




ARTICLE

# The acquisition of plain–emphatic consonant contrasts by Arabic-speaking children: An acoustic study

Anwar Alkhudidi<sup>1,2</sup> , Rebecca Holt<sup>1</sup> , Tuende Szalay<sup>3,4</sup>, Katherine Demuth<sup>1</sup> and Titia Benders<sup>1,5</sup> 

<sup>1</sup>Department of Linguistics, Macquarie University, <sup>2</sup>Department of English, Al-Leith University College, Umm al-Qura University, Makkah, Saudi Arabia, <sup>3</sup>Communication Sciences and Disorders, The University of Sydney, <sup>4</sup>School of Electrical Engineering & Telecommunications, University of New South Wales, Sydney, Australia and <sup>5</sup>Amsterdam Center for Language and Communication, University of Amsterdam, Amsterdam, The Netherlands

**Corresponding author:** Anwar Alkhudidi; Email: aakhudidi@uqu.edu.sa

(Received 28 October 2024; revised 30 May 2025; accepted 18 July 2025)

## Abstract

Arabic emphatic consonants are claimed to be late-acquired, likely due to their motoric complexity, involving both coronal and pharyngeal/uvular constrictions. Children's production has largely been studied using impressionistic data, with limited acoustic analysis. This study acoustically examines the acquisition of emphatic consonants in Saudi-Hijazi Arabic-speaking children aged 3–6 years. Thirty-eight children performed a real-word repetition task, after which consonantal and vocalic cues to the plain–emphatic contrast were measured. Results show that children produce both types of acoustic cues, with an age-related increase in the acoustic contrast and an overall alignment with adult patterns. Larger acoustic contrasts were found in vowels preceding rather than following consonants in word-medial positions, with no evidence for a difference between word-initial and word-final positions. The plain–emphatic contrast was greater for stops than fricatives and larger for female than male children. These findings are discussed in relation to the development of coarticulated consonants.

**Keywords:** emphatic consonants; Saudi Arabic; children; coarticulation

## ملخص الدراسة

اكتساب التباين بين الصوامت المطبقة وغير المطبقة لدى الأطفال الناطقين بالعربية: دراسة صوتية تتناول هذه الدراسة اكتساب الصوامت المطبقة لدى الأطفال الناطقين باللهجة السعودية الحجازية، في الفئة العمرية من 3 إلى 6 سنوات، مع التركيز على الجوانب الصوتية لهذا الاكتساب. رغم أن الصوامت المطبقة معروفة بتأخر اكتسابها لدى الأطفال نتيجة لتعقيدها النطقي الحركي، فإن معظم الأبحاث السابقة اعتمدت على الملاحظات الانطباعية دون التحليل الصوتي الدقيق. في هذه الدراسة، طلب من 38 طفل أداء مهمة تكرر كلمات حقيقية، وتم تحليل الخصائص الصوتية للصوامت المطبقة وغير المطبقة، والصوائت المصاحبة لها، للكشف عن الفروق الصوتية الناتجة عن التباين بينهما. أظهرت النتائج قدرة الأطفال

على إنتاج مؤشرات صوتية تعكس هذا التباين، كما لوحظ أن الفروق الصوتية تزداد مع تقدم الأطفال في العمر وتقترب تدريجيًا من أنماط النطق لدى البالغين. كما كشفت النتائج عن تباين أكبر في الصوائت السابقة للصوائت في المواضع الوسطى للكلمة مقارنة بالصوائت اللاحقة، مع عدم وجود فروق ملحوظة بين المواضع الأولية والنهائية للكلمة. بالإضافة إلى ذلك، كان التباين بين الصوائت المطبقة وغير المطبقة أوضح في الصوائت الانفجارية منه في الصوائت الاحتكاكية، وظهر بشكل أكبر لدى الإناث مقارنة بالذكور. وتمت مناقشة هذه النتائج في إطار النظريات المتعلقة باكتساب النظام الصوتي لدى الأطفال.

## 1. Introduction

Complex speech segments can be defined as single segments articulated with more than one constriction (Ladefoged & Maddieson, 1996). For example, the Arabic emphatic stop /tˤ/ is produced with a primary constriction in the alveolar region and a secondary constriction in the pharyngeal region (e.g., Al-Ani, 1970). Previous research indicates that complex speech segments are usually late-acquired (between 5 and 8 years old) – for example, liquids in English (Shriberg, 1993), doubly articulated stops in Igbo (Nwokah, 1986), and emphatic obstruents in Arabic (Amayreh, 2003; Amayreh & Dyson, 1998).

The current study aims to investigate how Arabic-speaking children aged 3–6 years acquire complex emphatic consonants. Through analysing the acoustic properties of these sounds in child-produced speech, we aim to gain insights into the developmental trajectory and linguistic factors that may influence acquisition.

### 1.1 Articulatory and acoustic characteristics of emphatics

Emphatic consonants, also known as “pharyngealised” or “uvularised” consonants, are a set of sounds in Arabic that are produced with two constrictions: a primary constriction in the coronal region and a secondary constriction whose exact location is subject to debate – some researchers refer to these sounds as pharyngealised, while others argue they are uvularised segments (Al-Ani, 1970; Al-Tamimi *et al.*, 2009; Al-Tamimi & Heselwood, 2011; Al-Tamimi, 2015, 2017; Bin-Muqbil, 2006; Ghazeli, 1977; Laufer & Baer, 1988; Zawaydeh, 1999; Zawaydeh & de Jong, 2011). In Modern Standard Arabic (MSA), which serves as the written and formal language of the Arab world, the four emphatic consonants /sˤ tˤ dˤ ðˤ/ are maintained and clearly distinguished from each other (e.g., Al-Ani, 1970). However, variations in the realisation of emphatic consonants are evident across different spoken varieties of Arabic. For example, /dˤ/ is realised as [ðˤ] in most of the Gulf and Tunisian dialects, while /ðˤ/ is realised as either [dˤ] or [zˤ] in Hijazi, Egyptian, and Lebanese dialects (Watson, 2002).

Arabic emphatics differ acoustically from their plain counterparts. The primary acoustic correlate of emphasis is the lowering of the second formant (F2) of adjacent vowels, most prominently in /a/ compared to /i/ and /u/ (Algryani, 2014; Bin-Muqbil, 2006; Bukshaisha, 1986; Card, 1983; Hassan & Esling, 2011; Israel *et al.*, 2012; Jongman *et al.*, 2011; Kalaldeh & Al-Shdaifat, 2019; Khat tab *et al.*, 2006; Laufer & Baer, 1988; Zawaydeh & de Jong, 2011). Other cues to emphasis may include raising of the first and third formant frequencies (F1 and F3) of adjacent vowels and, for stops, shorter voice onset time (VOT) and lower spectral mean of the burst (Abudalbuh, 2010; Al Malwi, 2017; Al-Khairy, 2005; Khat tab *et al.*, 2006). In addition to affecting immediately adjacent vowels, the emphasis feature may also spread to adjacent syllables with dialect-specific parameters for the domain and direction of spread, as well as the types of segments that block spreading (Alhammad, 2014; Al-Masri & Jongman, 2004; Card, 1983; S. Davis, 1995; Israel *et al.*, 2012; Younes, 1993).

Moreover, there may be some gender differences in the acoustic realisation of the plain–emphatic contrast. Findings, however, are inconsistent. Females were reported, in some studies, to show larger acoustic differences between emphatic and plain consonants than males – for example, greater lowering of F2 (Al-Masri & Jongman, 2004; Almbark, 2008; Almuhaimeed, 2021; Kahn, 1975). On the other hand, some studies have reported that males exhibit stronger cues to emphatics than females (Abudalbh, 2010; Lehn, 1963; Wahba, 1993).

The Arabic variety of interest in this study is the Saudi Arabian dialect, particularly the urban Hijazi dialect spoken in Taif City (Alzaidi et al., 2019). In this dialect, /t<sup>ʕ</sup>/ and /s<sup>ʕ</sup>/ are realised without allophonic variation and contrast phonemically with the plain (non-emphatic) counterparts /t/ and /s/, respectively (Table 1). The current study is, therefore, only concerned with the voiceless emphatics /t<sup>ʕ</sup>/ and /s<sup>ʕ</sup>/.

## 1.2 The acquisition of emphatics

Previous research on the phonological development of Arabic-speaking children classifies the emphatic consonants among the late-acquired sounds (Al-Awaji, 2014; Al-Buainain et al., 2012; Alqattan, 2015; Amayreh, 2003; Amayreh & Dyson, 1998; Ammar & Morsi, 2018; Ayyad, 2011; Dyson & Amayreh, 2000; Hamdan & Amayreh, 2007; Saleh et al., 2007). Typical production errors include emphatic omissions for very young children (1–2 years) or de-emphasis for older children (3–8 years), that is, no secondary articulation, resulting in neutralisation of the plain–emphatic contrast – for example, /t<sup>ʕ</sup>/ may be realised as [t] (Abdoh, 2011; Alajroush, 2019; Al-Awaji, 2014; Al-Buainain et al., 2012; Alqattan, 2015; Amayreh & Dyson, 1998; Ammar & Morsi, 2018; Dyson & Amayreh, 2000; Mashaqba et al., 2022; Saleh et al., 2007).

The generally late acquisition of emphatics has been linked to factors commonly suggested to account for late acquisition across other segment types: articulatory complexity, frequency of occurrence, and functional load (Contreras Kallens et al., 2023; Stokes & Surendran, 2005). First, the required precise coordination between articulatory gestures of the front part and root of the tongue may slow down the acquisition of emphatics, with de-emphasis often cited as evidence. Second, input frequency and functional load are widely recognised as key factors in early phonological development; sounds that occur more frequently in the input and serve to distinguish more minimal pairs are generally acquired earlier (Pye et al., 1987; Stokes & Surendran, 2005; Surendran & Niyogi, 2003). In the case of Arabic, emphatic consonants rank between 16th and 27th in frequency out of the 28 consonants in MSA (Amayera et al., 1999, as cited in Mashaqba et al., 2022), which Amayera et al. suggest may account for their late acquisition. Amayreh (2003) similarly attributes the delayed acquisition of emphatics in Jordanian Arabic to their low functional load, although no measures thereof were provided. However,

**Table 1.** The Arabic voiceless emphatic obstruents and their plain counterparts

Grapheme	IPA	Example	Gloss
ط	/t <sup>ʕ</sup> /	/t <sup>ʕ</sup> i:n/	<i>Mud</i>
ت	/t/	/ti:n/	<i>Fig</i>
ص	/s <sup>ʕ</sup> /	/s <sup>ʕ</sup> a:h/	<i>Screamed</i>
س	/s/	/sa:h/	<i>Melted</i>

assessing the impact of these factors for any spoken variety of Arabic, including Saudi-Hijazi Arabic, is complicated by the lack of a (digitised) phonemic lexicon and frequency counts (e.g., Al-Thubaity, 2015). As Ingram and Babatsouli (2024) note, a complete understanding of phonological development depends on research into under-represented languages, where many assumptions drawn from well-studied languages may not necessarily apply to less-documented languages, such as Arabic.

