CAMBRIDGE
UNIVERSITY PRESS

ARTICLE

Developmental differences in perceptual anticipation underlie different sensitivities to coarticulatory dynamics

Stella KRÜGER[1] and Aude NOIRAY[1,2] (iD)

[1]Linguistic Department, Laboratory for Oral Language Acquisition, University of Potsdam, Potsdam, Germany and [2]Haskins Laboratories, New Haven, CT, United States
Address for correspondence: Stella Krüger, Linguistic Department, University of Potsdam, Karl-Liebknecht-Strasse 24–25, 14476 Potsdam, II.14.2.40. E-mail: stelkrue@uni-potsdam.de

Abstract

Anticipatory coarticulation is an indispensable feature of speech dynamics contributing to spoken language fluency. Research has shown that children speak with greater degrees of vowel anticipatory coarticulation than adults – that is, greater vocalic influence on previous segments. The present study examined how developmental differences in anticipatory coarticulation transfer to the perceptual domain.

Using a gating paradigm, we tested 29 seven-year-olds and 93 German adult listeners with sequences produced by child and adult speakers, hence corresponding to low versus high vocalic anticipatory coarticulation degrees. First, children predicted vowel targets less successfully than adults. Second, greater perceptual accuracy was found for low compared to highly coarticulated speech. We propose that variations in coarticulation degrees reflect perceptually important differences in information dynamics and that listeners are more sensitive to fast changes in information than to a large amount of vocalic information spread across long segmental spans.

Keywords: speech perception; language development; vocalic coarticulation

Introduction

Coarticulation is an essential feature of speech dynamics allowing speech segments (e.g., words, sentences) to be produced in rapid, continuous, and meaningful speech streams. More specifically, coarticulation characterizes the temporal overlap of articulatory speech gestures for neighbouring segments (Browman & Goldstein, 1992; Fowler, 1980), which transfers into the acoustic speech signal via fine-grained intersegmental variations.

Despite being present very early on in infants' oral communication, coarticulatory patterning evolves while children develop spoken language fluency. In previous research we examined developmental trajectories in lingual vocalic anticipatory

CrossMark

coarticulation – that is, from the tongue, which is essential to the production of all vowels regardless of the language spoken. Thereby, lingual coarticulation degree reflects how much the tongue gesture for a target vowel overlaps with those for previous segments. We found that this degree decreased with increasing age both at the vicinity of segments or over longer temporal spans (intra-syllabic: Noiray, Abakarova, Rubertus, Krüger & Tiede, 2018; inter-syllabic: Rubertus & Noiray, 2018; both: Noiray, Wieling, Abakarova, Rubertus & Tiede, 2019. Young children in kindergarten exhibited a greater degree of anticipatory coarticulation than older children in primary school and adults. Results corroborated previous evidence of larger speech production units in children than in adults (reviews: Nittrouer, Studdert-Kennedy & Neely, 1996; Noiray et al., 2018, 2019). We also noted variations in vocalic coarticulation degree depending on the identity of preceding consonants and their speech motor demands. Those were present in both children's and adult's speech.

   While developmental differences in vocalic anticipatory coarticulation have been widely examined from the production perspective, little is known about their perception. In this study, we therefore investigated how developmental and contextual differences in vocalic anticipatory coarticulation transfer to the perceptual domain. To this end, we compared vocalic anticipatory perception in adult and seven-year-old child listeners. From the above-mentioned coarticulation studies (Noiray et al., 2018, 2019; Rubertus & Noiray, 2018), we selected speech stimuli from one adult, one age-matched child, and one three-year-old reflecting developmental differences in the degree of vocalic anticipation described above. With this approach, we aimed to elucidate whether adults and children differ in the perceptual integration of vocalic information present in segmental transitions. Results should further inform how contextually induced differences in vocalic anticipation influence listeners' sensitivity to anticipatory cues. Before laying out our empirical trajectory, we first review the empirical findings that motivated the present study.

Adult processing of coarticulatory information in adult and child speech

From a perceptual standpoint, vowels are particularly salient due to their long duration, loudness, and formant patterns (e.g., review: Nazzi & Cutler, 2019). In addition, the speech stream provides various vocalic cues that may be discernible beyond their acoustically defined temporal domain onto previous segments. Research has shown that adult listeners perceive those vocalic cues and can, for instance, predict upcoming vowel targets in anticipation of their acoustic onset (lingual & labial coarticulation: e.g., Katz, Kripke & Tallal, 1991; Nittrouer & Whalen, 1989; Waldstein & Baum, 1994; lingual coarticulation: e.g., Sereno & Lieberman, 1987; labial coarticulation: e.g., Sereno, Baum, Marean & Lieberman, 1987). However, it is unclear whether the developmental differences in vocalic anticipatory coarticulation reported here affect adult listeners' ability to predict upcoming vowels in anticipation of their acoustic onsets.

   Nittrouer and Whalen (1989) investigated differences in adult listeners' perception of fricative-vowel syllables produced by three-, four-, five-, seven-year-old children and adults. When only presented with a portion of the onset fricative, listeners predicted upcoming vowels more accurately and earlier in child speech – embedding a greater degree of vocalic anticipation than in adult speech. However, listeners did not perform well when presented with speech material produced by three-year-olds

despite exhibiting the highest coarticulation degrees among all child groups. Given that speech production is more variable in young children than adults (e.g., segment durations, acoustic and articulatory realizations: Goffman, Smith, Heisler & Ho, 2008; Green, Moore, Higashikawa & Steeve, 2000; Lee, Potamianos & Narayanan, 1999; Ménard, Schwartz, Boë & Aubin, 2007; Sadagopan & Smith, 2008), coarticulatory cues conveyed in the speech of three-year-old children may have been less informative for listeners in this study.

To summarize, anticipatory perception seems to benefit from children's large degree of vocalic anticipation, although this facilitative effect may be reduced in the case of young children due to the motor imprecision and variability characterizing their speech (e.g., Abakarova, Iskarous & Noiray, 2020; Zharkova, Hewlett & Hardcastle, 2011; Zharkova, 2020). This empirical discrepancy calls for in-depth investigations to uncover when developmental differences in coarticulatory patterns cease to facilitate adults' perception of child speech in comparison to adult speech.

Children's differences in the perception of coarticulatory information

Attention to segmental transitions, the most dynamical parts of the speech signal, is used for speech sound discrimination at the early stage of language development (e.g., Bertoncini & Boysson-Bardies, 2000; Houston & Jusczyk, 2003). However, to what extent older children use coarticulatory information for speech perception in comparison to adults remains an empirical question (reviews: Mayo, 2000; Mayo, Scobbie, Hewlett & Waters, 2003). Contrary to research addressing adults, most studies targeting children's perception of anticipatory coarticulation have used (semi-) synthesized speech partly because it is easier to manipulate specific acoustic parameters (e.g., Ohde, Haley & McMahon, 1996; Sussman, 2001). Sussman (2001) found that four- and five-year-old children could identify missing vowels in synthesized C_1VC syllables when the vocalic on- and offset transitions were preserved but they performed more poorly than adults. In addition, when listening to C_1VC syllables with C_1 transition information conflicting with the vocalic target, children's responses were in most cases influenced by the subsequent steady-state vowel formant information. Instead, adults were confused by conflicting vocalic information conveyed in the C_1 transition with respect to the subsequent steady-state vowel portion. As a result, their responses were less consistent than children. Sussman concluded that adults use fine-grained dynamic information to a greater extent than children – who may instead rely on longer, louder, and more prominent spectral information for speech processing. This finding is particularly important when it comes to children's processing of transitional cues because child speakers exhibit large degrees of vocalic coarticulation with vocalic information being transmitted over longer temporal extents than in adult speech.

