place at the Laboratory for Oral Language Acquisition at University of Potsdam (Germany) in an experimental room that is well suited for child studies and decorated to match our space journey storyline. All participants were recorded within the Sonographic and Optical Linguo-Labial Articulation Recording (SOLLAR) platform (Noiray, Ries, & Tiede, 2015). SOLLAR is a multi-data recording platform embedded into a spaceship to stimulate children's interest. This child-friendly platform allows for the recording of the audio speech signal (microphone Shure, sr.: 48 kHz), tongue movement via ultrasound imaging (Sonosite scanner, sr.: 48 Hz), and labial-shape tracking via video recording (camera SONY, sr.: 50 Hz). The same technical setup was used for adults to ensure similar experimental conditions for all participants. The audio signal was recorded in relation to two devices: (a) synchronous with the ultrasound device and (b) synchronous with the video camera. This information was used to generate a universal time code for all data. Video and acoustic signals were then synchronized using cross-correlation function within MATLAB. This method has been reliably used in previous speech production studies (e.g., in adults: Noiray, Cathiard, Ménard & Abry, 2011; Noiray, Iskarous, & Whalen, 2014; in children: Noiray, Cathiard, Abry, Ménard, & Savariaux, 2008; Rubertus et al., 2015).

In SOLLAR, the ultrasound probe is positioned below participants' chins to record the tongue surface contour on the midsagittal plane. It is placed in a custom-made probe holder that is constructed with a system of light springs and ball bearings to allow the probe to move smoothly down with the jaw while the participant speaks. It is mounted in an adjustable custom-made pedestal and on an electrical table to be adjustable for the participant's height. The child sits perpendicular to the probe holder with the small probe positioned below his chin between the maxillary bones. In this study, we did not use a fixed headset to maximize the naturalness of the speech recorded and avoid blocking jaw movements, which would require participants to modify their natural articulatory strategies. As we were interested in the maturation of coarticulatory patterns over age and as laboratory settings may already affect natural speech style, maximizing naturalness was a crucial criterion for us to decide against a fixed headset. A discussion of our setup's advantages and limitations can be found in the Limitation and Perspective section.

Upon arrival at the laboratory, children were familiarized with the experimenters and the SOLLAR platform. They were comfortably seated in a car seat that included seat belts as part of the SOLLAR spaceship. Two experimenters were involved in the recording to maximize the quality of the data collected. While one experimenter monitored the recording equipment and controlled for the quality of the data collection (e.g., quality of tongue images, position of the child via the video camera), the other experimenter maintained a face-to-face connection with the child, controlled for head movement, and executed the stimulus presentation. The two experimenters were well trained with the devices and with recording children. Prior to testing, we organized pilot recordings with adults and children to optimize the experimental procedure, timing, and general approach to be employed with children. Despite our efforts, articulatory data collection with young children remains challenging; hence, we could not guarantee experimental conditions as optimal as when testing adults. As in other developmental studies (e.g., Zharkova, 2017), we therefore conducted qualitative examinations of the video data post recording to select tokens for analyses (cf. Appendix).

The production task consisted of repeating the auditorily presented stimuli. Adults were recorded with the same experimental setup except that we excluded the space journey storyline. All participants were compensated for their participation in the study, and children received a present.

Data Processing

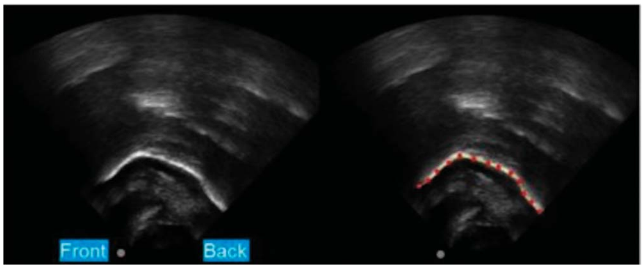
The acoustic speech signal was used as a reference to detect relevant time points in the articulatory signal recorded with ultrasound imaging. For adults, the segmentation was done semiautomatically using WebMAUSBasic (Kisler, Schiel, & Sloetjes, 2012) and subsequent manual adjustments. For children, two to three trained students at the linguistic department of University of Potsdam labeled segments from correct target CV syllables. Manual adjustments and labeling were done within Praat (Boersma & Weenink, 1996). Formant patterns (especially F2 and F3) and changes in periodic cycle were used to determine phoneme onsets and offsets (e.g., for vowel, the first pulse with the visible formant structure was used as reference for onsets and the end of the formant structure for offsets). Boundaries were systematically adjusted with the automatic function "move to nearest zero crossing" provided in Praat to guarantee consistency in boundary settings throughout labeling. Cases for which transcription was problematic were discussed with the labeling team. To measure CD differences, two time points were extracted from the acoustic speech signal: the temporal midpoint of the acoustically defined first consonant (hereafter referred to as *C50*) and the temporal midpoint of the acoustically defined vowel (V50).

Corresponding ultrasound images of the tongue were then extracted on the basis of the synchronized acoustic speech signal. More specifically, SOLLAR script selects the video frame with the time code most closely matching each of the two target time points. As the ultrasound data are recorded in 60 Hz, the interval between each ultrasound frame was 16.6667 ms. This means that, in the worst case, if the requested time falls exactly in between frames, the selected ultrasound frame would be 8.33 ms off with respect to the acoustical landmark. We judged this potential issue minimal (e.g., accuracy of ±40 ms in Zharkova & Hewlett, 2009). For each relevant frame (C50, V50), tongue contours were detected with SOLLAR custom-made scripts for MATLAB (Figure 1). For each tongue contour, an estimate of the tongue body position along the front–back dimension was obtained by extracting the *x* and *y* coordinates of the highest point of the tongue body. This point was taken as reference for vocalic gestures. We discarded /da/ sequences because /a/ being a low vowel, the highest point on the tongue body would occur more in the region of the tongue blade, which would not be representative of the vocalic gesture.

Statistical Analysis

A table providing an overview of the number of CV repetitions across and within age groups used for the statistical analyses can be found in the Appendix.

Figure 1. Example of ultrasound image recorded within Sonographic and Optical Linguo-Labial Articulation Recording platform. Left panel presents the initial tongue image as recorded on the ultrasound scanner; the right panel shows the highlighted tongue contours. In each image, the left portion corresponds to the anterior part of the tongue.