Most estimates of the age of acquisition for emphatics are based on adults' perceptual judgements of children's production and vary substantially across studies, from 2;06 to 7;11. These variations could be due to dialect differences, linguistic factors, and methodology (Mashaqba *et al.*, 2022), or a combination of factors. In terms of dialect, earlier acquisition of (some) emphatics has been observed for Egyptian and Kuwaiti children at 2;06 and 3;00, respectively (Alqattan, 2015; Morsi, 2003), and later acquisition for Ammani and Syrian children, at 4;06 – 4;11 or 6;00 (Amayreh & Dyson, 1998; Mashaqba *et al.*, 2022; Owaida, 2015). Linguistic factors, such as manner of articulation and word position, may also affect the age of acquisition, in line with well-documented trends cross-linguistically (Jakobson, 1968). The emphatic stops are typically acquired before the emphatic fricatives – for example, 3;10 for /t<sup>s</sup>/ versus 4;07 for /s<sup>s</sup>/ for Kuwaiti children (Ayyad, 2011), and word-initial (WI) emphatics are acquired earlier (between 6;00 and 6;11) than word-final (WF) (7;00 and 7;11) for Jordanian children (Mashaqba *et al.*, 2022). Such linguistic factors are not always systematically addressed in studies on emphasis acquisition, highlighting the need for more detailed investigation to fully capture their influence across different developmental contexts.

One prior study has supplemented perceptual judgements with acoustic analysis to investigate the acquisition of Arabic emphatic consonants in 2- to 7-year-old Ammani-Arabic-speaking children (Mashaqba *et al.*, 2022). Using the criteria for the *customary*, *acquisition*, and *mastery* stages of consonant development proposed by Sander (1972), emphatics were *customarily* produced by age 3 years – more than 50% of the children were able to produce these sounds correctly as judged by adults. The *acquisition* stage was considered to occur by age 4 years with at least 75% correct production. The *mastery* stage was by age 5 years, with over 90% of children producing emphatics correctly.

Despite the *mastery* stage starting at 5 years, the acoustic parameters of children's productions, particularly F2 – F1 values of adjacent vowels, were not adult-like for all manners of articulation and word positions until well over the age of 6 years.<sup>1</sup> For example, children produced acoustically adult-like emphatics earlier in WI and word-medial (WM) positions, while emphatics in WF positions were not adult-like until the age of 7;01 – 7;11. Moreover, the fact that children produced acoustically non-adult-like realisations, even when adults judged their productions largely to be correct, suggests that the fine motor control and specific articulatory gestures required to produce adult-like emphatics may take longer to develop (Zharkova, 2018). Conversely, for children who were perceived to neutralise the contrast, for example, their target /t<sup>s</sup>/ was perceived as [t], the acoustic analysis showed a statistically significant contrast in F2 – F1 between vowels adjacent to plain and emphatic consonants. This suggests that the phonological contrast was covertly produced (Barton & Macken, 1980).<sup>2</sup>

<sup>1</sup>Adult-like production was defined as F2 – F1 values not significantly different from those of adults, at  $p > .05$ .

<sup>2</sup>While Mashaqba *et al.* (2022) suggest a covert contrast, this term, as introduced by Scobbie (1998), refers to a phonological distinction that may not be overtly manifested in the acoustic signal in the same way it is in adult speech. Conversely, a “non-significant acoustic contrast” indicates that there is no detectable difference in the acoustic properties between two sounds based on statistical analysis. In the context of the article, the contrast was statistically insignificant for children below the age of 5 years but became significant from the age of 5 years onwards.

The work by Mashaqba et al. (2022) thus highlights that acoustic measurements can capture gradual, age-related changes during the acquisition process of emphatics, both before acquisition begins and after it is completed according to adult judgements. However, while Mashaqba et al. (2022) acoustically examined children's production of the plain–emphatic contrast, they focused solely on one acoustic cue to the contrast, that is, vocalic cues, which may not fully reflect the range of strategies children might use to realise these contrasts. For instance, children might realise the plain–emphatic contrast by only implementing a vocalic difference, such as producing a back vowel. Thus, it is important to consider both the consonantal and vocalic cues when examining children's production of the plain–emphatic contrast, which may shed light on how children realise this contrast.

### 1.3 The current study

The current study, therefore, aimed to build on Mashaqba et al.'s (2022) acoustic work by examining the developmental trajectory of the production of emphatic consonants in 3- to 6-year-old Saudi Arabic-speaking children, a dialect that is still understudied with respect to children's phonological and phonetic development in general (Abdoh, 2011; Alajroush, 2019; Turki, 2022). Productions were elicited for stops and fricatives, in WI, WM, and WF positions. Acoustic characterisation of children's emphatics was conducted by examining not only formant frequencies of adjacent vowels but also VOT durations of stop consonants.

We hypothesised that the acquisition process would be gradual (Macken & Barton, 1980; Mashaqba et al., 2022). Therefore, older children may show larger acoustic differences between plain and emphatic consonants than younger children. We additionally expected that the plain–emphatic contrast might be larger in initial than in final positions, especially for younger compared to older children (Amayreh, 2003; Mashaqba et al., 2022). For WM position, a larger acoustic contrast was expected for vowels immediately preceding than immediately following the target consonant (Al-Masri & Jongman, 2004; Zawaydeh, 1999). We also expected a larger plain–emphatic contrast for stops compared to fricatives (Amayreh, 2003; Jakobson, 1968). Finally, gender differences might be observed such that females may produce larger contrasts compared to males (Al-Masri & Jongman, 2004; Almbark, 2008; Almuhaimeed, 2021; Kahn, 1975). Because a comprehensive description of emphatics in the Saudi-Hijazi dialect is currently still missing from the literature, adult participants were also included, serving as reference data.

## 2 Methods

### 2.1. Participants

Thirty-eight monolingual Arabic-speaking children aged between 3;02 and 6;11 from Saudi Arabia participated ( $M_{\text{age}} = 5;01$ , female = 20, male = 18) – see Table A.1 in Supplementary Appendix A for more details. All children were reported by their parents to have typical cognitive and language development. Thirteen native Saudi Arabic-speaking adults between 17 and 38 years old ( $M_{\text{age}} = 24$ , female = 5, male = 8) were also recruited.

The study was approved by the Macquarie University Faculty of Medicine, Health, and Human Sciences Human Research Ethics Sub-Committee. Adult participants and parents or legal guardians of child participants signed an informed consent form before

the experiment. Verbal assent was obtained from children before the start of the procedure.

2.2. Stimuli

Target consonants examined in this study were the alveolar voiceless emphatic obstruents and their plain counterparts, that is, the oral-stop contrast /t<sup>ʕ</sup>/ vs. /t/ and the fricative contrast /s<sup>ʕ</sup>/ vs. /s/. Consonants were elicited in real words with a 'Ca.CaC structure, in WI and WM positions for the stops, and WI, WM, and WF positions for the fricatives.

The selection of stimulus items to elicit these targets was guided by the need to balance phonological control with developmental appropriateness. While minimal pairs are often preferred in phonetic research for isolating segmental contrasts, they were not always feasible for the current study due to constraints on lexical familiarity to young children, morphosyntactic structure (more details below), and picturability. As such, the study employed 15 near-minimal word pairs (for a total of 30 words), comprising 6 oral-stop pairs and 9 fricative pairs and contrasting an emphatic and plain target consonant within each pair (see Table 2 for examples; for a full list of stimuli, see Table A.2 in Supplementary Appendix A).

All target words were real, picturable nouns or verbs, and were selected to be familiar to young children. Familiarity was assessed through an informal lexical survey involving 14 Saudi-Hijazi Arabic-speaking parents of children aged 2;02–7;11 (*M*<sub>age</sub> = 4;02). A word was deemed familiar if at least 60% of parents reported that their child understood it.

In designing the stimulus set, additional morphosyntactic and phonological constraints were considered. In Arabic, WF /t/ usually occurs as a suffix to mark the third-person singular feminine as in (1) or the first-person singular as in (2). As a result, it was difficult to find word pairs where the plain /t/ in WF position was part of the stem, not a suffix/morpheme. Moreover, young Arabic-speaking children primarily use verbs that are marked for singular masculine, the default or unmarked verb form, as in (3), before developing the verbal agreement paradigm after the age of 4 years (Aljenaie, 2010;

**Table 2.** Stimuli sample of /t<sup>ʕ</sup>/ versus /t/and /s<sup>ʕ</sup>/ versus /s/ in word-initial, word-medial, and word-final positions

	Stops		Fricatives	
	Emphatic /t <sup>ʕ</sup> /	Plain /t/	Emphatic /s <sup>ʕ</sup> /	Plain /s/
Word-initial	طرف /t <sup>ʕ</sup> araf/ “edge” SG	ترك /tarak/ “left something” 3.SG.M	صحن /s <sup>ʕ</sup> ahan/ “plate” SG	سحب /sahab/ “pulled” 3.SG.M
Word-medial	بطل /bat <sup>ʕ</sup> al/ “hero” SG.M	فتح /fatah/ “opened” 3.SG.M	بصل /bas <sup>ʕ</sup> al/ “onion” SG	غسل /yasal/ “washed” 3.SG.M
Word-final			رقص /ragas <sup>ʕ</sup> / “danced” 3.SG.M	رفس /rafas/ “kicked” 3.SG.M

Abbreviations: SG, singular; 3, third person; M, masculine.

Omar, 1973). Therefore, a pair like (4a,b) is likely to be produced as (5a,b) by the youngest age group in this study (3 years old), where the final /t/ would be omitted, preventing comparison between the members of the pair.