Perception of varying coarticulatory degrees

As mentioned above, variations in vocalic anticipatory coarticulation have been observed as a function of the specific vowels and consonants assembled and whether they share similar articulatory demands on speech articulators (e.g., Abakarova, Iskarous & Noiray, 2018; Fowler, 1994; Noiray et al., 2018; Rubertus & Noiray, 2020; review: Recasens, 2018). While the production of all vowels necessitates a tongue gesture, some consonants do not (e.g., /b/ which recruits the lips), and therefore
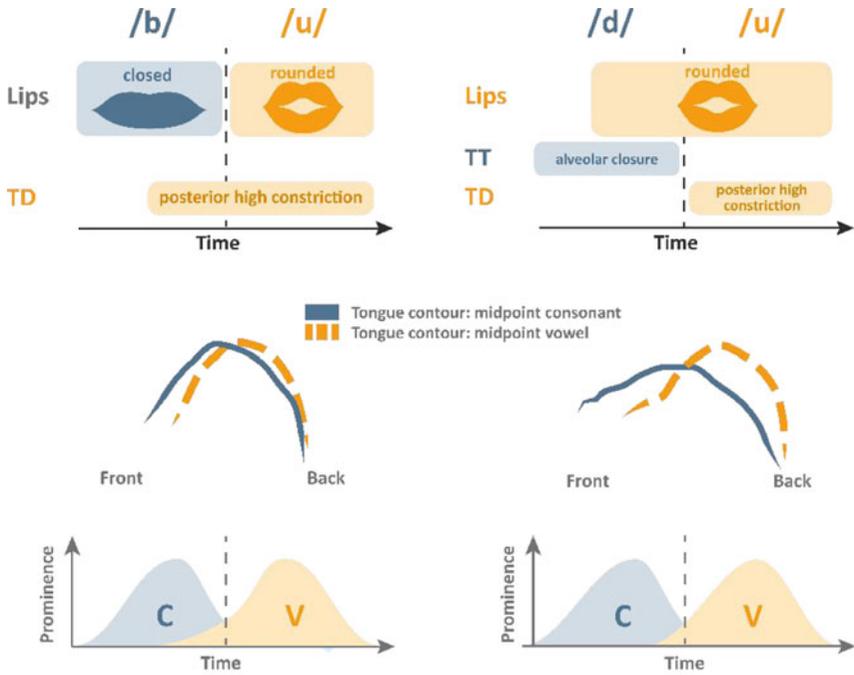
Figure 1. Conceptualization of the gestural organization for two articulators: lips and tongue (tongue dorsum, TD; tongue tip, TT) and segments' prominence in the production of two CV syllables: [bu] (left) and [du] (right). Tongue contours were plotted from production data of one adult female at consonant and vowel temporal midpoints. Dashed lines indicate the vocalic acoustic onset.

permit a large degree of vocalic overlap (e.g., /bu/ as illustrated in Figure 1). Other syllables (e.g., /du/ in Figure 1) instead allow minimal vocalic anticipation degree due to high demands on the tongue during both the consonant and vowel execution. In this case, the vocalic gesture cannot be initiated well in advance because it would compete with those for the planned consonant and therefore affect its intelligibility. This phenomenon has been largely described within the theory of COARTICULATORY RESISTANCE (in adults: e.g., Abakarova et al., 2018; Bladon & Al-Bamerni, 1976; review: Recasens, 2018). In the present study, we were interested whether this relationship translates into the perceptual domain.

In her perceptual investigation of coarticulatory resistance, Fowler (2005) used əCV disyllables for which consonants varied in their degree of resistance to vocalic anticipation. Stimuli were either spliced (appropriate vocalic information during /ə/) or cross-spliced (inappropriate vocalic information during /ə/). Differences in reaction time showed that listeners identified vowels in spliced disyllables with minimally resistant consonants better than with highly resistant consonants. The opposite pattern occurred for cross-spliced disyllables. These findings suggest that listeners' anticipatory perception is sensitive to transient coarticulatory effects from intervening consonants. However, to our knowledge no study so far has investigated effects of coarticulatory resistance on children's vocalic anticipatory perception, either when exposed to adult or to child speech.

Research questions and predictions

In the present study, we first asked whether listeners are sensitive to developmental differences in vocalic coarticulation degree (within-listener group comparisons). In our previous research, we had found greater coarticulation degrees in child compared to adult speech. Based on Nittrouer and Whalen's (1989) perception study, we expected listeners to benefit from greater coarticulation degrees and therefore identify vocalic targets to a greater extent in child speech compared to adult speech. However, given the absence of a facilitation effect observed with three-year-old speakers, we also anticipated lower perception accuracy when presented with three-year-old speech.

Second, we compared adult and child listeners' ability to use anticipatory cues and predict upcoming vowels (between-listener group comparisons). Due to less experience with their native language, we expected child listeners to perform more poorly than adults when presented with adult speech conveying lower coarticulation degrees and hence lesser anticipatory cues than child speech.

Last, we examined variations in vowel prediction accuracy as a function of whether its gestural specificities competed with those of preceding consonants. Given previous findings, we expected greater vowel accuracy in the context of /b/ that does not recruit the tongue articulator and hence is permeable to lingual coarticulatory influences than in the /d/ that recruits the tongue and therefore resists vocalic influences.

Method

Participants

We tested 29 seven-year-old children (range: 7;4–7;9 y; mean: 7;5 y) and 93 Adults (range: 18–55 y; mean: 31 y). All participants were monolingual German listeners living in the region of Berlin and Potsdam. None reported any language-related, hearing-related, or visual impairments, except one boy who was diagnosed with a red-green colour-blindness. All participants including children and their parents gave written consent to participate in the study. The Ethical Committee of the University of Potsdam approved the study.

Stimulus material

We used speech samples of the structure /ainə/ + CV produced by one woman (A), one seven-year-old girl at the end of the first grade (C7), and one three-year-old girl (C3). While the German indefinite article /aɪnə/ (schwa henceforth noted as @) was unaltered between samples, the segmental structure of the subsequent CV varied (C=/b/, /d/; V=/ iː, uː, yː, aː/), e.g., /aɪn@ biː/, /aɪn@ dyː/. To illuminate the temporal organization of anticipatory perception in child and adult listeners, we employed a gating paradigm (Grosjean, 1980) which consists in truncated speech sequences of increasing durations called 'gates' generated from the original speech samples.