Testing for Overall Developmental Differences in CD

First, we examined whether overall CD differences could be observed across the five cohorts investigated in this study. To achieve this, we fitted linear mixed-effects models using the "lme4" package in R (Bates, Mächler, Bolker, & Walker, 2015). We regressed the horizontal position of the highest point on tongue body at consonant midpoint (PEAKX_C1_050) on the horizontal position of the highest point on tongue body at vowel midpoint (PEAKX_V50), age cohort (COHORT), consonant (CONSONANT1), and the interaction of PEAKX_V50 and age cohort (PEAKX_V50: COHORT). Age cohort and consonant were treatment coded with C3 and /b/ as baselines, respectively. The structure of random effects for this model and for all the following linear mixed models presented in the article was determined following the approach suggested by Bates et al. (Bates, Kliegl, Vasishth, & Baayen, 2015). The approach combines principal components analysis (PCA) to determine the maximal number of dimensions for a model that is supported by the data ("RePsychLing" package; Bates et al., 2015) with likelihood ratio tests to assess goodness of fit. We began by testing the full random-effects structure for subject and word. If the maximal model converged, we used PCA to check whether this number of dimensions was supported by the data. If the PCA showed that the number of dimensions was not supported, we proceeded with dropping the smallest variance components. If the maximal model failed, we dropped variance components until the identification was achieved. As a result, random intercepts and random by-consonant slopes for subjects were included as random effects. The models' assumptions were checked by visual inspection of the residual plots. Outliers were checked individually and either removed (in case of experimental errors) or corrected (in case of processing errors). Removing outliers did not result in any changes in outcome pattern.

The *p* values were corrected to account for multiple comparisons following the truncated closed test procedure from Westfall (1997) as implemented in the *glht* function of "multcomp" package (Hothorn, Bretz, Westfall, & Heiberger, 2008). All pairwise comparisons for the PEAKX_V50: COHORT were obtained by manually setting the contrast matrix.

Testing for Consonant-Specific Effects on CD Across Age Groups

In the second step, to test whether CD for a specific consonant differed between age groups, we fitted a linear mixed model for each consonant with PEAKX_C1_050 as response variable and PEAKX_V50 and COHORT and interaction thereof as fixed effects. PEAKX_C1_050 was power transformed to better approximate normality using "BoxCox" function from R package "forecast" (Hyndman, 2017). Cohort was a five-level factor for all stops' models and a three-level factor for the alveolar fricative model, with treatment coding and C3 as baseline. For all the five models, the random effects explored consisted of the full random effects structure for subject and word. The resulting random-effects structure in all cases included by-subject random intercepts and slopes for PEAKX_V50.

Testing Consonant-Specific Effects on CD Within Age Groups

Third, we fitted linear mixed-effects models to statistically compare the effect of onset consonant identity on CD within each age group. All five models included PEAKX_C1_050 as response variable and PEAKX_V50, CONSONANT1, and their interaction (PEAKX_V50: CONSONANT1) as fixed effects. PEAKX_C1_050 was again power transformed with the BoxCox function from the R package forecast (Hyndman, 2017). For the 3- and 5-year-olds, consonant was a three-level categorical predictor (b, d, g). For 4- and 7-year-olds and adults, the consonant was a four-level predictor (b, d, g, z). We used dummy coding with /b/ as baseline. The structure of random effects for each model was determined following the strategy described above. For each of the five models, we began by testing the full random-effects structure for subject and word. We followed the same approach as above (cf. testing for differences in CD). All pairwise comparisons for the PEAKX_V50: CONSONANT1 were obtained by manually setting the contrasts matrix. As separate models were fitted for each cohort, results are only comparable within cohort.

All statistical analyses were carried out in R (Version 3.4.0, R Core Team, 2017). In the next section, we present the results of each statistical analysis as output of pairwise comparisons with a short descriptor of the comparisons, the effect estimates with associated standard errors, the test statistics (*t* value), the and multiplicity-adjusted *p* values. The ± signs of the estimates determine the direction of the effect. To take an example, the negative estimate comparing CD between 4- and 3-year-olds (noted as 4–3 in Table 1) means that the 4-year-olds show a lower slope as compared with the 3-year-olds.

Results

Developmental Differences in CD

Before addressing the question of consonant-specific effects on CD, we first evaluated whether overall CD differences could be observed across the five cohorts investigated in this study. Table 1 reports the results from the linear mixed-effects model comparing the effect of the horizontal tongue body position at vowel midpoint (V50) on the horizontal tongue body position at consonant midpoint (C50) between different age groups (adjusted *p* values). Results are pooled across consonants and do not include the alveolar fricative /z/ as this consonant was not collected in all age groups.

Table 1. Results from linear mixed-effects models across age groups for the horizontal position of the tongue at C50 with respect to V50.

Age group	Estimate	*SE*	*t* value	*p* value
4–3	−0.0182195	0.0207696	−0.877	.6483
5–3	−0.0187760	0.0211692	−0.887	.6483
G1–3	−0.0832647	0.0202244	−4.117	< .001***
Adult–3	−0.2209096	0.0186712	−11.832	< .001***
5–4	−0.0005565	0.0195833	−0.028	.97733
G1–4	−0.0650452	0.0185589	−3.505	.00139**
Adult–4	−0.2026901	0.0168517	−12.028	< .001***
G1–5	−0.0644888	0.0190023	−3.394	.00139**
Adult–5	−0.2021336	0.0173400	−11.657	< .001***
Adult–G1	−0.1376449	0.0161735	−8.511	< .001***

Note. Age groups include 3 years, 4 years, 5 years, Grade 1 (G1), and adult. C50 = first consonant; V50 = acoustically defined vowel. **$p < .01$. ***$p < .001$.

As can be noted in Table 1, significant differences in CD were found between all children cohorts regardless of age and adults with overall greater coarticulation in children relative to adults ($p < .001$). In addition, CD differed significantly between the oldest group of children at Grade 1 and the three younger groups of children at age 3 ($p < .001$), 4, or 5 years ($p < .01$). However, CD differences between the three groups of younger children did not yield significance.

Consonant-Specific Effects on CD Across Groups

Second, we tested for age-related differences in CD as a function of onset consonant identity. Table 2 reports results from the linear mixed-effects models comparing CD differences across the five age cohorts investigated. Here again, we observed substantial differences between adults and children.

In the case of labial-V coarticulation, all child cohorts differed significantly from adults ($p < .001$). In addition, 3-year-olds differed from Grade 1 children ($p < .05$). Velar-V coarticulatory patterns also showed greater CD in each child cohort compared with adults ($p < .001$ with the 3-, 4-, and 5-year-olds but $p < .05$ with Grade 1 children). Further, for velar-V sequences, both the 3- and 5-year-olds showed greater CD than children in Grade 1 ($p < .001$ and $p < .05$, respectively). For the alveolar stop /d/, CD was significantly greater in children relative to adults ($p < .01$ for the 3-, 4-, and 5-year-olds and $p < .05$ for Grade 1). However, this time, no difference was found between the 3-year-olds and

Table 2. Results from linear mixed-effects models across age groups for the horizontal position of the tongue at C50 with respect to V50.