- |     |  |                                     |
|-----|--|-------------------------------------|
| (1) | /ramat/                                    | “threw” 3.SG.F                      |
| (2) | /ram3:t/                                   | “threw” 1.SG                        |
| (3) | /rama/                                     | “threw” 3.SG.M                      |
| (4) | a. /rabat <sup>s</sup> / versus b. /ramat/ | “tied” 3.SG.M versus “threw” 3.SG.F |
| (5) | a. /rabat <sup>s</sup> / versus b. /rama/  | “tied” 3.SG.M versus “threw” 3.SG.M |

Therefore, word pairs with the target stops in the final position were excluded. Consequently, target stops were restricted to WI and WM positions, whereas target fricatives were elicited in all three word positions.

To minimise phonetic differences between stimulus words, all target stimulus words were controlled for number of syllables (all disyllabic with the structure 'CV.CVC) and vowels (all low front vowel /a/). Stress was fixed on the initial syllable, such that WI targets occurred in stressed syllables and medial/final targets in unstressed positions. Apart from the target emphatics, words did not include any other emphatic consonants, to avoid emphasis spread from non-target consonants, inducing F2 lowering. Additionally, to minimise any potential effect of adjacent non-target consonants, non-target consonants nearest to the target consonant within each pair were matched (where possible) for place of articulation.

Two high-frequency, picturable practice items (/mo:z/ “banana” and /saj:arah/ “car”) were included at the beginning of the task to familiarise participants with the elicitation format. All auditory stimuli were recorded by the first author, a native speaker of Saudi-Hijazi Arabic, using a high-quality lavalier microphone (*Rode Lavalier GO*) and *Zoom H2n* recorder in a quiet room. The pictures and audio recordings were then piloted with a small sample of adults and children to ensure that they effectively elicited the target sounds.

### 2.3. Procedure

Data collection was conducted in a hybrid in-person and online format. Speech recordings were collected by a trained research assistant who was present with participants to set up recording equipment and closely monitor the experiment. Stimuli were presented to participants in person on a MacBook Pro laptop. Experiment instructions were delivered to participants online via a videotelephony platform (*Zoom*). The first author monitored the progression of the experiment and participants' performance via the remote-control function in *Zoom*. Either before or after the experiment, adult participants and child participants' parents filled out a short language background questionnaire (see [Supplementary Appendix B](#)).

Data collection took place in schools, daycares, and participants' homes. In each location, recordings were made in the quietest room available. To improve the acoustic conditions of the recordings, foam isolation shields were placed to the sides and in front of participants to reduce background noise and reverberation. Participants' responses were recorded using a *Rode Lavalier GO* omnidirectional lavalier microphone connected to a *Zoom H2n* voice recorder.



Two practice trials were implemented first and were repeated until participants understood the task. Target consonants were then elicited through a single-word repetition task with a picture prompt. In this task, a picture on a computer screen coupled with an audio prompt of the picture's name was presented to participants. Participants were then instructed to repeat the word five times. To minimise any potential effects of presentation order, stimuli were pseudorandomised by participant, with the restriction that successive trials alternated between words with a target emphatic and words with a target plain consonant.

#### 2.4. Data selection and exclusion criteria

Out of the five repetitions of each target word, only the three most spontaneous and clear repetitions were considered for acoustic analysis (e.g., no hesitations, or extremely high or low intonations, etc.). These were mostly the three middle repetitions. This resulted in a total of 3,403 tokens produced by children and 1,170 tokens produced by adults.<sup>3</sup>

Tokens were then excluded from the consonant and vowel analyses if they exhibited excessive background noise, had a clipped waveform signal, or included mispronunciations or disfluencies that rendered the acoustic measurements unreliable for addressing the research questions (e.g., /fas<sup>s</sup>al/realised as /st<sup>s</sup>atar/) ( $n = 96$ ). Additionally, tokens where the target acoustic parameters could not be measured were excluded from the consonant analysis ( $n = 18$ ). In most cases, this was caused by the absence of a meaningful closure in a target stop consonant in medial position, which may imply that the target stop was realised as a fricative-like segment. This resulted in a total of 4,459 tokens (children = 3,305; adults = 1,154) selected for further analyses.

#### 2.5. Data analyses

**Acoustic annotation.** Selected data were automatically segmented using Munich Automatic Segmentation system (MAUS) forced aligner (Kisler *et al.*, 2017; Schiel, 1999) with the Arabic grapheme-to-phoneme converter (Al-Tamimi *et al.*, 2022). Segment boundaries were hand-corrected in Praat (Boersma & Weenink, 2023) according to the following acoustic criteria: VOT for stop consonants in initial and medial positions was marked from the onset of the burst release to the onset of periodicity in the waveform<sup>4</sup> (Figures 1 and 2). In cases where multiple bursts were observed, the first burst was taken as the onset of the burst release (e.g., Davis, 1994; Sučková, 2020; Tar, 2014; Yang, 2018).

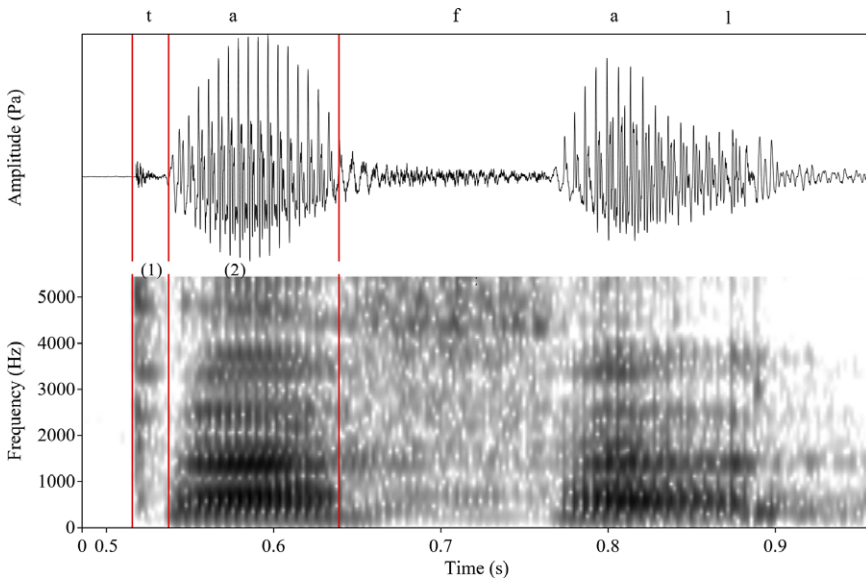
Vowel onsets were identified by the first presence of F2, while vowel offset was marked at the end of F2. For WI and WF tokens, only the vowel immediately adjacent to the target consonant was considered for analysis (Figures 1 and 3). For WM, both vowels preceding and following the target consonant were considered (Figure 2).

Blind to the original hand-correction, a second coder hand-corrected around 10% of the automatically segmented data (children = 343; adults = 112 tokens). For stop consonants, VOT measurements are typically considered in agreement when they differ by <10 ms (Fabiano-Smith & Bunta, 2012; Stoehr *et al.*, 2017). However, in our data, we

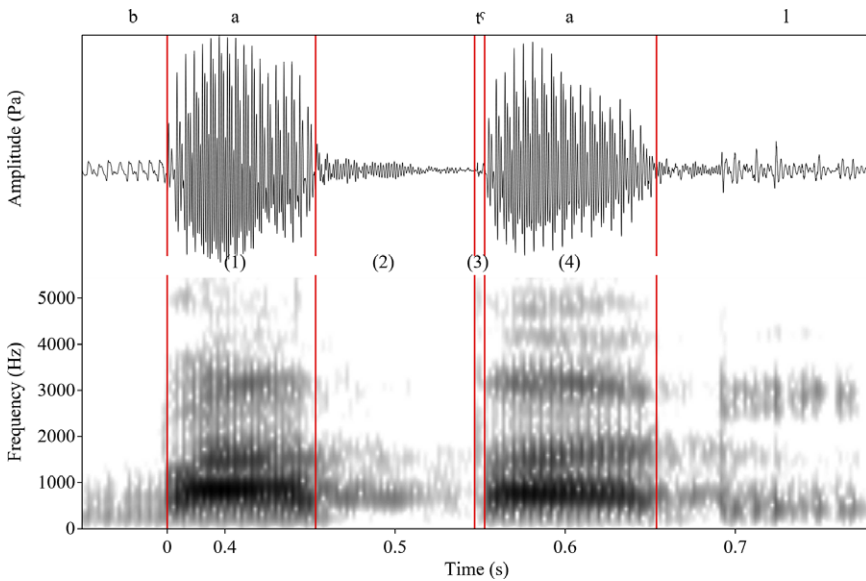
<sup>3</sup>In some cases, children did not comply with instructions and produced fewer than three repetitions. This resulted in 17 missing repetitions in total.

<sup>4</sup>For stops in WM, the closure duration was marked in addition to VOT (Figure 2); however, for this study, only VOT was considered for analysis.

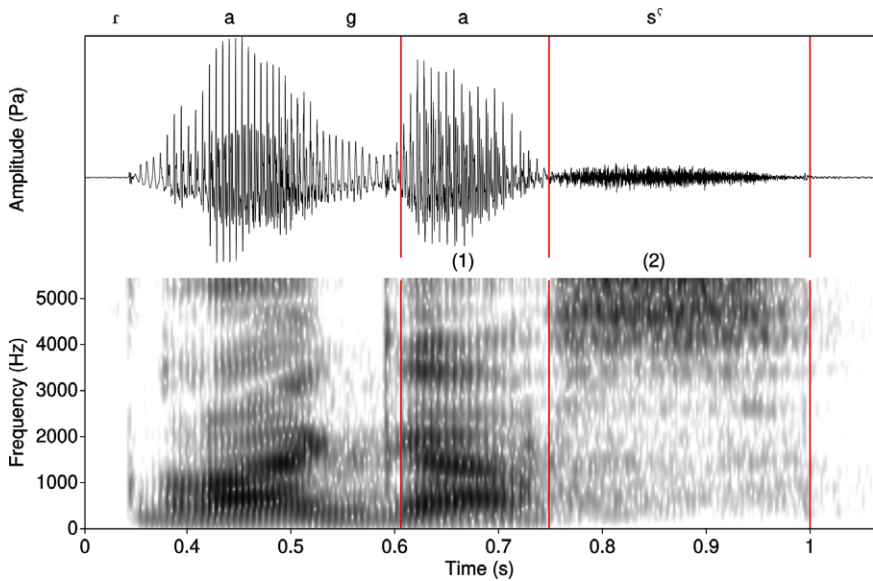




**Figure 1.** Example of the acoustic annotation of target segments for /tafal/ in word-initial position (1) represents the VOT interval for the target consonant /t/ and (2) represents the interval of the following target vowel /a/.