Four temporal gates were generated with Praat (Boersma & Weenink, 2020) using a custom-made script that extracted segment boundaries at the offset of schwa, the midpoint and offset of the consonant, and at the midpoint of the vowel. The gates always included the article (/ain@/) and an increasing portion of the original untruncated sample. While the shortest gate would only include /ain@/, the longest possible gate would include the article, the consonant (/b/ or /d/), and a portion of
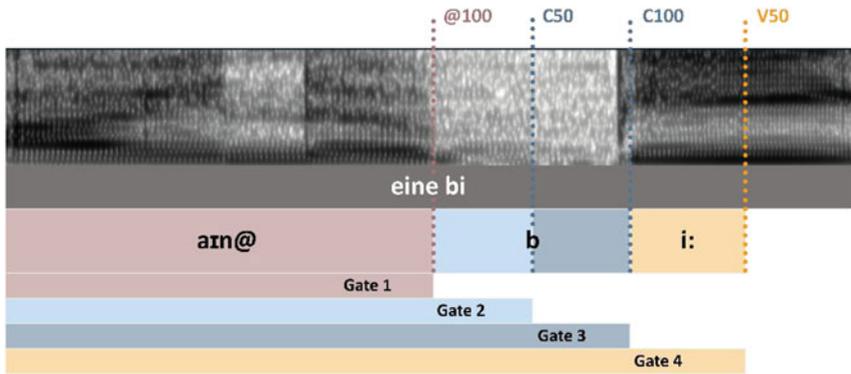
Figure 2. Temporal gates for each utterance (/ain@/+CV).

one of the four target vowels until its midpoint. Figure 2 depicts all gates within one possible sample. The midpoint of a segment is always tagged "50" and the offset tagged "100".

The speech stimuli were chosen from the coarticulation studies conducted in our laboratory: Noiray et al. (2018, 2019) and Rubertus and Noiray (2018, 2020). Details regarding the recording procedure, data processing, and analysis can be found there. For all speech stimuli the audio signal (sampling rate: 48kHz) and ultrasound images of the tongue surface (sampling rate: 48Hz) were recorded.

The selection of the speech stimuli was based on the developmental and contextually induced variations in coarticulation degree (e.g., Noiray et al., 2019). First, the speech of the youngest three-year-old child exhibited a globally higher coarticulation degree than the speech of the seven-year-old who in turn exhibited a higher coarticulation degree than the adult. Second, CDs varied as a function of intervocalic consonant with higher coarticulation degrees in the context of the labial consonant /b/ compared to the alveolar consonant /d/.

Coarticulation degrees were calculated via regression analyses, commonly used in coarticulation studies (e.g., Gibson & Ohde, 2007; Noiray, Ménard & Iskarous, 2013; Sussman, Duder, Dalston & Cacciatore, 1999). Regression slopes express systematic changes in the horizontal position of the tongue body during the consonant (X_c) dependent on those during the vowel (X_v). Table 1 reports coarticulation degrees (expressed as regression coefficient) for each of the speakers and consonants selected for the perception study. The seven-year-old shows a slightly higher coarticulation degree than the three-year-old in /d/ context.

In the next step, target speech samples were selected for each speaker controlling for segment durations. Peak amplitudes of each sound sample were normalized with Adobe Audition CC 2018.

Experimental procedure

All children and 42 out of the 93 adults were tested in the Laboratory for Oral Language Acquisition (LOLA) at the University of Potsdam. The perception test was run with OpenSesame (Mathôt, Schreij & Theeuwes, 2012), using a computer and headphones

Table 1. Lingual coarticulation degree (regression coefficients) between C50 and V50 for each consonantal context and speaker. The higher the coefficient, the greater the degree of lingual coarticulation.

Speaker	Consonant	Slope: C50 – V50
C3	b	0.92
	d	0.37
C7	b	0.73
	d	0.48
A	b	0.52
	d	0.19

(adults: over-ear super Lux headphones, frequency range: 10 Hz - 30 kHz, impedance: 32 Ω; children: on-ear JBL headphones, frequency range: 20 Hz - 20 kHz, impedance: 32 Ω). Participants were instructed either in written form on the screen (adults) or verbally (children) that they would hear utterances of different female speakers and different durations and would be asked to guess the upcoming vowel following the portion of speech heard as fast as possible. Responses were given via keypress on the keyboard. Four coloured keys were selected to represent the four target vowels. Vowel-colour associations were always shown on the screen. To avoid any bias due to handedness or favoured colour, we used two sets of colour-vowel associations which were counterbalanced across participants. Practice trials including both full length and gated speech items preceded each testing session to familiarize participants with the task. During the testing session, speech sequences were prompted in randomized order. Participants listened to a total of 96 sequences (three speakers * four gates * two consonants * four vowels). Each speech item was presented only once to consider children's limited attention span and keep the number of trials consistent between listeners. The perception test included breaks and the experimenter regularly encouraged children to stimulate their interest in the task.

Two testing conditions were employed with adult listeners: in our laboratory with the same experimental set-up used with children (43 adults) and online (51 adults) using an online version of the test implemented with SoSci Survey (Leiner, 2019). Participants were recruited via our university recruitment platform (and received credit points as part of their curriculum), social media, or within circles of acquaintances. In this version, participants used the computer mouse to select target vowels. The online test was designed to be fully self-explanatory. Prior to the perception test, we asked a few technical questions (e.g., type of headphone, type of device used). Commentary fields were included at the end of the test to assess the difficulty of the task and collect participants' overall opinions. While these precautionary measures may not guarantee the exact same experimental conditions across all participants, they allowed us to monitor adults' testing. Furthermore, when testing an effect of experimental environment (in the laboratory or online) on adult listeners' responses, results did not yield any significant difference in vowel predictions at any gate. All adult listeners were therefore grouped together for subsequent analysis.

Data analysis

Six adults and five children were excluded from data analysis for the following reasons: a) less than half of the number of observations in at least two conditions (five adults), b) accuracy for V50 below chance ($p = .25$; one adult and one child), c) technical issues (one child), or d) participant reportedly did not understand the task (three children). Therefore, 87 adults and 24 children were included in the analysis.

To assess accuracy in vowel anticipatory perception over different gates we employed generalized linear mixed models fitted in R (R Core Team, 2020) with the *lme4* package (version 1.1.23; Bates, Maechler, Bolker & Walker, 2015). Modelling was conducted with the *glmer* function for binomial data to assess probabilities of correct answers in each listener group (transforming our response variable to log odds or short "logits"). To adjust the *p*-value for pairwise multiple comparisons we used the *emmeans* function of the *emmeans* package (version 1.4.8; Lenth, 2020) using Tukey as the method for adjustment. The random effect structure was determined by the design of the study (repeated measurements for each listener and for item across listeners), visual inspection of grouping trends within the data, distribution of variances in model outputs, and theoretical considerations as suggested by Winter (2019). In models for which we had no interest in the specific effect of consonants (*m1* and *m2*) we combined consonants and vowels to the factor syllables. In addition, the complexity of the random effect structure was reduced in a stepwise fashion for cases where it was not supported by our data (singular fit or converge problems). Random effect structures were adapted for each model to fit these criteria.

Results

Perceptual sensitivity to developmental differences in vocalic anticipation

To determine whether listeners' anticipatory perception differ across speakers' age, we compared logits between speakers (A, C7, C3) for each temporal gate with listener groups (children or adults) and consonantal contexts (b, d) pooled together. For the random effect structure, we created the variable *nListSyl* that explicitly nested listener groups (child and adult) with syllable. We ran the following model:

$$m1 <- glmer(correct \sim Speaker^*Time + (1|participant) + (1 + Speaker|nListSyl)$$
$$+ (1|item), data = dat, family =' binomial', control = glmerControl(optimizer$$
$$= c("bobyqa")))]$$

Figure 3 represents listeners' vowel prediction functions for all temporal gates and speakers. Logits were transformed back to probabilities of the answer being correct, to ease interpretation.