Cons	Age group	Estimate	*SE*	*t* value	*p* value
b	4–3	−0.0103	0.0080	−1.284	.404
	5–3	−0.0101	0.0083	−1.215	.404
	G1–3	−0.0247	0.0080	−3.082	.011**
	Adult–3	−0.0736	0.0077	−9.616	< .001****
	5–4	0.0002	0.0078	0.026	.979
	G1–4	−0.0144	0.0075	−1.931	.130
	Adult–4	−0.0633	0.0071	−8.952	< .001****
	G1–5	−0.0146	0.0078	−1.881	.130
	Adult–5	−0.0635	0.0074	−8.589	< .001****
	Adult–G1	−0.0489	0.0071	−6.924	< .001****
d	4–3	0.0240	0.1532	0.157	.96175
	5–3	−0.0166	0.1549	−0.107	.96175
	G1–3	−0.1120	0.1509	−0.742	.79641
	Adult–3	−0.4997	0.1488	−3.358	.00433***
	5–4	−0.0406	0.1526	−0.266	.96175
	G1–4	−0.1360	0.1485	−0.916	.79641
	Adult–4	−0.5237	0.1465	−3.576	.00320***
	G1–5	−0.0954	0.1502	−0.635	.79641
	Adult–5	−0.4831	0.1482	−3.261	.00433***
	Adult–G1	−0.3877	0.1439	−2.693	.02631**
g	4–3	−0.2019	0.0897	−2.249	.0632*
	5–3	−0.0802	0.0923	−0.868	.3853
	G1–3	−0.3345	0.0888	−3.769	< .001****
	Adult–3	−0.5393	0.0864	−6.245	< .001****
	5–4	0.1217	0.0900	1.352	.2336
	G1–4	−0.1326	0.0863	−1.536	.2336
	Adult–4	−0.3374	0.0839	−4.023	< .001****
	G1–5	−0.2543	0.0890	−2.857	.0121**
	Adult-5	−0.4591	0.0866	−5.300	< .001****
	Adult–G1	−0.2048	0.0828	−2.473	.0486**
z	G1–4	−1.6770	2.4080	−0.696	.486
	Adult–4	−2.9100	2.3470	−1.240	.430
	Adult–G1	−1.2340	2.3350	−0.528	.597

Note. Results (adjusted *p* values) are presented per consonant. Age groups include 3 years, 4 years, 5 years, Grade 1 (G1), and adult. C50 = first consonant; V50 = acoustically defined vowel; Cons = consonant.

$^{*}p < .1$. $^{**}p < .05$. $^{***}p < .01$. $^{****}p < .001$.

children in Grade 1. Finally, as regards the alveolar fricative /z/, no difference in CD between the 4-year-olds, the Grade 1 children, and the adults was found.

Consonant-Specific Effects on CD Within Groups

Table 3 presents results from linear mixed-effects models assessing the effect of onset consonant identity on CD pattern within each age group.

Results show a similar decreasing degree of coarticulation for all cohorts from labial to velar, followed by the alveolar stop consonant. With respect to the alveolar fricative and stop, the 4-years-olds and the Grade 1 children showed a similar direction of effects with greater CD in the fricative compared with the stop context. However, this difference was not significant in the 4-year-old children contrary to older children at the end of Grade 1 ($p < .05$). Finally, neither the 4-years-olds nor the Grade 1 children showed a difference between the velar /g/ and fricative /z/, whereas adults did.

Table 3. Results from linear mixed-effects models within age group for the horizontal position of the tongue at C50 with respect to V50.

Age group	Cons	Estimate	*SE*	*t* value	*p* value
3 years	d-b	−0.0274	0.0029	−9.498	< .001***
	g-b	−0.0149	0.0026	−5.786	< .001***
	g-d	0.0125	0.0030	4.201	< .001***
4 years	d-b	−0.1310	0.0152	−8.633	< .0001***
	g-b	−0.1197	0.0136	−8.779	< .0001***
	z-b	−0.1224	0.0156	−7.842	< .001***
	g-d	0.0112	0.0166	0.678	.776
	z-d	0.0085	0.0177	0.481	.776
	z-g	−0.0027	0.0167	−0.163	.870
5 years	d-b	−0.2168	0.0208	−10.422	< .0001***
	g-b	−0.1384	0.0203	−6.810	< .0001***
	g-d	0.0784	0.0216	3.639	.000273***
Grade 1	d-b	−8.6322	0.7235	−11.931	< .001***
	g-b	−6.7567	0.6846	−9.870	< .001***
	z-b	−6.4479	0.7885	−8.177	< .001***
	g-d	1.8755	0.7643	2.454	.0272*
	z-d	2.1843	0.8489	2.573	.0272*
	z-g	0.3088	0.8171	0.378	.7055
Adults	d-b	−1.3134	0.1153	−11.390	< .001***
	g-b	−0.6636	0.1086	−6.113	< .001***
	z-b	−0.9059	0.1127	−8.041	< .001***
	g-d	0.6497	0.1249	5.202	< .001***
	z-d	0.4074	0.1299	3.135	.00172**
	z-g	−0.2423	0.1227	−1.974	.04842*

Note. Results (adjusted *p* values) are presented per consonant. C50 = first consonant; V50 = acoustically defined vowel; Cons = consonants.

$^{*}p < .05$. $^{**}p < .01$. $^{***}p < .001$.

Discussion

This study investigated the maturation of intrasyllabic coarticulatory organization in TD German children from 3 to 7 years of age. Investigation of coarticulation in this period is not only relevant for understanding the maturation of spoken language fluency or speech motor control but it also provides normative data necessary for disentangling atypical trajectories from the typical variability observed in the young age (e.g., in CAS: Nijland et al., 2002, 2003; children who stutter: Soo-Eun et al., 2002; speech sound disorders: Cleland, Scobbie, & Wrench, 2015; phonological disorders: Gibbon, 1999; hearing impairment: Bernhardt, Gick, Bacsfalvi, & Adler-Bock, 2005). Indeed, if we understand well how coarticulatory mechanisms mature in TD children, we should be better equipped to detect deviations from typical trajectories early in the child's language development.

In light of previous research, we were specifically interested in the following questions. Does intrasyllabic lingual coarticulation in German children differ between children and adults? If so, is this difference observed across consonants? Finally, does CD vary on the basis of the onset consonants' identity within children, as it does in German adults (Abakarova et al., 2017; Iskarous et al., 2013)? In the next two sections, results are discussed with respect to the two overarching themes they address, namely, the general ontogeny of coarticulation and its relations to speech motor control development and its implication for ongoing discussions about the units of speech organization. Finally, we discuss the limitations and perspectives of the study.