**Figure 2.** Example of the acoustic annotation of target segments for /batʰal/ in word-medial position (1) represents the interval of the preceding target vowel /a/, (2) represents the interval of the closure duration for the target consonant /tʰ/, (3) represents the VOT interval, and (4) represents the interval of the following target vowel /a/.



**Figure 3.** Example of the acoustic annotation of target segments for /ragasʰ/ in word-final position (1) represents the interval of the preceding target vowel /a/ and (2) represents the interval of the target consonant /sʰ/.

chose a smaller range (5 ms) to increase the chances of finding VOT differences associated with the emphatic-plain contrast. For vowels, measurements were considered in agreement when they differed by <15 ms for adults and 20 ms for children. Percentage agreement for adults and children was as follows: for onset bursts, percentage agreement was 98% for adults and 100% for children. For vowel onsets, percentage agreement was 94% for adults and 98% for children. For vowel offsets, percentage agreement was 95% for adults and 99% for children.

**Formant measurements.** Formant trajectories were first automatically estimated, then hand-corrected using an interactive Praat script (Szalay *et al.*, 2022) that visualises the trajectories overlaid on wide-band spectrograms and enables researchers to identify and manually correct misalignments in the formant trajectories. To minimise any potential effect of the adjacent non-target consonants, formant values were averaged over the initial 35% of the vowel following the target consonant in WI and WM positions, as well as the final 35% of the vowel preceding the target consonant in WM and WF positions.

**Statistical analyses.** All visualisations and analyses were conducted in R statistical software (R Core Team, 2021) using the following packages: *lme4* (Bates, Mächler, *et al.*, 2015) and *lmerTest* (Kuznetsova *et al.*, 2017) for linear mixed-effect models (LMMs), *sjPlot* to generate tables (Ludecke *et al.*, 2015), and *ggplot2* (Wilkinson, 2011) to generate plots.

Six LMMs were fitted separately in R for adults and children with either logged VOT (one adult model and one child model) or the F2:F1 ratio (two adult models and two child models) as the dependent variable. In the vowel models, the ratio of the first two formants (F2:F1) was taken as the dependent variable (rather than the difference [F2 – F1], as Mashaqba *et al.* [2022] used). This is considered a normalisation technique because

formant ratios are less sensitive to developmental changes in the size of the vocal tract and other physiological properties that often vary with age and sex, making it more reliable for reducing interspeaker variability (Sapir et al., 2010).

All models included fixed main effects for Consonant type (Emphatic [E] =  $-0.5$ , Plain [P] =  $+0.5$ ) and Gender (Male [M] =  $-0.5$ , Female [F] =  $+0.5$ ). For the children's model, we also included two orthogonal polynomial terms for Age (in months, created using R's *poly()* function), linear and quadratic, as visualisation of the children's data showed curvilinear trends with age. One or two further fixed main effects were included in the analyses, depending on the exact data subset (as detailed below), and all interactions between all main effects were included.

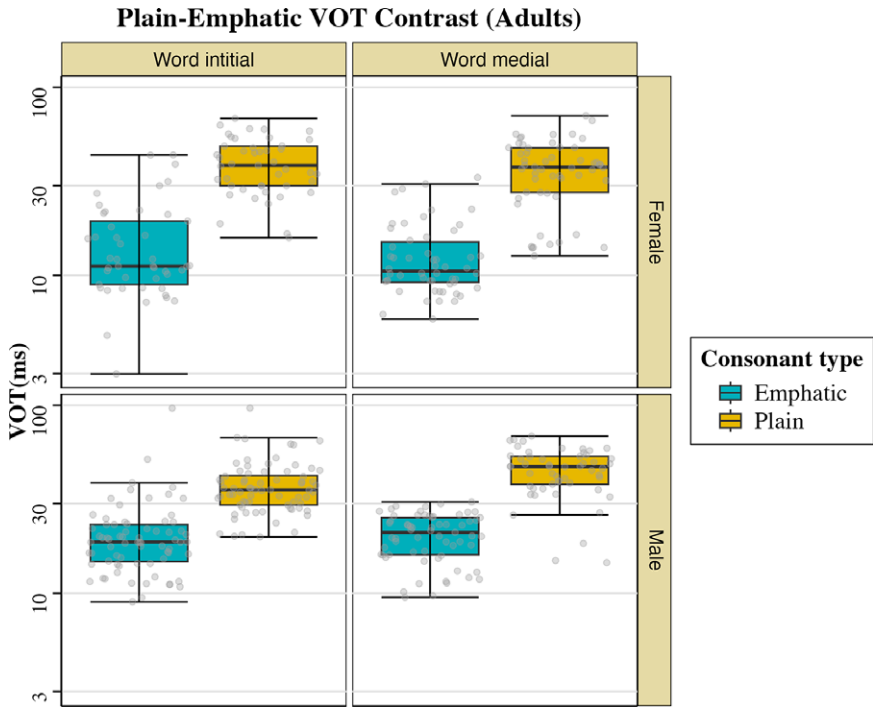
The consonant analysis, on VOT, only included stops not in the WF position. The additional fixed main effect was thus Word Position (WM =  $-0.5$  and WI =  $+0.5$ ).

The first vowel analysis was conducted on WI and WF tokens, including both stop and fricative consonants for WI positions but only fricative consonants for WF positions. The additional fixed main effect for these analyses was termed WpMoA, a combination of Word position and Manner of articulation variables, coded with a sliding-difference contrast with two coefficients. The first coefficient WF Fricatives (WFFs) estimates to what extent the F2:F1 ratio is higher in WI Fricatives (WIFs) than WFFs (estimate =  $\text{WIF} - \text{WFF}$ ), and the second coefficient WI Stops (WISs) estimates to what extent the F2:F1 ratio is higher in WISs than WIFs (estimate =  $\text{WIS} - \text{WIF}$ ). This factor has three levels ( $\text{WIS} = \text{WIF} - \text{WFF} = +1/3$  and  $\text{WIS} = \text{WIF} = +2/3$ ;  $\text{WIF} = \text{WIF} - \text{WFF} = +1/3$  and  $\text{WIS} = \text{WIF} = -1/3$ ; and  $\text{WFF} = \text{WIF} - \text{WFF} = -2/6$  and  $\text{WIS} = \text{WIF} = -1/3$ ).

The second vowel analysis was conducted on WM tokens with both stop and fricative target consonants. The additional fixed main effects for these analyses were thus MoA (MoA: Fricative [f] =  $-0.5$ ; Stop [s] =  $+0.5$ ) and Vowel (Following the target consonant [foll] =  $-0.5$ ; Preceding the target consonant [pre] =  $+0.5$ ).

For each LMM, the final model's random structure was determined through an iterative procedure of model simplification (Bates, Kliegl, et al., 2015), starting with a maximal model incorporating the exhaustive combination of random effects based on the study's design. Random intercepts were always set for both participant and target word (where "Target word" is defined as a lexical item). By-participant random slopes were included for all within-subject effects (the within-subject main effects and their interactions), and by-item random slopes for within-item effects (the within-item main effects and their interactions if applicable). When the maximal model failed to converge or was singular, indicating an over-parameterisation (Bates, Kliegl, et al., 2015), a systematic simplification process was initiated. Throughout this iterative process, the random slope for "Consonant type" was consistently retained due to its importance to the research questions. The simplification process began with the removal of correlations between random effects. If the model did not converge or was singular, the next step involved eliminating complex interactions and main effects, focusing on interactions first and higher-order interactions before lower-order ones. For multiple potential removals, separate models were fitted for each random slope removal. If multiple models converged and were not singular, the model with the lowest Akaike Information Criterion was selected. If none of those models converged or were singular, each model was further simplified and assessed for convergence.

To address the research questions about the difference (acoustic contrast) between plain and emphatic consonants, only the main effect of Consonant type and its interactions are relevant. Only those effects will thus be reported in the text. The reader is referred to the tables in Supplementary Appendix C for full model results on all effects.



**Figure 4.** Boxplots showing adult VOT durations for the emphatic /tʰ/ and the plain /t/ across Gender and Word position. The line within the box represents the median. The whiskers represent the 1.5 interquartile range (1.5\*IQR). Grey dots are individual data points.

**Table 3.** LMM results for adults' logged VOT

Predictors	Estimates	Std. error	t
Intercept	3.20 ***	0.08	38.78
Consonant type <sub>-E + P</sub>	0.88 ***	0.11	8.00
Consonant type <sub>-E + P</sub> *Word position <sub>-WM + WI</sub>	−0.08	0.20	−0.40
Consonant type <sub>-E + P</sub> *Gender <sub>-M + F</sub>	0.31 *	0.13	2.49
Consonant type <sub>-E + P</sub> *Word position <sub>-WM + WI</sub> *Gender <sub>-M + F</sub>	0.13	0.20	0.66

Note: \*  $p < 0.05$ ; \*\*  $p < 0.01$ ; \*\*\*  $p < 0.001$ .

### 3. Results

#### 3.1. Adult model (VOT)

Results of the Adult VOT model (Figure 4 and Table 3) show a positive significant main effect of Consonant type, suggesting, as expected, that VOT is longer in plain than in emphatic stop consonants. The positive significant interaction between Consonant type and Gender suggests that the VOT difference between the plain and emphatic consonants is larger in females than in males. All the other two-way and higher-order interactions with Consonant type were not significant.

### 3.2. Child model (VOT)

Results of the child VOT model (Figure 5 and Table 4) show a significant positive main effect of Consonant type, suggesting that VOT is longer in plain than in emphatic stop consonants. The positive two-way interaction between Consonant type and Age<sub>Linear</sub> suggests that as age increases, the acoustic contrast increases. However, this increase in the acoustic contrast with age is stronger before than after the sample's mean age (60 months), as suggested by the significant negative interaction between Consonant type and Age<sub>Quadratic</sub>. The significant positive two-way interaction between Consonant type and Gender suggests that females produce a larger acoustic contrast than males. This larger female contrast is increasingly greater in WM than WI position as children are closer to the sample's mean age, as suggested by the negative four-way interaction between Consonant type, Age<sub>Quadratic</sub>, Word position, and Gender.