First, regardless of speaker group and gate, mean probabilities for correct vowel predictions were above the 25% chance level (red dashed line). Still, prediction accuracy was greater close to the acoustic onset of the target vowel compared to earlier gates. This increase was not characterized by a strictly linear shape over time: it was flatter between the earlier gates (@100 versus C50) and the longer gates (C100 versus V50) but steeper between the midpoint and the offset of the consonant (C50 versus C100).

Second, when comparing listeners' perceptual performances across speakers, we found a reliable speaker's effect at the offset of the consonant between the three-year-old (yellow line) and adult (blue line) (logit coefficient: +1.19, SE = 0.32,
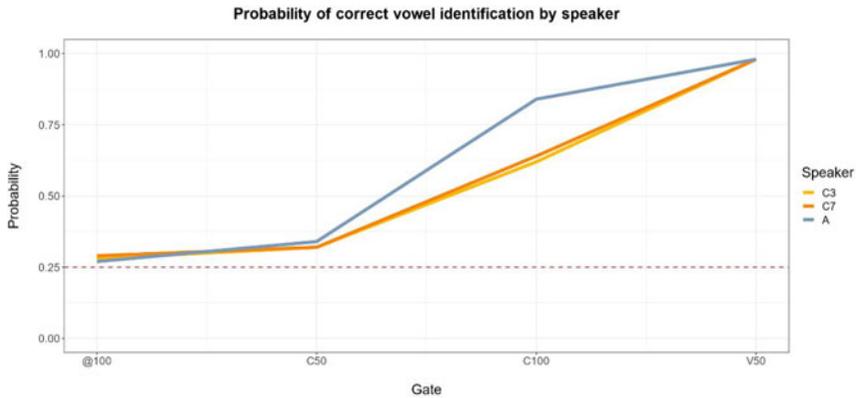
Figure 3. Probability functions for correct vowel prediction for each speaker across all gates with listener groups and consonantal contexts pooled together. Red dashed line indicates chance level performance.

Table 2. *Pairwise comparison of vowel prediction accuracy between speakers for each gate. Results were obtained via emmeans function with Tukey p-value adjustment. Cohort abbreviations are C3: three-year-old child, C7: seven-year-old child, and A: adult. Significance codes '***': p < .001; '**': p < .01; '*': p < .05; '.': p < 0.1*

Gate	Comparison	Logit coefficient	SE	Z	p-value	
@100	A - C3	−0.0388	0.312	−0.124	0.9915	
	A - C7	−0.0829	0.296	−0.280	0.9578	
	C3 - C7	−0.0441	0.274	−0.161	0.9858	
C50	A - C3	0.0656	0.312	0.210	0.9759	
	A - C7	0.1080	0.296	0.210	0.9759	
	C3 - C7	0.0424	0.274	0.155	0.9869	
C100	A - C3	1.1942	0.321	3.720	0.0006	***
	A - C7	1.0978	0.306	3.592	0.0010	***
	C3 - C7	−0.0964	0.277	−0.348	0.9353	
V50	A - C3	0.0403	0.429	0.094	0.9952	
	A - C7	−0.0126	0.420	−0.030	0.9995	
	C3 - C7	−0.0528	0.400	−0.132	0.9904	

z = 3.72, *p < .05*) and the seven-year-old (orange line) and the adult (logit coefficient: +1.10, SE = 0.31, z = 3.59, *p < .05*). For the two child speakers, listeners were equally successful at anticipating vowel targets. Results of pairwise comparisons between the speakers for each gate are summarized in Table 2.

Differences in vocalic anticipatory perception across listener groups

Next, we compared child and adult listeners' anticipatory vowel perception, to elucidate whether children at the age of seven already have similar predicting abilities as adults.
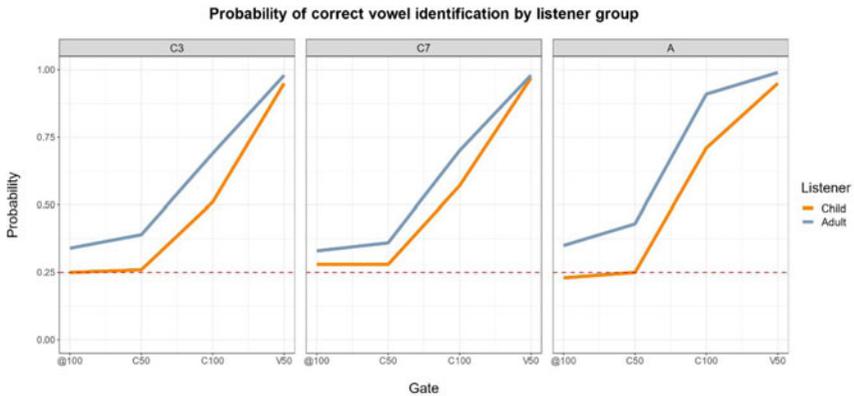
Probability of correct vowel identification by listener group

Figure 4. Probability functions for correct vowel prediction for each listener group across all gates separated by speakers (from left to right with ascending age) with consonantal contexts pooled together. Red dashed line indicates chance level performance.

We fitted a new model:

$$m2 <- glmer(correct \sim ListSpeak^{*}Gate + (1|Participant) + (1|Syllable)$$
$$+ (1|Item), data = dat, family$$
$$='binomial', control = glmerControl \ (optimizer ='bobyqa'))$$

For this model we created the nominal predictor ListSpeak which combines all levels of predictors associated to listeners' age and speakers' age (i.e., Child listener: C3 speaker, Adult listener: C3 speaker, …, Adult listener: Adult speaker) to avoid overfitting of the data and simplify p-value adjustment. Pairwise between listener group comparisons were carried out within speaker groups and with consonantal context pooled together.

Figure 4 reports probabilities of correct vowel prediction for each listener group. Overall, adult listeners (blue line) identified vocalic targets in all three speakers and from the earliest temporal gates onward better than child listeners. However, when comparing probability slopes between listener group adults' perceptual advantage was not uniform across all temporal gates. For instance, when presented with speech of the three-year-old speaker (left panel), adult listeners were reliably more successful than child listeners within the temporal domain of the consonant (C50: logit coefficient: +0.63, SE = 0.20, z = 3.17, $p = .0189$; C100: logit coefficient: +0.78, SE = 0.20, z = 3.98, $p = .001$) but not at the earliest gate (@100). For the seven-year-old speaker, listeners groups only differed significantly in anticipatory vowel perception starting from the offset of the consonant (middle panel). When tested with adult speech (right panel), adult listeners predicted the vowel significantly better than child listeners for all gates (for respective logit coefficient, SE, z, and p-value see Table 3). Results of pairwise comparisons between listeners by speaker for each gate are provided in Table 3.