Developmental Differences in Coarticulation and Speech Motor Control Development

The challenging question of whether children organize their speech in segments versus syllables versus phonological words or lexical items is twofold: It requires finding the phonological units guiding children's speech production and the motor units embedding those higher level units. From a motor perspective, a main goal for children is to learn to discretize continuous articulatory movements into distinct articulatory gestures conveying the specific phonetic properties of their native language. In that perspective, coarticulation is an interesting mechanism to investigate as it offers a window into understanding how speech motor control develops with age and experience with the native language. In this study, we could not assess the direct role of individual experience; however, by employing a cross-sectional design with four tightly grouped cohorts of German children spreading the kindergarten period (at an average of 3, 4, and 5 years of age) and the beginning of primary school (at 7 years of age), we could estimate the apparent changes in CD patterns across childhood.

First, as predicted, results indicate that lingual patterns in all children groups differ from adults with greater CD in children relative to adults. In addition, we found that the three youngest groups of children in kindergarten differed from older children at the end of Grade 1. This suggests that the period between the end of kindergarten and the end of the first year into primary school corresponds to an important transition with respect to children's phonological and speech motor control development, that is, children have decreased the spatiotemporal overlap for CV articulatory gestures. However, by the end of the first school year, German children's CD patterns are still greater than adults'. The maturation of lingual coarticulation is still ongoing.

Second, we found age-related differences in CD across the three consonantal contexts (/b/, /d/, and /g/) in the direction of greater CD in children than in adults. This suggests that vocalic gestures invade the temporal domain of the syllable more than in adults. Interestingly, in labial context, 3-year-olds differed from children at Grade 1. Given that jaw patterns are controlled the earliest in age (e.g., review in Green et al., 2002), that labials are dominant in the babbling repertoire (e.g., in German: Fox-Boyer, 2006, 2009), one may expect CD in syllables involving labials to approximate adults' CD patterns early in age. Yet, our results do not support this hypothesis. In fact, they show that, at the age of 7 years, children have certainly more mature lingual CD patterns than at age 3 years, but they still differ significantly from adults as to the phasing between two articulatory gestures: a well-practiced motor routine involving the jaw (with support of the lips) for the labial consonant and another gesture involving the tongue for the achievement of the vowel. This result somehow contradicts previous reports that, by age 6 years, children's coupling of the jaw and lips approximates those of adults (e.g., Green et al., 2002). This may be due to the fact that, in their study, the authors only examined spontaneous productions of /baba/ sequences that are prototypical in the babbling and first words of children while we examined /b/-V syllables involving various vocalic contexts in a repetition task. In our study, age-related differences were also found for the velar stop /g/ with greater vowel-related influence upon CD in all groups of children compared with adults and in the 3- and 5-year-olds compared with older children at age 7 years. Hence, in homorganic syllables involving a single organ for the production of both consonantal and vocalic gestures, such as in /g/-V, results suggest that at age 7 years, children are on their way to achieve adultlike CD patterns, but they are not quite yet like adults. For the alveolar stop /d/, age-related differences between the groups of children disappear; the only difference in CD pattern was noted between children and adults. Finally, results for the fricative alveolar /z/ did not reveal any significant difference within the groups of children or as compared with adults. This result does not corroborate previous reports (e.g., in German based on measures of spectral center of gravity in 4- to 6-year-old children; Kleber, 2015). Discrepancy in results may be due to methodological differences between acoustical and articulatory studies along with the age and particular choice of stimuli investigated.

However, our final observation that children displayed consonantal effects on CD patterns within age groups provides further details on children's speech motor control and, most specifically, on the absence of age-related difference for the alveolar fricative. We indeed found greater CD in the context of the labial stop /b/, followed by the velar /g/, then the alveolar stop /d/. As briefly outlined in the introduction, a main explanation for these differences in CD stems from consonants' degree of flexibility versus resistance to overlap with neighboring vowels. While some consonants allow sizeable coarticulation with neighbors (e.g., labials), others make strong articulatory demands on the tongue to preserve intelligibility (e.g., alveolars) and, therefore, resist blending with adjacent vowels (e.g., Fowler, 1994; Fowler & Brancazio, 2000; Recasens, 1985; Recasens & Espinosa, 2009). Previous work has shown that, during the alveolar stop /d/, the tongue body and tongue tip function as a collaborative network (or functional synergy) so that the base of the tongue moves the front to support the anterior part of the tongue into achieving the alveolar occlusion (e.g., Iskarous et al., 2010, 2011). Results from this study seem to support the hypothesis that children as young as 3 years of age display synergistic relationships among the functional subparts of the tongue to achieve the alveolar constriction (e.g., for /d/). However, the significant differences in CD noted between children and adults indicate that these synergies are not yet mature by the end of Grade 1. Such discrepancy is also exemplified by the finding that both 4-year-olds and children at the end of Grade 1 did not show any difference in CD between the velar /g/ and alveolar fricative /z/, whereas adults did. The horizontal position of the tongue body in syllables including /z/ was affected by the subsequent vocalic gesture as much as in velar context. This suggests that the articulation of alveolar fricatives remains challenging throughout childhood as exemplified in other studies (e.g., in German: Fox-Boyer, 2009; in English: review in Maas & Mailend, 2017). Beyond the positioning of the tongue tip and tongue blade in the alveolar region, fricatives require fine glottal control and, therefore, precise coordination between the supralaryngeal and laryngeal areas, which are not yet mastered at age 10 years (e.g., Koenig, Lucero, & Perlman, 2008). The observations made for the two alveolars in this study are interesting but evidently preliminary and call for more detailed examinations of lingual synergies for individual articulatory gestures. In future studies, other developmental factors, such as anatomical growth of the vocal tract, should also be carefully considered.

Taken together, our findings indicate that, by the end of their first year in primary school, children have substantially changed the organization of their lingual coarticulatory patterns in comparison to the kindergarten period, but they do not yet approximate adults' patterns. This suggests that the process of learning to coordinate the spatiotemporal properties of articulatory gestures to produce adultlike patterns of intrasyllabic CD is protracted and not uniform across childhood (e.g., Smith, 2010) as found for other motor behaviors (e.g., Thelen & Smith, 1994). This study provided a first account in German. More research is obviously needed to assess the age at which children transition toward adultlike patterns of lingual coarticulation and to fully understand the complexities underlying the maturation of intraorgan and interorgan coordination.

Units of Intrasyllabic Organization: Segments? Syllables? Articulatory Gestures?

An overarching motivation to conduct this research stemmed from the conflicting reports found in the literature as to whether intrasyllabic coarticulation is organized in children as a single chunk with a large vocalic influence observed throughout the syllable (e.g., Goodell, & Studdert-Kennedy, 1993; Nijland et al., 2002; Nittrouer & Whalen, 1989; Rubertus et al., 2015) or in a more sequential manner with minimal overlap between both segments (Gibson & Ohde, 2007; Green et al., 2002; Kent, 1983; Zharkova, Hewlett, & Hardcastle, 2012). Examinations of consonant-related effects on CD across and within age groups provide new insight on this question.