### 3.3. Adult model (F2:F1 Word-initial and Word-final)

Results of the Adult F2:F1 model for WI and WF tokens (Figure 6 and Table 5) show a positive significant main effect of consonant type, suggesting that F2:F1 is larger in vowels next to a plain than an emphatic consonant. None of the two-way and three-way interactions with Consonant type were significant.

### 3.4. Child Model (F2:F1 Word-initial and Word-final)

Results of the Child F2:F1 model for WI and WF tokens (Figure 7 and Table 6) show a positive significant main effect of Consonant type, suggesting that F2:F1 is larger in vowels next to a plain compared to an emphatic consonant. The significant positive interaction between Consonant type and Age<sub>Linear</sub> suggests that the acoustic contrast

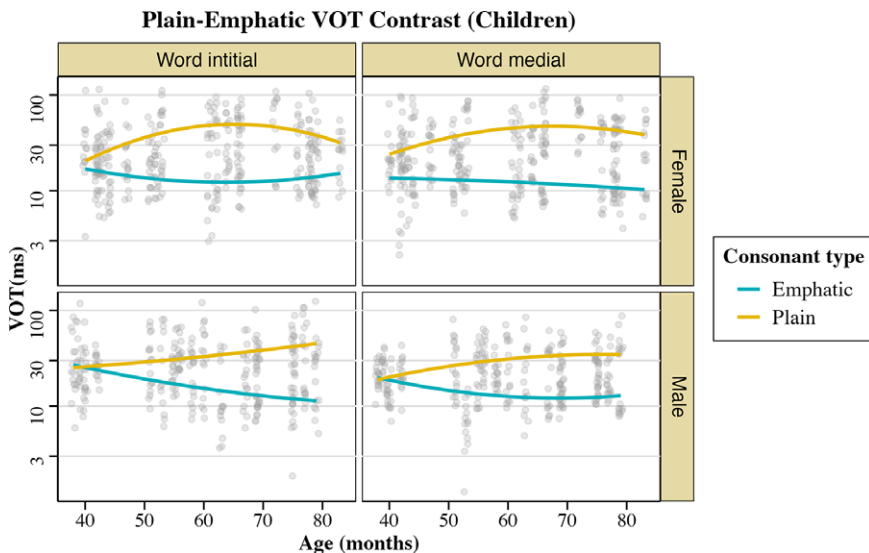


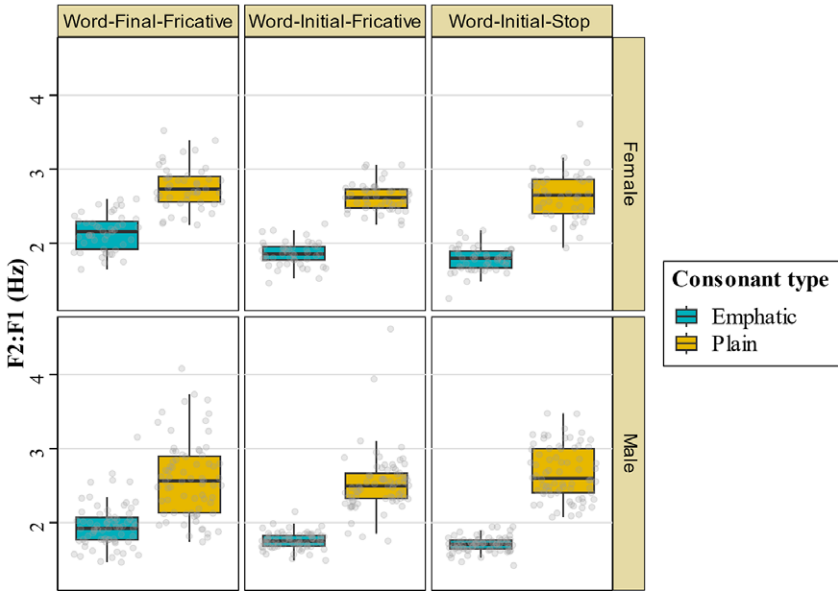
Figure 5. Curvilinear plot showing children's VOT durations of the emphatic /tʰ/ and the plain /t/ across Age, Word position, and Gender. Grey dots are individual data points.

**Table 4.** LMM results of children’s logged VOT

Predictors	Estimates	Std. error	<i>t</i>
Intercept	3.08***	0.05	57.21
Consonant type <sub>-E + P</sub>	0.89***	0.10	8.73
Consonant type <sub>-E + P</sub> * Age <sub>Linear</sub>	11.94***	2.36	5.07
Consonant type <sub>-E + P</sub> * Age <sub>Quadratic</sub>	−5.25*	2.34	−2.25
Consonant type <sub>-E + P</sub> * Word Position <sub>-WM + WI</sub>	−0.10	0.18	−0.56
Consonant type <sub>-E + P</sub> * Gender <sub>-M + F</sub>	0.31*	0.13	2.38
Consonant type <sub>-E + P</sub> * Age <sub>Linear</sub> * Word Position <sub>-WM + WI</sub>	3.30	2.93	1.13
Consonant type <sub>-E + P</sub> * Age <sub>Quadratic</sub> * Word Position <sub>-WM + WI</sub>	−0.69	2.91	−0.24
Consonant type <sub>-E + P</sub> * Age <sub>Linear</sub> * Gender <sub>-M + F</sub>	−3.89	4.66	−0.83
Consonant type <sub>-E + P</sub> * Age <sub>Quadratic</sub> * Gender <sub>-M + F</sub>	−5.10	4.66	−1.09
Consonant type <sub>-E + P</sub> * Word Position <sub>-WM + WI</sub> * Gender <sub>-M + F</sub>	−0.14	0.17	−0.84
Consonant type <sub>-E + P</sub> * Age <sub>Linear</sub> * Word Position <sub>-WM + WI</sub> * Gender <sub>-M + F</sub>	−6.74	5.79	−1.17
Consonant type <sub>-E + P</sub> * Age <sub>Quadratic</sub> * Word Position <sub>-WM + WI</sub> * Gender <sub>-M + F</sub>	−12.25*	5.79	−2.12

Note: \**p* < 0.05; \**p* < 0.01; \*\*\**p* < 0.001.

**Plain-Emphatic F2:F1 in Word-initial and Word-final (Adults)**



**Figure 6.** Boxplots showing adults’ F2:F1 ratios between plain–emphatic consonants in word-initial and word-final positions across MoA and Gender. The line within the box represents the median. The whiskers represent the 1.5 interquartile range (1.5\* IQR). Grey dots are individual data points.

**Table 5.** LMM results for adults' F2:F1 in word-initial and word-final positions

Predictors	Estimates	Std. error	t
Intercept	2.25***	0.06	38.83
Consonant type <sub>E + P</sub>	0.77***	0.10	7.60
Consonant type <sub>E + P</sub> * WpMoA <sub>-WFF + WIF</sub>	0.15	0.24	0.65
Consonant type <sub>E + P</sub> * WpMoA <sub>-WIF + WIS</sub>	0.13	0.24	0.55
Consonant type <sub>E + P</sub> * Gender <sub>-M + F</sub>	-0.05	0.09	-0.55
Consonant type <sub>E + P</sub> * WpMoA <sub>-WFF + WIF</sub> * Gender <sub>-M + F</sub>	-0.01	0.17	-0.04
Consonant type <sub>E + P</sub> * WpMoA <sub>-WIF + WIS</sub> * Gender <sub>-M + F</sub>	-0.09	0.17	-0.54

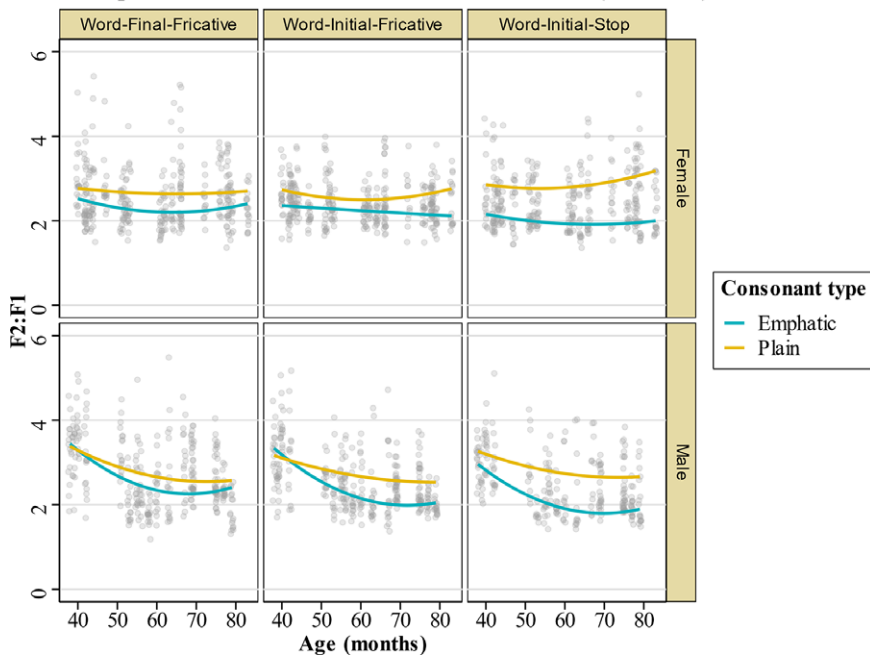
Note: \* $p < 0.05$ ; \*\* $p < 0.01$ ; \*\*\* $p < 0.001$ .

increases with age, with older children showing a larger acoustic contrast relative to younger children. None of the other two-way and higher-order interactions with Consonant type were significant.

### 3.5. Adult model (F2:F1 Word-medial)

Results of the Adult F2:F1 model for WM tokens (Figure 8 and Table 7) show a significant positive main effect of Consonant type, suggesting that F2:F1 is larger in vowels next to a

#### Plain-Emphatic F2:F1 in Word-initial and Word-final (Children)



**Figure 7.** Curvilinear plot showing children's F2:F1 between emphatic and plain consonants in word-initial and word-final positions across Age, MoA, and Gender. Grey dots are individual data points.