Contextual effects on vocalic anticipatory perception

Next, we investigated whether the consonantal context preceding the vocalic targets modulated the observed differences in listeners' anticipatory perception. To address this question, we fitted a model with all possible combinations of listeners (children,

Table 3: Pairwise comparison of vowel prediction accuracy between listener groups for each speaker by gate. Results were obtained via emmeans function with Tukey *p*-value adjustment. Cohort abbreviations are C3: three-year-old child, C7: seven-year-old child, and A: adult. Significance codes '***': p < .001; '**': p < .01; '*': p < .05; '.': p < 0.1

Gate	Speaker	Logit coefficient	SE	Z	*p*-value	
@100	A	0.6146	0.198	3.099	0.0238	*
	C3	0.4543	0.197	2.308	0.1905	
	C7	0.2671	0.194	1.379	0. 7397	
C50	A	0. 8180	0.194	4.214	0.0004	***
	C3	0.6296	0.198	3.172	0.0189	*
	C7	0.3835	0.197	1.946	0.3737	
C100	A	1.3771	0.219	6.276	<.0001	***
	C3	0.7786	0.195	3.984	0.0010	***
	C7	0.5676	0.196	2.901	0.0431	*
V50	A	1.3357	0.449	2.973	0.0350	*
	C3	1.1746	0.416	2.825	0.0536	.
	C7	0.6955	0.443	1.569	0.6193	

adults), speakers (adult, three-, seven-year-old), and consonantal contexts (/b/, /d/) in a predictor called ListSpeakCon.

$$m3 <-glmer(correct \sim ListSpeakCons^*Time + (1|Participant) + (1|Vowel)$$
$$+ (1|Item), data = dat, family =' binomial', glmerControl(optimizer$$
$$=' optimx', optCtrl = list(method =' nlminb')))$$

Figure 5 illustrates correct prediction probability functions for child (upper row) and adult listeners (lower row) for both consonants by speaker (panels are ordered from left to right with ascending age). Even though visual inspection of Figure 5 suggests both listeners groups better predicted vowels in the context of /d/ than /b/, all pairwise comparisons did not yield any significant effect (Table 4).

Discussion

In previous research, we found that vocalic coarticulation degree differed as a function of age and phonetic contexts (Noiray et al., 2018, 2019; Rubertus & Noiray, 2018). The main goal of this study was to investigate whether the observed developmental and contextual changes in intersegmental speech organizations transfer in perception. More specifically, we asked whether and how children and adults listeners process variations in coarticulatory information in the speech stream. To address these questions, we investigated anticipatory vowel perception in 24 German seven-year-old and 87 adult listeners.

Developmental differences in coarticulatory information dynamics

Based on Nittrouer and Whalen's (1989) findings we predicted that highly coarticulated child speech – that is, with large vocalic influences on previous segments would
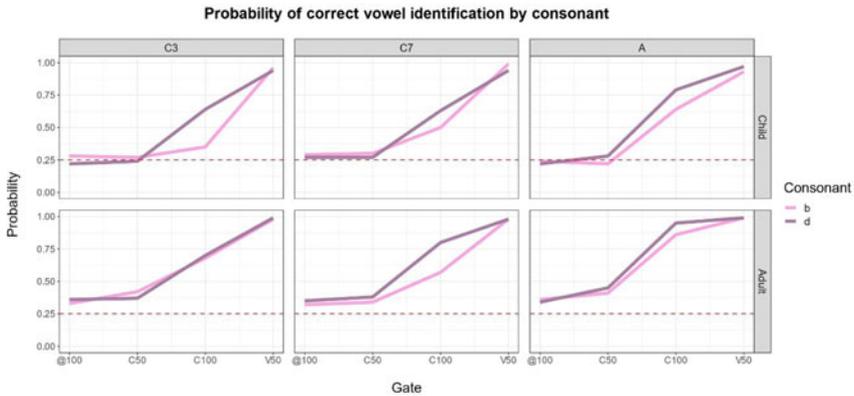
Figure 5. Probability functions for correct vowel anticipation for each consonant across all gates separated by speakers (from left to right with ascending age) for child (1st row) and adult (2nd row) listeners. Red dashed line indicates chance level performance.

facilitate listeners' vowel predictions when compared to lower coarticulated adult speech. Our findings did not support this hypothesis. While all listeners were able to predict upcoming vowels in speech sequences produced by either child or adult speakers from the midpoint of the consonant latest, the expected influence of coarticulation degree onto perceptual processing was reversed: perceptual accuracy was enhanced in adult as compared to child speech. This result aligns with previous reports of more variable speech motor patterns (e.g., the lips, jaw) (e.g., Green, Nip & Maassen, 2010; Kent, 1976; Lee et al., 1999; Ménard & Noiray, 2011). This is also reflected in children's speech outputs that remain quite variable compared to adults (e.g., Vorperian & Kent, 2007; Eguchi, 1969; Ménard et al., 2007). The speaker differences in anticipatory perception observed in our study may therefore be attributed to children's inconsistency in speech motor coordination making the production of individual segments more variable and thus anticipatory vocalic information less helpful than in adult speech (Nittrouer & Whalen, 1989). Nevertheless, children's imprecise production patterns may not solely explain why all listeners identified vowels produced by the three- and seven-year-old child equally well, especially given the large age gap between the two speakers.

We propose an alternative explanation for this unpredicted result. Previous research on anticipatory coarticulation has shown that, unlike adults, young children may initiate vocalic gestures two segments before the acoustically defined temporal domain of the vowel (Rubertus & Noiray, 2018; Noiray et al., 2019). Therefore, children do not only exhibit greater coarticulation degrees compared to adults, but they also show longer temporal extents of vocalic influences. As a consequence, vocalic information spreads from an earlier time point in child speech as compared to adults'. Vocalic information is more diffuse throughout the entire sequence and hence probably less salient than more compact information conveyed within a short temporal span. This pattern is schematized in the left panel of Figure 6. The amount of vocalic information over the four temporal gates (from left to right) is represented by a grey dashed line for highly coarticulated child speech and by a solid black line for minimally coarticulated adult speech.

Table 4. Pairwise comparisons of vowel prediction accuracy between consonantal contexts (/b/ was set as baseline) for each gate grouped by listener and speaker. Results were obtained via emmeans function with Tukey *p*-value adjustment. Cohort abbreviations are C3: three-year-old child, C7: seven-year-old child, and A: adult. Significance codes '***': p < .001; '**': p < .01; '*': p < .05; '.': p < 0.1

Listener	Speaker	Gate	Logit coefficient	SE	z	p-value
Child	C3	@100	0.3122	0.570	0.548	1
		C50	0.1443	0.570	0.253	1
		C100	−1.1638	0.571	−2.039	0.6663
		V50	0.2439	0.773	0.316	1
	C7	@100	0.1045	0.564	0.185	1
		C50	0.1795	0.568	0.316	1
		C100	−0.5329	0.560	−0.952	0.9985
		V50	1.9074	1.157	1.648	0.8912
	A	@100	−0.5115	0.531	−0.963	0.9984
		C50	−0.3105	0.568	−0.547	1
		C100	−0.7386	0.573	−1.289	0.9806
		V50	−0.9268	0.835	−1.110	0.9943
Adult	C3	@100	−0.1235	0.484	−0.255	1
		C50	0.2393	0.484	0.495	1
		C100	−0.1079	0.492	−0.219	1
		V50	−0.4230	0.733	−0.577	1
	C7	@100	−0.1232	0.484	−0.254	1
		C50	−0.1939	0.485	−0.400	1
		C100	−1.1043	0.492	−2.244	0.5175
		V50	−0.2549	0.701	−0.364	1
	A	@100	0.0926	0.483	0.192	1
		C50	−0.1510	0.482	−0.313	1
		C100	−1.0171	0.537	−1.894	0.7633
		V50	0.3856	0.791	0.488	1