In our adult reference group, the domain of coarticulatory organization varied as a function of the onset consonant identity, which suggests that intrasyllabic organization is not uniform across CV syllables but, instead, depends on the articulatory demands associated with consecutive phonemes. In German children, results overall corroborate previous research arguing for a broader-size organization in which the vowel exerts more influence on the previous consonant relative to adults. This general finding supports a holistic perspective by which children start with a word-based or syllable-based unit of speech organization and progressively reduce their domain of coarticulatory organization as they learn to abstract phonemes from the words and syllables in which they are embedded and their associated articulatory gestures (e.g., Nittrouer et al., 1996). However, the fact that we observe consistent consonantal effects on CD within age groups is a strong indication that, at the age of 3 years, children have departed from a purely holistic organization to integrate articulatory gestures associated with individual segments.

Results suggest that children do not coarticulate consonants and vowels with an unvarying degree of overlap irrespective of the segments combined (e.g., Gibson & Ohde, 2007; Katz & Bharadwaj, 2001; Noiray et al., 2013; Zharkova et al., 2015). In fact, results show that the degree with which lingual vocalic gestures invade the temporal domain of previous consonants depends on whether the production of both the consonant and vowel involves the tongue. Taken together, previous research and present findings suggest that a crucial step in becoming mature speakers may not be for children to globally increase or decrease coarticulation but to achieve flexible patterns of coarticulation depending on the combination of segments. This hypothesis is in line with the theory of articulatory phonology (Browman & Goldstein, 1992), which argues that articulatory (or constriction) goals represent dynamic units of action guiding the activity of speech articulators. From a developmental standpoint, articulatory phonology provides an interesting framework to explain differences in intrasyllabic coarticulatory organization, which departs from traditional phonological descriptions. In the first years of language acquisition, children have minimal structural knowledge about their native language, that is, they have limited awareness that words can be decomposed into segments and recombined into new meaningful forms (e.g., Gillon, 2007; Liberman, Shankweiler, Fischer, & Carter, 1974). Yet, they are able to concatenate sounds into an intelligible flow to communicate with the world. Hence, one reason for the unresolved controversy over children's units of speech production may stem from the premise that children's productions can be described in terms of adults' categories (e.g., segments, syllables). However, up to date, there has not been any strong evidence showing that children organize their speech in categories similar to those used to describe adults' productions. Instead, children may organize their speech in very different units—that are neither segments nor syllables—that researchers fail to capture because of the inadequacy of the methodologies employed focusing on finding adult-based units or because of the limitations in the linguistic material tested (often one or two contrasts). Results from adult studies suggest that even adults do not organize their speech along traditional phonological lines but, rather, in terms of articulatory goals that are dynamically organized to integrate contextual variability (e.g., Browman & Goldstein, 1992; Iskarous et al., 2011; Recasens & Espinosa, 2009).

Results from this cross-sectional investigation support the hypothesis of an articulatory basis for speech organization in preschool and school-aged children. They further support the view that CD varies within a continuum with gradient distinctions resulting from the coproduction of articulatory gestures. When a consonant imposes strong demands on an articulator that is shared with its neighboring vowel, coarticulatory resistance leads toward a more phone-sized organization. When no such constraint exists between consecutive phonemes, speakers may employ a broader domain of articulation. More in-depth investigations are evidently needed to uncover the impact of articulatory goals in spoken language organization. Fortuitously, in the last decade, optimization of ultrasound imaging for child studies has allowed for a direct and user-friendly access to children's articulation. More recently, a platform combining a talking head with ultrasound images has been developed (e.g., Fabre, Hueber, Girin, Alameda-Pineda, & Badin, 2017), which provides new opportunities for advancing our understanding of early language organization in typical and atypical populations beyond acoustic analyses or standard assessment of articulatory capabilities. In a new project, we have also started examining the maturation of coarticulatory organization in relation to speech motor control development in greater details and the impact of phonemic awareness on coarticulatory (re)organization in late childhood.

Limitations of the Study and Perspectives

In this study, we were interested in consonant-related effects on coarticulatory organization in children and in the phenomenon of consonants resistance across a range of vowels. We therefore conducted analyses across vocalic contexts. Other studies have investigated the specific role of individual vowels upon lingual organization (e.g., in Catalan adults: Recasens & Rodríguez, 2016). Future studies should expand on such examinations to disentangle consonant-related and vowel-related effects on children's coarticulatory patterns as compared with adults.

Another important aspect to discuss regards the experimental approach endorsed in this study. Controlling for head movement has been a longstanding challenge for psycholinguists aiming to collect data from children's speech articulators. Each approach (fixed headset, handheld, microphone stand, probe holder) presents certain advantages and drawbacks we briefly outline below. Using a fixed headset prevents the head from moving with respect to the jaw and, therefore, makes comparisons across tongue curves more reliable than when the ultrasound probe is free to move. However, this method has also issues, which were outlined in the method section. Importantly, fixed headsets block the jaw and, therefore, require children to employ articulatory strategies that may not reflect their actual speech. Given previous research showing the importance of the jaw in early childhood (e.g., Green et al., 2002; Grigos, 2009; Smith & Zelaznik, 2004) and its use in adults for vowel contrast (e.g., Noiray et al., 2014), it appeared fundamental for us to not prevent speakers from moving their jaw freely, especially in the young age. Some studies have used microphone stands to attach ultrasound probes (e.g., Noiray et al., 2013). Although this device does not block jaw motion, it encourages the participant to move his head up to compromise for the impossibility to move the jaw down. Finally, other studies have used a handheld probe approach (e.g., Zharkova et al., 2015) and reported on noticeable differences with the headset approach. Such difference may result from the hand sliding in the lateral and horizontal dimensions and from inconsistent contact with the chin floor affecting the probe's vertical dimension and, more generally, the image quality (e.g., shorter tongue curve, blurry tongue curve line).

Our customized probe holder also presents advantages and limitations. It was designed to maximize naturalness of speech while controlling for probe motion (see Method section for details). Although this setup certainly succeeds at not impeding natural motor movements from the face, it may affect the reliability of tongue curvature comparisons because of a physically induced misalignment of tongue shape rather than linguistically induced distinctions. We used various strategies to address this limitation (e.g., seatbelts, two experimenters, an experimenter and a star as a visual fixation point in front of the child) and conducted qualitative examinations of the tongue contour using video data postrecording to discard data for which children moved as in other studies (e.g., Zharkova, 2017). Of course, qualitative examinations remain somewhat subjective, but we remain confident that, with the quantitative dataset we collected, our results bring insightful knowledge to the field.