**Table 6.** LMM results for children’s F2:F1 in word-initial and word-final positions

Predictors	Estimates	Std. error	<i>t</i>
Intercept	2.51***	0.08	29.79
Consonant type <sub>-E + P</sub>	0.47***	0.14	3.34
Consonant type <sub>-E + P</sub> * Age <sub>Linear</sub>	5.08*	2.15	2.36
Consonant type <sub>-E + P</sub> * Age <sub>Quadratic</sub>	−2.58	2.13	−1.21
Consonant type <sub>-E + P</sub> * WpMoA <sub>WFF + WIF</sub>	0.10	0.32	0.33
Consonant type <sub>-E + P</sub> * WpMoA <sub>WIF + WIS</sub>	0.46	0.36	1.29
Consonant type <sub>-E + P</sub> * Gender <sub>-M + F</sub>	0.13	0.11	1.16
Consonant type <sub>-E + P</sub> * Age <sub>Linear</sub> * WpMoA <sub>WFF + WIF</sub>	3.34	3.04	1.10
Consonant type <sub>-E + P</sub> * Age <sub>Quadratic</sub> * WpMoA <sub>WFF + WIF</sub>	2.86	3.00	0.95
Consonant type <sub>-E + P</sub> * Age <sub>Linear</sub> * WpMoA <sub>WIF + WIS</sub>	−0.56	3.01	−0.19
Consonant type <sub>-E + P</sub> * Age <sub>Quadratic</sub> * WpMoA <sub>WIF + WIS</sub>	−0.57	3.01	−0.19
Consonant type <sub>-E + P</sub> * Age <sub>Linear</sub> * Gender <sub>-M + F</sub>	−2.63	4.30	−0.61
Consonant type <sub>-E + P</sub> * Age <sub>Quadratic</sub> * Gender <sub>-M + F</sub>	5.57	4.26	1.31
Consonant type <sub>-E + P</sub> * WpMoA <sub>WFF + WIF</sub> * Gender <sub>-M + F</sub>	−0.12	0.18	−0.66
Consonant type <sub>-E + P</sub> * WpMoA <sub>WIF + WIS</sub> * Gender <sub>-M + F</sub>	0.17	0.18	0.90
Consonant type <sub>-E + P</sub> * Age <sub>Linear</sub> * WpMoA <sub>WFF + WIF</sub> * Gender <sub>-M + F</sub>	−6.08	6.09	−1.00
Consonant type <sub>-E + P</sub> * Age <sub>Quadratic</sub> * WpMoA <sub>WFF + WIF</sub> * Gender <sub>-M + F</sub>	8.18	6.00	1.36
Consonant type <sub>-E + P</sub> * Age <sub>Linear</sub> * WpMoA <sub>WIF + WIS</sub> * Gender <sub>-M + F</sub>	6.72	6.02	1.12
Consonant type <sub>-E + P</sub> * Age <sub>Quadratic</sub> * WpMoA <sub>WIF + WIS</sub> * Gender <sub>-M + F</sub>	−3.50	6.01	−0.58

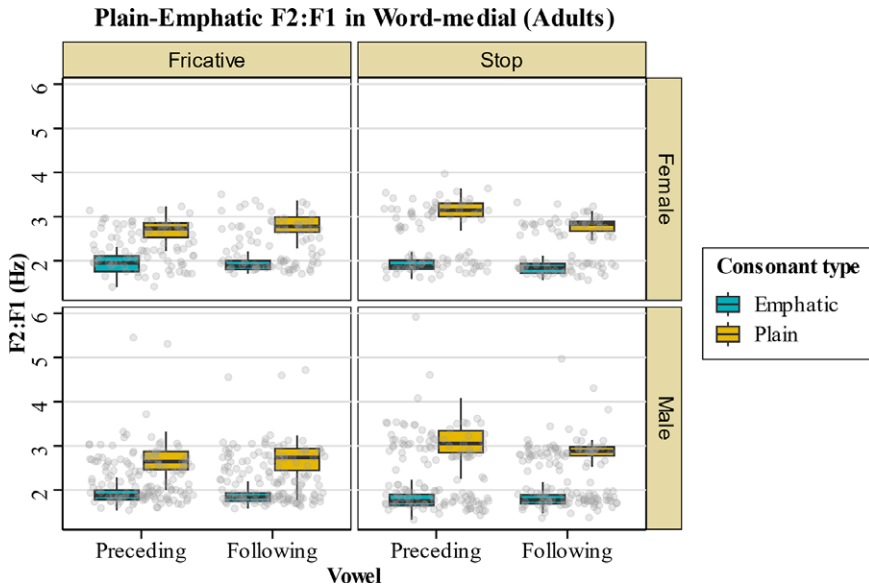
Note: \**p* < 0.05; \*\**p* < 0.01; \*\*\**p* < 0.001.

plain compared to an emphatic consonant. The significant positive two-way interaction between Consonant type and Vowel suggests that the acoustic contrast is larger in preceding compared to following vowels. The significant positive two-way interaction between Consonant type and MoA suggests that the acoustic contrast is larger in vowels next to stops compared to fricatives. The positive three-way interaction between Consonant type, Vowel, and MoA suggests that the extent to which the acoustic contrast is larger in preceding than following vowels is more apparent in stops compared to fricatives. None of the other two-way and higher-order interactions with Consonant type were significant.

### 3.6. Child model (F2:F1 Word-medial)

Results of the Child F2:F1 model for WM tokens model (Figure 9 and Table 8) show a significant positive main effect of Consonant type, suggesting that F2:F1 is larger in vowels next to a plain compared to an emphatic consonant.

The positive two-way interaction between Consonant type and Age<sub>Linear</sub> suggests that the acoustic contrast increases as children grow older. The negative two-way interaction,



**Figure 8.** Boxplots showing adults' F2:F1 between plain and emphatic consonants in word-medial position across MoA, Vowel, and Gender. The line within the box represents the median. The whiskers represent the 1.5 interquartile range (1.5\*IQR). Grey dots are individual data points.

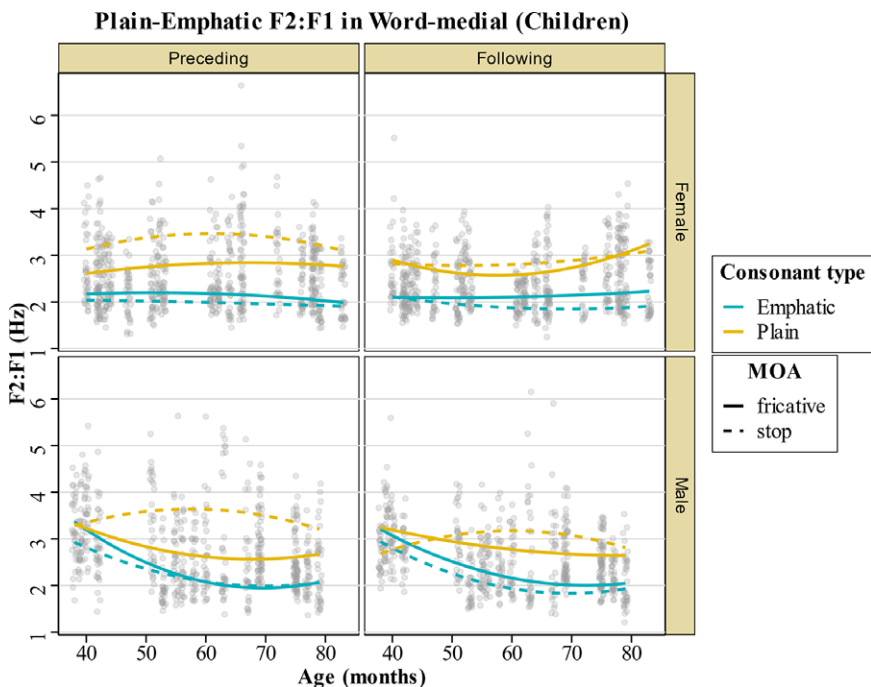
**Table 7.** LMM results for adults' F2:F1 in word-medial position

Predictors	Estimates	Std. error	<i>t</i>
(Intercept)	2.38***	0.04	65.18
Consonant type <sub>E + P</sub>	1.02***	0.08	13.60
Consonant type <sub>E + P</sub> * Vowel <sub>fol + pre</sub>	0.08 *	0.04	2.17
Consonant type <sub>E + P</sub> * MoA <sub>f + s</sub>	0.35***	0.10	3.56
Consonant type <sub>E + P</sub> * Gender <sub>M + F</sub>	-0.09	0.13	-0.70
Consonant type <sub>E + P</sub> * Vowel <sub>fol + pre</sub> * MoA <sub>f + s</sub>	0.30***	0.08	4.05
Consonant type <sub>E + P</sub> * Vowel <sub>fol + pre</sub> * Gender <sub>M + F</sub>	-0.04	0.08	-0.47
Consonant type <sub>E + P</sub> * MoA <sub>f + s</sub> * Gender <sub>M + F</sub>	-0.09	0.12	-0.76
Consonant type <sub>E + P</sub> * Vowel <sub>fol + pre</sub> * MoA <sub>f + s</sub> * Gender <sub>M + F</sub>	0.17	0.15	1.12

Note: \* $p < 0.05$ ; \*\* $p < 0.01$ ; \*\*\* $p < 0.001$ .

however, between Consonant type and Age<sub>Quadratic</sub> indicates that the acoustic contrast is increasing more before than after the sample's mean age (60 months).

The positive two-way interaction between Consonant type and Vowel suggests that the acoustic contrast is larger in preceding than following vowels. The positive two-way interaction between Consonant type and MoA suggests that the acoustic contrast is larger in vowels next to stops compared to fricatives. This larger contrast in vowels next to stops is greater in preceding compared to following vowels – as suggested by the positive three-way interaction between Consonant type, Vowel, and MoA.



**Figure 9.** Curvilinear plot showing children's F2:F1 between emphatic and plain consonants in word-medial position across Age, MoA, Vowel, and Gender. Grey dots are individual data points.

The negative three-way interaction between Consonant type, Age<sub>Quadratic</sub>, and MoA suggests that the extent to which the acoustic contrast is larger in vowels next to stops compared to fricatives reaches its peak around children's mean age. The positive three-way interaction between Consonant type, Age<sub>Quadratic</sub>, and Gender suggests that the extent to which the acoustic contrast is larger in females compared to males is at its minimum around children's mean age. The negative four-way interaction between Consonant type, Age<sub>Linear</sub>, Vowel, and MoA suggests that the extent to which the preceding vowel advantage is larger in vowels next to stops than in fricatives decreases as children grow older.