More importantly, developmental differences in the amount of vocalic information over time also affect the shape of segmental transitions – in other words, its dynamics. Those dynamics are schematized in Figure 6 for child speech (middle panel) and adult speech (right panel). In child speech, because the lingual vocalic gesture is activated early, vocalic information is conveyed from the earliest temporal gate (@100) onward. The amount of vocalic information is not considerably lower than at the midpoint of the consonant (C50), which in turn is not considerably lower than at the offset of the consonant (C100). In sum, differences in the amount of vocalic information between adjacent gates do not change by much; therefore, the magnitude of change in vocalic information over time remains small (light green
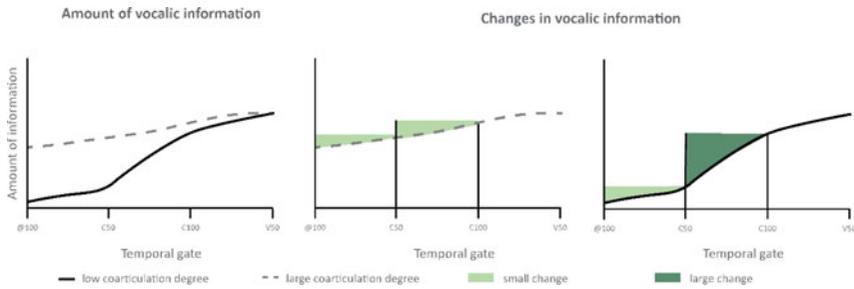
Figure 6. Schematized levels of prominence of vocalic information and changes in vocalic information in low (adult) and high (child) coarticulated speech over time (gate).

area). In adult speech however, the dynamics characterizing changes in vocalic information over time can be predicted differently (right panel): on the one hand, those changes are minimal between @100 and C50 (light green area). On the other hand, greater differences in vocalic information occur between C50 and C100 (dark green area) because at the offset of the consonant a high amount of vocalic information is now present in the signal.

Thus, when vocalic gestures are activated as early as in child speech, changes in vocalic information over time are less dynamic than in adult speech and hence presumably less perceptually salient. Assuming listeners parse the incoming speech signal incrementally, they may not pay much attention to vocalic information occurring far in advance from the vowel's acoustic onset because it might interfere with the perception of on-going segments (e.g., consonants).

Mayo and Turk's (2005) findings provide further support for the hypothesis that dynamical changes in vocalic information underlie listeners' better vowel predictions. They found that both adult and child listeners' perceptual sensitivity towards transitional cues was modulated by its spectral informativeness, i.e., its extent and duration. In a consonant categorization task, listeners assigned greater weight to transitional information when those cues were spectrally distinct across a consonantal pair (e.g., /do/-/bo/) than when spectrally similar (/de/-/be/). Similarly, listeners may show greater predictive ability in low as compared to high coarticulated speech because, in the former case, the acoustic speech signal conveys large changes in vocalic information during the consonant-vowel transition (see Figure 6, right panel: dark green area from C50 to C100) that are perceptually more salient than when spread over longer speech samples (see Figure 6, middle panel: light green area from @100 to C50 and C50 to C100). Furthermore, note that for the earliest gates (@100 to C50) dynamic changes in vocalic information are almost equal between speaker groups (compare light green area in the middle and right panel). These different patterns of information dynamics are also reflected in the finding of discrepancies between speaker groups as a function of speaker over time, with speaker differences not emerging before the offset of the consonant (cf. Figure 3).

Going a step further, we investigated whether the coarticulation patterns illustrated in Figure 6 are supported by actual lingual production data of the three speakers. To this end, horizontal tongue dorsum positions were normalized following Noiray et al. (2019). The most anterior tongue dorsum position during V50 among all trials for each speaker was set to zero (front in Figure 7) and the most posterior to one (back
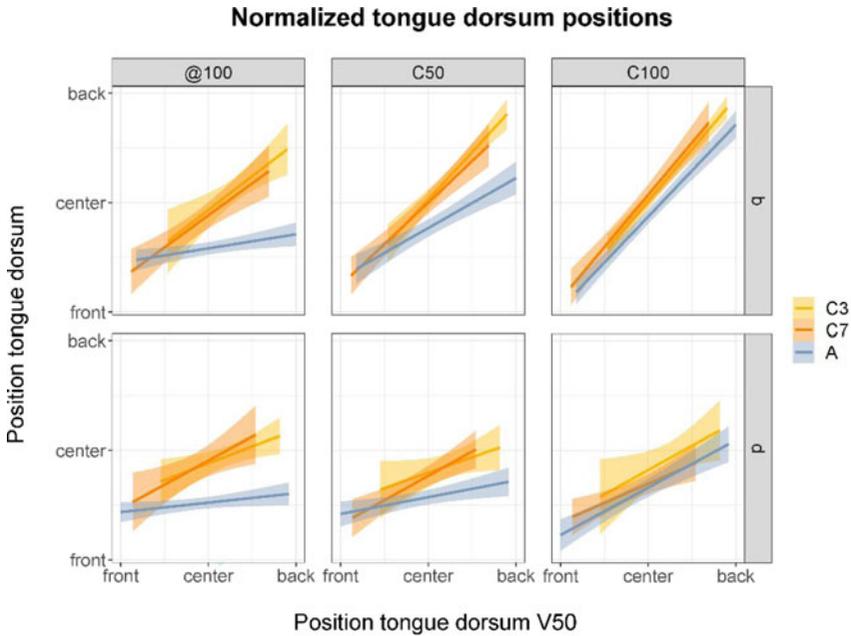
Normalized tongue dorsum positions

Figure 7. Dependence of the horizontal tongue dorsum positions (y-axis) across gates (@100: left panels, C50: mid panels, and C100: right panels) on the tongue dorsum position during the vowel midpoint (V50, x-axis) for each speaker. Tongue positions are indicated by the labels: front, center and back on the y-axis. 95% confidence intervals are shown.

in Figure 7). All other tongue dorsum positions were then scaled in relation to these references.

Figure 7 reports correlations between the horizontal tongue dorsum position at the midpoint of the vowel (x-axis) and the various temporal gates (y-axis) at @100, C50, and C100 (left, mid and right panels) by fitting a linear smooth using the *ggplot* function from the ggplot2 package (Version: 3.3.2; Wickham, 2016). Coarticulation degree is high if tongue position at a specific gate (e.g., @100) resembles those of the upcoming vowel midpoint. This relationship is captured by a steep slope.

The figure shows that both child speakers' slopes are higher than those for the adult speaker from the earliest gate (@100) and increase more continuously over time compared to the adult speaker, whose slopes instead increase abruptly between C50 and C100. Coarticulation slopes therefore captures well the patterns depicted in Figure 6.

Developmental differences in anticipatory perception

Next, we investigated whether the two groups of listeners differed in vocalic anticipatory perception. First, all listeners used vocalic coarticulatory information available in the temporal domain of the consonant to predict upcoming vowels. This finding aligns with previous studies (adults: e.g., Katz et al., 1991; Nittrouer & Whalen, 1989; Sereno & Lieberman, 1987; Waldstein & Baum, 1994; children: e.g., Ohde et al., 1996; Sussman, 2001). It further indicates that if a developmental shift occurs in children's attention to signal properties, seven-year-old child listeners at least behaved in similar

ways as adults. Furthermore, our finding suggests that children at this age can use sub-phonemic information and that their phonetic representations are already fine-grained (e.g., Cross & Joanisse, 2018; Desmeules-Trudel, Moore & Zamuner, 2020).