For our analyses, we used the highest point on the tongue body as used in previous research (e.g., Noiray et al., 2013). This approach allowed us to depart from the end curves of the tongue that can be inaccurate regardless of the recording (e.g., headset, microphone stand, probe holder) or analysis approach (e.g., automatic tongue contour detection, manual detection). Other approaches exist, such as examining the whole tongue shape to account for coarticulatory differences (e.g., Zharkova et al., 2015). The reliability of this method may be affected by the quality of the tongue imaging at the two ends of the tongue curve across tokens and speakers, but it provides insightful results as to the tongue bunching. Hence, both approaches provide complementary information on the maturation of articulatory organization in the young age.

Various parameters may affect the quality of tongue imaging irrespective of the method employed. The experimenter must decide on the method that best suits the population and theoretical questions addressed. Considering the limitations outlined above, future studies should prioritize large-scale descriptions across childhood employing the same experimental and analytical approaches for all subjects. Including comparative data with adults is also essential to have a reference regarding the maturity of the production patterns investigated. Most quantitative analyses of lingual coarticulation in children (including ours) do not provide absolute values easily comparable to the adult literature but rather compare numerical values for one age cohort relative to the others within-study.

Conclusion

This study addressed the maturation of lingual coarticulatory patterns in TD German children in comparison to adults. Results show that, already at age 3 years, differences in the degree of intrasyllabic coarticulation correlate with differences in phonetic properties of the onset consonant. Overall, results provide evidence that German children's speech is not uniformly organized along abstract phonological lines, in either segments or syllables as often argued in the literature (at least for English learning children). Instead, coarticulatory organization is sensitive to the articulatory properties associated with individual segments and their compatibility once combined into a continuous speech flow. This result replicates previous findings in adults and older children and expands it to German children from age 3 to 7 years. This is an important result because it not only concurs to clarify typical maturation of coarticulatory mechanism but it could also potentially supplement assessment of atypical production patterns. Future research should further test the role of lingual gestures for coarticulatory organization in TD children and examine whether articulatory demands observed for consecutive articulatory gestures may impede acquisition of spoken language fluency.

Acknowledgments

This research was supported by the Deutsche Forschungsgemeinschaft (1098 and 255676067, recipient: Dr. Aude Noiray). The authors are grateful to Carol Fowler for stimulating discussions at various stages of this research and to Martijn Wieling for his statistical insights. The authors are grateful to Jan Ries (University of Potsdam) for his assistance in developing MATLAB scripts, to the Baby-Lab at University of Potsdam (in particular to Barbara Hoehle and Tom Fritzsch) for helping with participants' recruitment, and the team at the Laboratory for Oral Language Acquisition involved in data recording and processing. Importantly, the authors thank all the participants, adults, and children without whom this research would not have been possible. Finally, we are grateful to the three reviewers for their helpful comments and suggestions.

References

Abakarova, D., Iskarous, K., & Noiray, A. (2017, July). *Quantification of coarticulation resistance in German with ultrasound.* Paper presented at the Speech Motor Control conference, Groningen, the Netherlands.

American Speech-Language-Hearing Association. (2007). *Childhood apraxia of speech* [Technical report]. American Speech-Language-Hearing Association. Retrieved from http://www.asha.org/policy

Barbier, G., Perrier, P., Ménard, L., Payan, Y., Tiede, M., & Perkell, J. (2015, September). *Speech planning in 4-year-old children versus adults: Acoustic and articulatory analyses.* Paper presented at the 16th Annual Conference of the International Speech Communication Association (Interspeech 2015), Dresden, Germany.

Bates, D., Kliegl, R., Vasishth, S., & Baayen, H. (2015). *Parsimonious mixed models.* Manuscript submitted for publication.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Bernhardt, B., Gick, B., Bacsfalvi, P., & Adler-Bock, M. (2005). Ultrasound in speech therapy with adolescents and adults. *Clinical Linguistics & Phonetics, 19*(6–7), 605–617.

Bladon, R. A., & Al-Bamerni, A. (1976). Coarticulation resistance in English /l/. *Journal of Phonetics, 4,* 135–150.

Boersma, P., & Weenink, D. (1996). Praat, a system for doing phonetics by computer, version 3.4. Retrieved from http://www.praat.org

Browman, C. P., & Goldstein, L. (1992). Articulatory phonology: An overview. *Phonetica, 49*(3–4), 155–180.

Cheng, H. Y., Murdoch, B. E., Goozée, J. V., & Scott, D. (2007). Physiologic development of tongue–jaw coordination from childhood to adulthood. *Journal of Speech, Language, and Hearing Research, 50*(2), 352–360.

Cleland, J., Scobbie, J. M., & Wrench, A. A. (2015). Using ultrasound visual biofeedback to treat persistent primary speech sound disorders. *Clinical Linguistics & Phonetics, 29*(8–10), 575–597.

Fabre, D., Hueber, T., Girin, L., Alameda-Pineda, X., & Badin, P. (2017). Automatic animation of an articulatory tongue model from ultrasound images of the vocal tract. *Speech Communication, 93,* 63–75.

Fowler, C. A. (1994). Invariants, specifiers, cues: An investigation of locus equations as information for place of articulation. *Attention, Perception, & Psychophysics, 55*(6), 597–610.

Fowler, C. A., & Brancazio, L. (2000). Coarticulation resistance of American English consonants and its effects on transconsonantal vowel-to-vowel coarticulation. *Language and Speech, 43*(1), 1–41.

Fowler, C. A., & Saltzman, E. (1993). Coordination and coarticulation in speech production. *Language and Speech, 36*(2–3), 171–195.

Fox-Boyer, A. (2006). Evidence from German-speaking children. In Z. Hua, & B. Dodd. (Eds.), *Phonological development and disorders in children: A multilingual perspective* (pp. 56–80). Clevedon, United Kingdom: Multilingual Matters.

Fox-Boyer, A. (2009). *Kindliche Aussprachestörungen: phonologischer Erwerb, Differenzialdiagnostik, Therapie* [Diagnostic tool for assessing phonological development in German children]. Idstein, Germany: Schulz-Kirchner.

Gibbon, F. E. (1999). Undifferentiated lingual gestures in children with articulation/phonological disorders. *Journal of Speech, Language, and Hearing Research, 42*(2), 382–397.

Gibson, T., & Ohde, R. N. (2007). F2 locus equations: Phonetic descriptors of coarticulation in 17- to 22-month-old children. *Journal of Speech, Language, and Hearing Research, 50*(1), 97–108.

Gillon, G. T. (2007). *Phonological awareness: From research to practice.* New York, NY: Guilford.

Goodell, E. W., & Studdert-Kennedy, M. (1993). Acoustic evidence for the development of gestural coordination in the speech of 2-year-olds: A longitudinal study. *Journal of Speech and Hearing Research, 36*(4), 707–727.