#### 4. Discussion

This study investigated the production of emphatic consonants in real words by Saudi-Hijazi Arabic-speaking children aged 3–6 years. We hypothesised a gradual acquisition process, with older children exhibiting a larger acoustic contrast between plain and emphatic consonants than younger children. We additionally predicted that the plain–emphatic acoustic contrast would be larger in initial than in final positions and larger for vowels preceding than following the target consonants for WM position. Children were also expected to produce larger acoustic contrasts in stops compared to fricatives. Finally, we predicted that females may produce a larger acoustic contrast compared to males.

The results suggest that children in the 3–6 years age range, on average, produce an acoustic contrast between plain and emphatic consonants that increases with age, with the

**Table 8.** LMM results for children's F2:F1 in word-medial position

Predictors	Estimates	Std. error	t
(Intercept)	2.57***	0.05	53.77
Consonant type <sub>-E + P</sub>	0.81***	0.07	12.05
Consonant type <sub>-E + P</sub> * Age <sub>Linear</sub>	7.63***	2.16	3.54
Consonant type <sub>-E + P</sub> * Age <sub>Quadratic</sub>	-5.20*	2.13	-2.44
Consonant type <sub>-E + P</sub> * Vowel <sub>folll + pre</sub>	0.18***	0.04	4.37
Consonant type <sub>-E + P</sub> * MoA <sub>f + s</sub>	0.53***	0.12	4.34
Consonant type <sub>-E + P</sub> * Gender <sub>M + F</sub>	0.14	0.09	1.49
Consonant type <sub>-E + P</sub> * Age <sub>Linear</sub> * Vowel <sub>folll + pre</sub>	-1.22	2.09	-0.59
Consonant type <sub>-E + P</sub> * Age <sub>Quadratic</sub> * Vowel <sub>folll + pre</sub>	-2.62	2.07	-1.26
Consonant type <sub>-E + P</sub> * Age <sub>Linear</sub> * MoA <sub>f + s</sub>	3.05	3.23	0.94
Consonant type <sub>-E + P</sub> * Age <sub>Quadratic</sub> * MoA <sub>f + s</sub>	-8.58**	3.20	-2.68
Consonant type <sub>-E + P</sub> * Vowel <sub>folll + pre</sub> * MoA <sub>f + s</sub>	0.43***	0.07	6.10
Consonant type <sub>-E + P</sub> * Age <sub>Linear</sub> * Gender <sub>M + F</sub>	-6.28	4.31	-1.46
Consonant type <sub>-E + P</sub> * Age <sub>Quadratic</sub> * Gender <sub>M + F</sub>	10.87*	4.26	2.55
Consonant type <sub>-E + P</sub> * Vowel <sub>folll + pre</sub> * Gender <sub>-M + F</sub>	0.03	0.08	0.31
Consonant type <sub>-E + P</sub> * MoA <sub>f + s</sub> * Gender <sub>-M + F</sub>	-0.06	0.15	-0.41
Consonant type <sub>-E + P</sub> * Age <sub>Linear</sub> * Vowel <sub>folll + pre</sub> * MoA <sub>f + s</sub>	-7.33*	3.63	-2.02
Consonant type <sub>-E + P</sub> * Age <sub>Quadratic</sub> * Vowel <sub>folll + pre</sub> * MoA <sub>f + s</sub>	1.28	3.61	0.35
Consonant type <sub>-E + P</sub> * Age <sub>Linear</sub> * Vowel <sub>folll + pre</sub> * Gender <sub>-M + F</sub>	2.00	4.17	0.48
Consonant type <sub>-E + P</sub> * Age <sub>Quadratic</sub> * Vowel <sub>folll + pre</sub> * Gender <sub>-M + F</sub>	-7.97	4.15	-1.92
Consonant type <sub>-E + P</sub> * Age <sub>Linear</sub> * MoA <sub>f + s</sub> * Gender <sub>-M + F</sub>	-2.55	6.47	-0.39
Consonant type <sub>-E + P</sub> * Age <sub>Quadratic</sub> * MoA <sub>f + s</sub> * Gender <sub>-M + F</sub>	4.54	6.41	0.71
Consonant type <sub>-E + P</sub> * Vowel <sub>folll + pre</sub> * MoA <sub>f + s</sub> * Gender <sub>-M + F</sub>	0.01	0.14	0.08
Consonant type <sub>-E + P</sub> * Age <sub>Linear</sub> * Vowel <sub>folll + pre</sub> * MoA <sub>f + s</sub> * Gender <sub>-M + F</sub>	-0.86	7.26	-0.12
Consonant type <sub>-E + P</sub> * Age <sub>Quadratic</sub> * Vowel <sub>folll + pre</sub> * MoA <sub>f + s</sub> * Gender <sub>-M + F</sub>	0.92	7.23	0.13

Note: \* $p < 0.05$ ; \*\* $p < 0.01$ ; \*\*\* $p < 0.001$ .

rate of increase tapering off as children grow older. Notably, the acoustic contrast was detected on both the consonant (stop VOTs) and the immediately adjacent vowels (formant frequency ratios) in similar patterns as for adults, suggesting that children are targeting the emphatic consonants, at least for emphatic stops, rather than implementing a different strategy – for example, producing only a back vowel.<sup>5</sup>

<sup>5</sup>We would like to acknowledge the researchers at the International Congress of Phonetic Sciences (ICPhS) 2023 who, upon hearing children's speech samples, made the observation that children may not be producing the emphatic consonant but rather only producing a vocalic difference.

The increase in acoustic contrast with age is consistent with Mashaqba *et al.* (2022), who noted that although Ammani-speaking children produced plain–emphatic acoustic contrasts by age 5 years, adult-like production emerged later, around 6–7 years of age. The age-related increase in the contrasts in our study similarly suggests an ongoing refinement of speech production abilities as children grow older (Sadagopan & Smith, 2008; Walsh & Smith, 2002) and may reflect increased lingual control as children mature (Zharkova *et al.*, 2014). Given the articulatory complexity associated with emphatics and cross-linguistic findings of protracted development of complex speech segments, such as English liquids (Howson & Redford, 2021), we hypothesise that the development of these segments may continue after 6;11 – that is, the age of the oldest children tested in this study.

The gradual development of the plain–emphatic contrast may explain why some previous studies relying on perceptual judgements reported “no acquisition” of emphatics until over 6 years old (Amayreh, 2003; Amayreh & Dyson, 1998; Dyson & Amayreh, 2000). Listeners in those studies may have identified children’s still-developing emphatic realisations as not yet meeting their threshold for an emphatic percept. In contrast, the first author of the present study, a native Saudi-Hijazi Arabic speaker, perceives most of the youngest children in the present sample to be producing the target emphatics. This apparent discrepancy between studies poses some questions about children’s emphatic production and adults’ perception thereof. For example, were children in previous studies, from different dialects, producing weaker acoustic contrasts than the children in our sample? Which auditory cues do adult native speakers across dialects rely on when evaluating children’s production of emphatics? Are there inherent biases or variations in perceptual thresholds that might skew interpretations of whether children are indeed producing emphatics accurately? These questions highlight the need for future research to combine instrumental measures with perceptual evaluations, ensuring a comprehensive understanding of the development of emphatics.

#### 4.1. *The acoustic contrast and linguistic factors*

The current study also explored the effects of the manner of articulation and the target segment word position on the production of emphatics. Consistent with our hypotheses, both children and adults produced larger acoustic contrasts in vowels next to stops compared to fricatives. As for positional effects, there was no evidence for or against a difference between children’s (or adults’) productions in WIFs and WFFs. For WM position, both children and adults produced larger plain–emphatic acoustic contrasts in vowels immediately preceding compared to immediately following the target consonant.

The most parsimonious explanation for this similarity between adults’ and children’s patterns is that children are aligning their production with the adult input. In what follows, we discuss the relevant articulation mechanisms, explaining how these might give rise to each of the observed linguistic effects, even in adults, and might affect children in particular. While the present data do not permit the disentanglement of input factors from articulation development, we aim to embed the observed linguistic and developmental patterns for emphatics into the general literature on phonological and articulatory development.

Regarding the manner of articulation effect, recall that the larger acoustic contrast in vowels next to stops compared to fricatives was observed in both adults and children, aligning with previous findings on other dialects of Arabic (Ghazeli, 1977; Jongman *et al.*,

2011). One possible explanation for this general difference between emphatic stops and emphatic fricatives is the differential degree of tongue involvement. Fricative production requires a precise, narrow constriction to allow for continuous airflow that generates turbulence at the place of constriction. Maintaining this aerodynamic requirement might limit the extent to which the tongue can move, for example, assume a secondary pharyngeal constriction, without disrupting the fricative noise. In contrast, stops require a complete blockage of airflow, allowing the tongue greater freedom to move to different positions before and after the closure without influencing sound production. This flexibility might result in more pronounced spectral differences for emphatic stops compared to fricatives (Jongman et al., 2011).

In the context of language development, the controlled constriction required for fricatives might account for the universal tendencies for children to acquire fricatives later than stops (Gildersleeve-Neumann et al., 2000; Jakobson, 1968; Stoel-Gammon, 1985). The challenge of producing an emphatic fricative is further increased due to the additional complexity of maintaining the secondary gesture. Most studies agree that emphatic fricatives are acquired after emphatic stops (Alqattan, 2015; Ammar & Morsi, 2018; Ayyad, 2011; Morsi, 2003). In the current data, the fact that the rate of increase in the size of the acoustic contrast slows down more in stops compared to fricatives may suggest that the plain–emphatic contrast stabilises earlier in stops than in fricatives, which is in line with previous findings showing that fricative coarticulation development is more protracted than stops (Zharkova, 2018).

Regarding the difference between vowels before and after target consonants, the analysis of WM position shows that both children and adults produced larger plain–emphatic acoustic contrasts in vowels immediately preceding compared to immediately following the target consonant. This aligns with previous findings on emphatics, which show stronger anticipatory compared to carryover coarticulation in adults (e.g., Al-Masri & Jongman, 2004; Jongman et al., 2011; Zawaydeh, 1999). This can be explained through the different mechanisms claimed to give rise to anticipatory and carryover coarticulation, which may follow different developmental trajectories. Anticipatory coarticulation, where speakers prepare for the production of upcoming sounds, has been hypothesised to reflect planning strategies that need to be learned (Hertrich & Ackermann, 1995; Recasens, 1984; Waldstein & Baum, 1991). In contrast, carryover coarticulation is believed to be more reflexive, arising from the physical properties and limitations of the articulators, often called *mechanical inertia constraints* (Baum & Waldstein, 1991; Daniloff & Hammarberg, 1973; MacNeilage & DeClerk, 1969; Recasens, 1984).