Yet, vowel prediction accuracy was greater in adult listeners compared to child listeners. There is evidence that children employ different strategies when making phonemic identifications; they are in general less consistent in their usage of minimal phonemic information than adults (e.g., Parnell & Amerman, 1978). Thus, children may be less successful in extracting or processing sub-phonemic information conveyed in a consonant because it is short and spectrally quite inconspicuous (aside from the burst) compared to vowels. Noiray et al. (2019) have argued that vowels may serve as "attractors" in children's gestural planning because vowels are acquired early and are perceptually more salient than most consonants. Turning to perception, Sussman (2001) found that children rely more on steady-state formants than transitional information if both conflicted with each other. This indicates that vowels also serve as attractors in children's perception.

Given that speech conveys all sorts of information, informative (e.g., anticipatory vocalic information) and uninformative, children must gain experience with their native language not only to become fluent speakers but also to become fluent listeners. The role of experience may also be particularly relevant when it comes to processing of truncated speech. Research found that perception is less robust to signal distortions in children than in adult speech (e.g., Eisenberg, Shannon, Schaefer Martinez, Wygonski & Boothroyd, 2000; Elliott, 1979). Even though the quality of spectral properties in the speech stimuli we used were not manipulated, the gating method employed is unusual. Adults are more experienced listeners than children; they can therefore better cope with all sorts of degraded speech (e.g., in noisy environments, truncated speech) while children may need long and clear speech chunks.

Perception of context variation in coarticulation degree

Last, we investigated how variations in vocalic coarticulation degree due to consonantal contexts affected listeners' anticipatory perception. Results suggest that consonant specific differences in coarticulation degree may be too subtle to be reflected in listeners' anticipatory vowel perception, at least with the experimental design selected in this study.

However, contrary to our expectations, listeners seemed to predict vowels better in /d/ than in /b/ contexts. This finding may provide further evidence for differences in information dynamics as outlined in Figure 6. Indeed, when preceded by the resistant lingual consonant /d/, vocalic gestures can only be initiated at the end of the consonant while it may be initiated earlier when co-produced with a labial consonant allowing large coarticulatory effects (e.g., /b/; Rubertus & Noiray, 2018). Thus, the amount of vocalic information is overall greater during the consonant /b/ (Figure 6, middle panel) than it is during /d/ (Figure 6, right panel) context, and changes in segmental information from the consonant to the vowel therefore less dynamic as in the context of /d/. Therefore, vocalic information may be perceptually more salient in the context of /d/ than in the context of /b/.

Conclusion

This study investigated developmental and consonant-related effects on both adults and children's vocalic anticipatory perception. Overall, both children and adult listeners

used coarticulatory information to predict target vowels in the speech of speakers from various ages. This finding not only demonstrates that both populations can utilize such information, it also suggests that, by the age of seven, children already have detailed phonetic representations and use them when parsing speech. Yet, children are not as successful as adults in using anticipatory coarticulatory information. This discrepancy may be due to differences in the way children process coarticulatory information, in auditory skills, in brain maturity to track speech and/or in their experience as a listener. Last, results suggest that anticipatory perception is sensitive to information dynamics – that is, listeners assign greater weight to the magnitude of change in the acoustic signal rather than to the overall amount of vocalic information spread throughout a speech sequence. This finding provides new insight on the role of segmental transitions for speech processing both from an adult and child perspective. However, more empirical research is needed to replicate the finding and inform how sensitivity to information dynamics contributes to the development of spoken language fluency.

References

Abakarova, D., Iskarous, K., & Noiray, A. (2018). Quantifying lingual coarticulation in German using mutual information: An ultrasound study. *The Journal of the Acoustical Society of America*, 144(2), 897–907. doi: https://doi.org/10.1121/1.5047669

Abakarova, D., Iskarous, K., & Noiray, A. (2020). Articulatory Strategies and Coarticulatory Patterns Across Age. Paper presented at the 12th International Seminar on Speech Production, Retrieved January 10, 2021, from https://issp2020.yale.edu/S03/abakarova_03_14_147_poster.pdf

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, 67(1), 1–48. doi: https://doi.org/10.18637/jss.v067.i01.

Bertoncini, J., & Boysson-Bardies, B. D. (2000). La perception et la production de la parole avant deux ans. *L'acquisition du langage*, 1, 95–136. doi: https://doi.org/10.3917/puf.fayol.2000.01.0095

Bladon, R. A. W., & Al-Bamerni, A. (1976). Coarticulation resistance in English/l. *Journal of Phonetics*, 4(2), 137–150. doi: https://doi.org/10.1016/S0095-4470(19)31234-3

Boersma, P., & Weenink, D. (2020). Praat: doing phonetics by computer [Computer program]. Version 6.1.41, retrieved 4 October 2020 from http://www.praat.org/

Browman, C. P., & Goldstein, L. (1992). Articulatory phonology: An overview. *Phonetica*, 49(3–4), 155–180. doi: https://doi.org/10.1159/000261913

Cross, A. M., & Joanisse, M. F. (2018). Eyetracking of coarticulatory cue responses in children and adults. *Language, Cognition and Neuroscience*, 33(10), 1315–1324. doi: https://doi.org/10.1080/23273798.2018.1484148

Desmeules-Trudel, F., Moore, C., & Zamuner, T. S. (2020). Monolingual and bilingual children's processing of coarticulation cues during spoken word recognition. *Journal of Child Language*, 1–18. doi: https://doi.org/10.1017/S0305000920000100

Eguchi, S. (1969). Development of speech sounds in children. *Acta Otolaryngol*, 257, 1–51.

Eisenberg, L. S., Shannon, R. V., Schaefer Martinez, A., Wygonski, J., & Boothroyd, A. (2000). Speech recognition with reduced spectral cues as a function of age. *The Journal of the Acoustical Society of America*, 107(5), 2704–2710. doi: https://doi.org/10.1121/1.428656

Elliott, L. L. (1979). Performance of children aged 9 to 17 years on a test of speech intelligibility in noise using sentence material with controlled word predictability. *The Journal of the Acoustical Society of America*, 66(3), 651–653. doi: https://doi.org/10.1121/1.383691

Fowler, C. A. (1980). Coarticulation and theories of extrinsic timing. *Journal of Phonetics*, *8*(1), 113–133.

Fowler, C. A. (1994). Invariants, specifiers, cues: An investigation of locus equations as information for place of articulation. *Perception & psychophysics*, *55*(6), 597–610. doi: https://doi.org/10.3758/BF03211675

Fowler, C. A. (2005). Parsing coarticulated speech in perception: Effects of coarticulation resistance. *Journal of Phonetics*, *33*(2), 199–213. doi: https://doi.org/10.1016/j.wocn.2004.10.003

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. doi: https://doi.org/10.1044/1092-4388(2007/008)

Goffman, L., Smith, A., Heisler, L., & Ho, M. (2008). The breadth of coarticulatory units in children and adults. *Journal of Speech, Language, and Hearing Research*, *51*(6), 1424–1437. doi: https://doi.org/10.1044/1092-4388(2008/07-0020)

Green, J. R., Moore, C. A., Higashikawa, M., & Steeve, R. W. (2000). The physiologic development of speech motor control: Lip and jaw coordination. *Journal of Speech, Language, and Hearing Research*, *43*(1), 239–255. doi: https://doi.org/10.1044/jslhr.4301.239

Green, J. R., Nip, I. S., & Maassen, B. (2010). Some organization principles in early speech development. *Speech Motor Control: New Developments in Basic and Applied Research*, *10*, 171–188. doi: https://doi.org/10.1093/acprof:oso/9780199235797.003.0010