Graetzer, S. (2006). Consonantal coarticulation resistance in vowel-consonant-vowel sequences in two Australian languages. In *Proceedings of the 11th Australasian International Conference on Speech Science and Technology, 1*(1), 270–275.

Green, J. R., Moore, C. A., & Reilly, K. J. (2002). The sequential development of jaw and lip control for speech. *Journal of Speech, Language, and Hearing Research, 45*(1), 66–79.

Grigos, M. I. (2009). Changes in articulator movement variability during phonemic development: A longitudinal study. *Journal of Speech, Language, and Hearing Research, 52*(1), 164–177.

Hardcastle, W. J., & Hewlett, N. (2006). *Coarticulation: Theory, data and techniques*. Cambridge, United Kingdom: Cambridge University Press.

Hothorn, T., Bretz, F., Westfall, P., & Heiberger, R. M. (2008). Simultaneous inference in general parametric models. *Biometric Journal, 50*(3), 346–363.

Hyndman, R. J. (2017). *Forecast: Forecasting functions for time series and linear models. R package version 8.1*. Retrieved from http://github.com/robjhyndman/forecast

Iskarous, K., Fowler, C. A., & Whalen, D. H. (2010). Locus equations are an acoustic expression of articulator synergy. *The Journal of the Acoustical Society of America, 128*(4), 2021–2032.

Iskarous, K., Mooshammer, C., Hoole, P., Recasens, D., Shadle, C. H., Saltzman, E., & Whalen, D. H. (2013). The coarticulation/invariance scale: Mutual information as a measure of coarticulation resistance, motor synergy, and articulatory invariance. *The Journal of the Acoustical Society of America, 134*(2), 1271–1282.

Iskarous, K., Shadle, C. H., & Proctor, M. I. (2011). Articulatory–acoustic kinematics: The production of American English/s. *The Journal of the Acoustical Society of America, 129*(2), 944–954.

Katz, W. F., & Bharadwaj, S. (2001). Coarticulation in fricative-vowel syllables produced by children and adults: A preliminary report. *Clinical Linguistics & Phonetics, 15*(1–2), 139–143.

Katz, W. F., Kripke, C., & Tallal, P. (1991). Anticipatory coarticulation in the speech of adults and young children: Acoustic, perceptual, and video data. *Journal of Speech and Hearing Research, 34*(6), 1222–1232.

Kent, R. D. (1983). The segmental organization of speech. In *The production of speech* (pp. 57–89). New York, NY: Springer.

Kisler, T., Schiel, F., & Sloetjes, H. (2012, July). *Signal processing via web services: The use case WebMAUS*. Talk presented at the Digital Humanities Conference 2012, Hamburg, Germany.

Kleber, F. (2015, June). *Direction of coarticulation in vowel-fricative sequences in L1-German children*. Poster presented at the International Child Phonology Conference 2015, St. John's, Canada.

Klein, H. B., Byun, T. M., Davidson, L., & Grigos, M. I. (2013). A multidimensional investigation of children's /r/ productions: Perceptual, ultrasound, and acoustic measures. *American Journal of Speech-Language Pathology, 22*(3), 540–553.

Koenig, L. L., Lucero, J. C., & Perlman, E. (2008). Speech production variability in fricatives of children and adults: Results of functional data analysis. *The Journal of the Acoustical Society of America, 124*(5), 3158–3170.

Liberman, I. Y., Shankweiler, D., Fischer, F. W., & Carter, B. (1974). Explicit syllable and phoneme segmentation in the young child. *Journal of Experimental Child Psychology, 18*(2), 201–212.

Lindblom, B., & Sussman, H. M. (2012). Dissecting coarticulation: How locus equations happen. *Journal of Phonetics, 40*(1), 1–19.

Maas, E., & Mailend, M. L. (2017). Fricative contrast and coarticulation in children with and without speech sound disorders. *American Journal of Speech-Language Pathology, 26*(2S), 649–663.

Ménard, L., & Noiray, A. (2011). The development of lingual gestures in speech: Comparing synthesized vocal tracts with natural vowels. *Faits de langue, 37,* 189–202.

Munson, B. (2004). Variability in /s/ production in children and adults: Evidence from dynamic measures of spectral mean. *Journal of Speech, Language, and Hearing Research, 47*(1), 58–69.

Nijland, L., Maassen, B., & van der Meulen, S. (2003). Evidence of motor programming deficits in children diagnosed with DAS. *Journal of Speech, Language, and Hearing Research, 46*(2), 437–450.

Nijland, L., Maassen, B., van der Meulen, S. V. D., Gabreëls, F., Kraaimaat, F. W., & Schreuder, R. (2002). Coarticulation patterns in children with developmental apraxia of speech. *Clinical Linguistics & Phonetics, 16*(6), 461–483.

Nittrouer, S. (1993). The emergence of mature gestural patterns is not uniform: Evidence from an acoustic study. *Journal of Speech and Hearing Research, 36*(5), 959–972.

Nittrouer, S. (1995). Children learn separate aspects of speech production at different rates: Evidence from spectral moments. *The Journal of the Acoustical Society of America, 97*(1), 520–530.

Nittrouer, S., Studdert-Kennedy, M., & McGowan, R. S. (1989). The emergence of phonetic segments: Evidence from the spectral structure of fricative-vowel syllables spoken by children and adults. *Journal of Speech and Hearing Research, 32*(1), 120–132.

Nittrouer, S., Studdert-Kennedy, M., & Neely, S. T. (1996). How children learn to organize their speech gestures: Further evidence from fricative-vowel syllables. *Journal of Speech and Hearing Research, 39*(2), 379–389.

Nittrouer, S., & Whalen, D. H. (1989). The perceptual effects of child–adult differences in fricative-vowel coarticulation. *The Journal of the Acoustical Society of America, 86*(4), 1266–1276.

Noiray, A., Cathiard, M. A., Abry, C., Ménard, L., & Savariaux, C. (2008). Emergence of a vowel gesture control attunement of the anticipatory rounding temporal pattern in French children. In S. Kern, F. Gayraud, & E. Marsico (Eds.), *Emergence of linguistic abilities* (pp. 100–116). Newcastle, United Kingdom: Cambridge Scholars Publishing.

Noiray, A., Cathiard, M. A., Ménard, L., & Abry, C. (2011). Test of the movement expansion model: Anticipatory vowel lip protrusion and constriction in French and English speakers. *Journal of the Acoustical Society of America, 129*(1), 340–349.

Noiray, A., Iskarous, K., & Whalen, D. H. (2014). Variability in English vowels is comparable in articulation and acoustics. *Laboratory Phonology, 5*(2), 271–288.

Noiray, A., Ménard, L., & Iskarous, K. (2013). The development of motor synergies in children: Ultrasound and acoustic measurements. *The Journal of the Acoustical Society of America, 133*(1), 444–452.

Noiray, A., Ries, J., & Tiede, M. (2015). Sonographic & Optical Linguo-Labial Articulation Recording system (SOLLAR). *Oral presentation at Ultrafest VII,* Hong Kong.

Prather, E., Hendrick, D., & Kern, C. (1975). Articulation development in children aged two to four years. *Journal of Speech and Hearing Disorders, 40,* 179–191.

R Core Team. (2017). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

Recasens, D. (1985). Coarticulatory patterns and degrees of coarticulatory resistance in Catalan CV sequences. *Language and Speech, 28*(2), 97–114.

Recasens, D. (1999). Lingual coarticulation. In W. J. Hardcastle & N. Hewlett (Eds.), *Coarticulation: Theory, data and techniques* (pp. 80–100). Cambridge, United Kingdom: Cambridge University Press.

Recasens, D., & Espinosa, A. (2006). Articulatory, positional and contextual characteristics of palatal consonants: Evidence from Majorcan Catalan. *Journal of Phonetics, 34*(3), 295–318.

Recasens, D., & Espinosa, A. (2009). An articulatory investigation of lingual coarticulatory resistance and aggressiveness for consonants and vowels in Catalan. *The Journal of the Acoustical Society of America, 125*(4), 2288–2298.

Recasens, D., & Rodríguez, C. (2016). An investigation of lingual coarticulation resistance using ultrasound. *Journal of Phonetics, 59,* 58–75.

Reidy, P. F. (2015). *The spectral dynamics of voiceless sibilant fricatives in English and Japanese* (Doctoral dissertation). Retrieved from https://etd.ohiolink.edu/

Repp, B. H. (1986). Some observations on the development of anticipatory coarticulation. *The Journal of the Acoustical Society of America, 79*(5), 1616–1619.

Rubertus, E., Abakarova, D., Tiede, M., & Noiray, A. (2015, December). *Development of coarticulation in German children: Acoustic and articulatory locus equations*. Poster presented at Ultrafest VII, Hong Kong.

Sereno, J. A., Baum, S. R., Marean, G. C., & Lieberman, P. (1987). Acoustic analyses and perceptual data on anticipatory labial coarticulation in adults and children. *The Journal of the Acoustical Society of America, 81*(2), 512–519.

Smith, A. (2010). Development of neural control of orofacial movements for speech. In W. J. Hardcastle, J. Laver, & F. E. Gibbon (Eds.), *The handbook of phonetic sciences* (2nd ed., pp. 251–296). Chichester, United Kingdom: Wiley-Blackwell.

Smith, A., & Zelaznik, H. N. (2004). Development of functional synergies for speech motor coordination in childhood and adolescence. *Developmental Psychobiology, 45*(1), 22–33.

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. *Journal of Phonetics, 41*(3), 281–295.

Soo-Eun, C., Ohde, R. N., & Conture, E. G. (2002). Coarticulation and formant transition rate in young children who stutter. *Journal of Speech, Language, and Hearing Research, 45*(4), 676–688.

Stoel-Gammon, C., & Dunn, C. (1985). *Normal and disordered phonology in children*. Austin, Texas: Pro-Ed.

Sussman, H. M., Duder, C., Dalston, E., & Cacciatore, A. (1999). An acoustic analysis of the development of CV coarticulation. A case study. *Journal of Speech, Language, and Hearing Research, 42*(5), 1080–1096.

Sussman, H. M., Hoemeke, K. A., & Ahmed, F. S. (1993). A cross-linguistic investigation of locus equations as a phonetic descriptor for place of articulation. *The Journal of the Acoustical Society of America, 94*(3), 1256–1268.

Terband, H. (2017, July). *Deviant coarticulation in childhood apraxia of speech (CAS) does not include hyperartion*. Paper presented at the Speech Motor Control conference, Groningen, the Netherlands.

Terband, H., Maassen, B., Van Lieshout, P. H. H. M., & Nijland, L. (2011). Stability and composition of functional synergies for speech movements in children with developmental speech disorders. *Journal of Communication Disorders, 44*(1), 59–74.

Thelen, E., & Smith, L. B. (1994). *A dynamic systems approach to the development of perception and action*. Cambridge, MA: MIT Press.

Westfall, P. H. (1997). Multiple testing of general contrasts using logical constraints and correlations. *Journal of the American Statistical Association, 92*(437), 299–306.

Zharkova, N. (2017). Voiceless alveolar stop coarticulation in typically developing 5-year-olds and 13-year-olds. *Clinical Linguistics & Phonetics, 31,* 503–513.

Zharkova, N., Hardcastle, W. J., Gibbon, F., & Lickley, R. (2015). Development of lingual motor control in children and adolescents. *Proceedings of the 18th ICPhS, Glasgow, 2015,* 1–5.

Zharkova, N., & Hewlett, N. (2009). Measuring lingual coarticulation from midsagittal tongue contours: Description and example calculations using English /t/ and /ɑ/. *Journal of Phonetics, 37*(2), 248–256.

Zharkova, N., Hewlett, N., & Hardcastle, W. J. (2012). An ultrasound study of lingual coarticulation in /sV/ syllables produced by adults and typically developing children. *Journal of the International Phonetic Association, 42*(2), 193–208.

Zharkova, N., Hewlett, N., Hardcastle, W. J., & Lickley, R. J. (2014). Spatial and temporal lingual coarticulation and motor control in preadolescents. *Journal of Speech, Language, and Hearing Research, 57*(2), 374–388.

Appendix

Number of Data Points per Syllable

Syllable	C3	C4	C5	G1	A
ba	84	87	77	91	111
be	80	97	81	88	109
bi	81	96	81	79	109
bo	89	82	87	93	110
bu	79	98	77	98	107
by	71	89	88	92	110
Total /b/ syllables	**484**	**549**	**491**	**541**	**656**
de	83	94	83	93	108
di	70	92	80	93	110
do	84	85	87	93	107
du	72	101	83	95	111
dy	76	89	89	96	108
Total /d/ syllables	**385**	**461**	**422**	**470**	**544**
ga	86	92	74	89	109
ge	82	96	78	97	111
gi	73	91	82	102	107
go	83	85	89	93	105
gu	60	93	75	87	109
gy	73	86	82	98	110
Total /g/ syllables	**457**	**543**	**480**	**566**	**651**
za	—	85	—	97	108
ze	—	—	—	—	108
zi	—	92	—	86	107
zo	—	84	—	82	107
zu	—	81	—	91	109
zy	—	78	—	84	107
Total /z/ syllables	—	**420**	—	**440**	**646**

Note. Age groups include 3 years (C3), 4 years (C4), 5 years (C5), Grade 1 (G1), and adult (A). Em dashes indicate data not obtained.