Grosjean, F. (1980). Spoken word recognition processes and the gating paradigm. *Perception & psychophysics*, *28*(4), 267–283. doi: https://doi.org/10.3758/BF03204386

Houston, D. M., & Jusczyk, P. W. (2003). Infants' Long-Term Memory for the Sound Patterns of Words and Voices. *Journal of Experimental Psychology: Human Perception and Performance*, *29*(6), 1143–1154. doi: https://doi.org/10.1037/0096-1523.29.6.1143

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, Language, and Hearing Research*, *34*(6), 1222–1232. doi: https://doi.org/10.1044/jshr.3406.1222

Kent, R. D. (1976). Anatomical and neuromuscular maturation of the speech mechanism: Evidence from acoustic studies. *Journal of speech and hearing Research*, *19*(3), 421–447. doi: https://doi.org/10.1044/jshr.1903.421

Lee, S., Potamianos, A., & Narayanan, S. (1999). Acoustics of children's speech: Developmental changes of temporal and spectral parameters. *The Journal of the Acoustical Society of America*, *105*(3), 1455–1468. doi: https://doi.org/10.1121/1.426686

Leiner, D. J. (2019). SoSci Survey (Version 3.1.06) [Computer software]. Retrieved from https://www.soscisurvey.de

Lenth, R. (2020). emmeans: Estimated Marginal Means, aka Least-Squares Means. https://CRAN.R-project.org/package=emmeans. (R package version 1.4.8.)

Mathôt, S., Schreij, D., & Theeuwes, J. (2012). OpenSesame: An open-source, graphical experiment builder for the social sciences. *Behavior Research Methods*, *44*(2), 314–324. doi: https://doi.org/10.3758/s13428-011-0168-7

Mayo, C. (2000). *The relationship between phonemic awareness and cue weighting in speech perception: Longitudinal and cross-sectional child studies* (Doctoral dissertation). Retrieved from Queen Margaret University College, Edinburgh.

Mayo, C., Scobbie, J., Hewlett, N., & Waters, D. (2003). The influence of phonemic awareness development on acoustic cue weighting strategies in children's speech perception. *Journal of Speech, Language, and Hearing Research*, *46*(5), 1184–1196. doi: https://doi.org/10.1044/1092-4388(2003/092)

Mayo, C., & Turk, A. (2005). The influence of spectral distinctiveness on acoustic cue weighting in children's and adults' speech perception. *The Journal of the Acoustical Society of America*, *118*(3), 1730–1741. doi: https://doi.org/10.1121/1.1979451

Ménard, L., Schwartz, J. L., Boë, L. J., & Aubin, J. (2007). Articulatory–acoustic relationships during vocal tract growth for French vowels: Analysis of real data and simulations with an articulatory model. *Journal of Phonetics*, *35*(1), 1–19. doi: https://doi.org/10.1016/j.wocn.2006.01.003

Ménard, L., & Noiray, A. (2011). The development of lingual gestures in speech: an experimental approach to language development. *Faits de langues*, *37*(1), 189–202. doi: https://doi.org/10.1163/19589514-037-01-900000011

Nazzi, T., & Cutler, A. (2019). How consonants and vowels shape spoken-language recognition. *Annual Review of Linguistics*, *5*, 25–47. doi: https://doi.org/10.1146/annurev-linguistics-011718-011919

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. doi: https://doi.org/10.1121/1.398741

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, Language, and Hearing Research*, *39*(2), 379–389. doi: https://doi.org/10.1044/jshr.3902.379

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. doi: https://doi.org/10.1121/1.4763983

Noiray, A., Abakarova, D., Rubertus, E., Krüger, S., & Tiede, M. (2018). How do children organize their speech in the first years of life? Insight from ultrasound imaging. *Journal of Speech, Language, and Hearing Research*, *61*(6), 1355–1368. doi: https://doi.org/10.1044/2018_JSLHR-S-17-0148

Noiray, A., Wieling, M., Abakarova, D., Rubertus, E., & Tiede, M. (2019). Back from the future: Nonlinear anticipation in adults' and children's speech. *Journal of Speech, Language, and Hearing Research*, *62*(8S), 3033–3054. doi: https://doi.org/10.1044/2019_JSLHR-S-CSMC7-18-0208

Ohde, R. N., Haley, K. L., & McMahon, C. W. (1996). A developmental study of vowel perception from brief synthetic consonant–vowel syllables. *The Journal of the Acoustical Society of America*, *100*(6), 3813–3824. doi: https://doi.org/10.1121/1.417338

Parnell, M. M., & Amerman, J. D. (1978). Maturational influences on perception of coarticulatory effects. *Journal of speech and hearing research*, *21*(4), 682–701. doi: https://doi.org/10.1044/jshr.2104.682

R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Retrieved from https://www.R-project.org/.

Recasens, D. (2018). "Coarticulation" in *The Oxford research encyclopedia of linguistics*.

Rubertus, E., & Noiray, A. (2018). On the development of gestural organization: A cross-sectional study of vowel-to-vowel anticipatory coarticulation. *PloS one*, *13*(9). doi: https://doi.org/10.1371/journal.pone.0203562

Rubertus, E., & Noiray, A. (2020). Vocalic activation width decreases across childhood: Evidence from carryover coarticulation. *Laboratory Phonology: Journal of the Association for Laboratory Phonology*, *11*(1), 7. doi: https://doi.org/10.5334/labphon.228

Sadagopan, N., & Smith, A. (2008). Developmental changes in the effects of utterance length and complexity on speech movement variability. *Journal of Speech, Language, and Hearing Research 51*(5), 1138–1151. doi: https://doi.org/10.1044/1092-4388(2008/06-0222)

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. doi: https://doi.org/10.1121/1.394917

Sereno, J. A., & Lieberman, P. (1987). Developmental aspects of lingual coarticulation. *Journal of Phonetics*, *15*(3), 247–257. doi: https://doi.org/10.1016/S0095-4470(19)30569-8

Sussman, J. E. (2001). Vowel perception by adults and children with normal language and specific language impairment: Based on steady states or transitions?. *The Journal of the Acoustical Society of America*, *109*(3), 1173–1180. doi: https://doi.org/10.1121/1.1349428

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. doi: https://doi.org/10.1044/jslhr.4205.1080

Vorperian, H. K., & Kent, R. D. (2007). Vowel acoustic space development in children: A synthesis of acoustic and anatomic data. *Journal of Speech, Language, and Hearing Research*, *50*(6), 1510–1545. doi: https://doi.org/10.1044/1092-4388(2007/104)

Waldstein, R. S., & Baum, S. R. (1994). Perception of coarticulatory cues in the speech of children with profound hearing loss and children with normal hearing. *Journal of Speech, Language, and Hearing Research*, *37*(4), 952–959. doi: https://doi.org/10.1044/jshr.3704.952

Wickham, H. (2016). ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York. Retrieved from https://ggplot2.tidyverse.org

Winter, B. (2019). *Statistics for linguists: An introduction using R*. Routledge.

Zharkova, N., Hewlett, N., & Hardcastle, W. J. (2011). Coarticulation as an indicator of speech motor control development in children: An ultrasound study. *Motor Control*, *15*(1), 118–140. doi: https://doi.org/10.1123/mcj.15.1.118

Zharkova, N. (2020). Development of the voiceless sibilant fricative contrast in three-year-olds: an ultrasound and acoustic study. *Journal of Child Language*, 1–24. doi: https://doi.org/10.1017/S0305000920000653