Whether these positional effects in children's production might reflect aspects of articulatory development depends on one's theory of coarticulation development. On the one hand, given that children's speech production is still in development, children's larger acoustic contrasts in vowels preceding (rather than following) emphatics might reflect that the anticipatory coarticulation patterns are learned before the maturation of mechanical inertia constraints responsible for carryover coarticulation. Children might thus be learning planning strategies before the mature motor control required for carryover coarticulation. Recent research, on the other hand, claims that both anticipatory and carryover coarticulation might share a common mechanism, and thus, developmental patterns should be similar (Noiray et al., 2018; Rubertus & Noiray, 2020). This view is also compatible with our findings, given that children might use this single mechanism to target the adult pattern of stronger anticipatory than carryover coarticulation. Moreover, the stronger anticipatory emphasis in the adult target might make anticipatory cues to

emphatics more perceptually salient and, therefore, a more prominent target for children's production. The present data, thus, do not permit the disentanglement of different coarticulatory accounts from each other and input factors. Our ongoing work investigates the acquisition of emphasis spread, where the effects of the emphatic consonant extend bidirectionally beyond the immediately adjacent vowel, which could provide more insight into coarticulation development in children and the relative impact of input versus articulation mechanisms.

Regarding word position effects, previous findings indicate that Ammani-speaking children achieve adult-like production of emphatics earlier in WI and WM than in WF positions (Amayreh & Dyson, 1998; Mashaqba *et al.*, 2022), reflecting general cross-linguistic development patterns in consonant acquisition (Levelt *et al.*, 2000; Levelt & de Vijver, 2004). The current study, however, provides no evidence for (or against) a difference between WISs and WIFs nor between WIFs and WFFs in children's (or adults') productions. This is particularly surprising as WI (but not WF) target consonants were in a stressed syllable, which would be expected to make the contrast in WI position more prominent and, thus, further increase children's contrast in this position (Beckman, 1998; Echols & Newport, 1992; Rose, 2002). The apparent lack of a word position effect might be due to the opposite effects of coarticulation versus word position and stress. Because WI contrasts reflect carryover coarticulation with the following vowel, while WF contrasts reflect anticipatory coarticulation with the preceding vowel, coarticulation effects in isolation would lead to a larger contrast in WF compared to WI position, similar to the pattern observed in WM position. However, the prominence of the WI contrast, being in a stressed syllable, might have mitigated the effects of coarticulatory direction, or vice versa. Future work may explore this possibility by examining how prosodic effects, such as stress, interact with the coarticulatory development of the emphatic consonant.

#### 4.2 *The acoustic contrast and gender*

While previous findings suggest that, among adults, males exhibit stronger cues to emphatics than females (Abudalbh, 2010; Lehn, 1963; Wahba, 1993), other studies reported the opposite, with females producing stronger cues to emphatics than males (Kahn, 1975; Al-Masri & Jongman, 2004). Our adult results are in partial agreement with the latter studies, as we observed that the VOT contrast in stops was larger for females than males. However, no evidence for or against gender-based differences in formant frequency ratios was found. Thus, we cannot claim that females consistently produce stronger cues to emphatics than males. The relatively small sample size (females = 5 and males = 8) might have left us unable to reliably estimate differences between female and male speakers in Hijazi Arabic.

In contrast, female children displayed larger acoustic contrasts than male children for both VOT and vowel formants, with the gender effect for vowel formants most apparent at the upper and lower limits of the age range studied. Because sex-related differences in the vocal tract may not emerge until after puberty (Barbier *et al.*, 2015; Vorperian *et al.*, 2011) and because the vowel formant measure used in this study (F2:F1 ratio) serves as a normalised metric that reduces interspeaker variability, including that related to age and sex (Sapir *et al.*, 2010), the observed differences in the child data are unlikely to reflect anatomical differences. Taken together, the sex-related trends in our data with children under 7 years of age may instead be interpreted as learning of gendered speech patterns



(Ford et al., 2018; Foulkes et al., 2005; Munson et al., 2015). In the future, more targeted studies into this issue could collect more comprehensive information about adults' and children's social background, as factors such as economic status, education level, and cultural norms may moderate the extent to which gender is associated with acoustic cues to emphatics (Abudalbh, 2010; Al Malwi, 2017; Al-Masri & Jongman, 2004; Alzoubi, 2017; Khattab et al., 2006).

#### 4.3 Limitations and future directions

In the current study, we acoustically examined the consonantal and vocalic cues in children's production of the plain–emphatic contrast. The acoustic measurements of the vocalic cues included only 35% of the vowel immediately adjacent to the target consonant, to minimise influences from adjacent non-target consonants. This is a potential limitation, as the analysis window does not allow us to examine the extent to which emphasis is maintained across the entire vowel immediately adjacent to the target emphatic and whether it spreads to vowels further away (e.g., Zawaydeh & de Jong, 2011). Moreover, our specific choice of words allowed us to examine the effect of emphasis only in the context of the low front vowel /a/, which has been previously shown to be more influenced by emphatic consonants than other vowels (e.g., Jongman et al., 2011). Our ongoing work will explore whether the current findings can be generalised across other vowel contexts.

While we examined both consonantal and vocalic cues to emphatics for stops, we only considered vocalic cues to emphatics for fricatives. This decision was motivated by (1) the lack of evidence in previous literature for a spectral difference between plain and emphatic fricatives in adults (Al-Khairy, 2005; Jongman et al., 2011) and (2) the unsuitability of our recordings to perform reliable fine-grained spectral analyses. Although speech was recorded via a professional voice recorder and every effort was taken to improve the acoustic conditions in each recording location, some mild background noise was also captured during elicitation, which may impact spectral measures.<sup>6</sup> While the quality of the recording may also impact formant measurements, these are typically more robust, and our hand correction further ensures their accuracy. Future work, with recordings made in better sound-attenuated environments, may explore whether children exhibit a spectral difference between plain and emphatic consonants, using measures such as spectral mean and spectral tilt (Al-Khairy, 2005; Al-Tamimi, 2015; Jongman et al., 2011). Understanding these spectral differences could provide more insights into the development of speech production in children, shedding light on how they acquire and refine articulatory patterns.

We also chose not to assess “adult-likeness” in the present study because determining adult-like patterns based on null-hypothesis significance testing may be misleading. While a significant *p*-value can indicate a difference between adults' and children's productions, a non-significant *p*-value does not necessarily mean they are identical and that the children's productions are, therefore, “adult-like.” Future research may incorporate both qualitative and quantitative techniques, perhaps alongside longitudinal data

<sup>6</sup>This point was based on consultation with an expert phonetician who observed that it is not possible to filter out noise, particularly high-frequency noise, without filtering out parts of the speech spectrum, which can significantly impact spectral analyses of fricatives, for example, the centre of gravity (COG).

collection, to offer a more comprehensive perspective on “adult-likeness” and its relevance in child speech development.

A final limitation to our interpretation of these data is due to the absence of corpus-based frequency data or a comprehensive, searchable dictionary for Saudi-Hijazi Arabic. As such, we cannot investigate or comment on the effects of frequency of lexical items or segmental frequency and functional load – all factors that are known to influence acquisition (Ambridge *et al.*, 2015; Pye *et al.*, 1987; Stokes & Surendran, 2005; Surendran & Niyogi, 2003). Future research is needed to investigate how these factors impact the accuracy and phonetic contrast with which children produce emphatic consonants.

While the focus of the present study was on the acoustic realisation of emphatic contrasts, the non-target-like productions observed during data annotation raise questions for future research. Among 3- and 4-year-olds, emphatic fricatives were occasionally realised with an inserted [t], producing forms such as [statan] for /s<sup>ʕ</sup>adam/ and [gafas<sup>ʕ</sup>] for /gafas<sup>ʕ</sup>/. Final consonant omission was also observed, as in /gafas<sup>ʕ</sup>/ realised as [gafa], suggesting possible articulatory simplification strategies. One 5-year-old produced [qafas<sup>ʕ</sup>] for /gafas<sup>ʕ</sup>/, which may reflect emphasis overspread affecting the place of articulation of /g/ or a style shift towards the MSA pronunciation. In Hijazi Arabic, /q/ is typically realised as [g]; the child’s production of [q], therefore, is possibly triggered by increased exposure to MSA forms in formal settings, for example, schools. A similar observation was reported for Palestinian Arabic children and adolescents, who showed diglossic style shifting towards the Standard Arabic [q] in a school-like picture-naming task (Shetewi *et al.*, 2024). These patterns were not analysed systematically (and were excluded from the current analysis), as they fall outside the acoustic focus of the current study. However, they raise important questions for future research – particularly about the link between phonetic patterns and phonological representations in a diglossic context.

As a final direction for future research, determining whether children are attempting the secondary constriction gesture when producing emphatics, as expected from an adult speaker, requires a closer inspection of the tongue. Future articulatory work may be necessary to confirm the articulatory gestures underlying the current acoustic results and provide a more comprehensive understanding of the developmental trajectory of these sounds.

## 5. Conclusion

This study is the first to assess the acquisition of emphatics by typically developing Saudi-Hijazi Arabic-speaking children between 3 and 6 years old. It is the first, to the best of the authors’ knowledge, to instrumentally examine both consonantal and vocalic cues to the plain–emphatic contrast. Our results show that children produce both consonantal and vocalic cues to the contrast and progressively enhance these contrasts over time. Additionally, the overall patterns observed in children’s productions closely align with those observed in adults, suggesting the potential impact of adult input on children’s production of emphatics.

Although this study was not designed with applied outcomes as its primary aim, the findings have potential relevance for speech-language therapists and educators working with Arabic-speaking children. The emergence of an acoustic contrast by age 3 years suggests that some children may be developing articulatory targets earlier than perceptual

judgements alone would suggest (e.g., Amayreh, 2003; Amayreh & Dyson, 1998). Awareness of the potential of covert contrast (Macken & Barton, 1980; Scobbie, 1998) may help clinicians in accurately assessing children's abilities and challenges.

**Supplementary material.** The supplementary material for this article can be found at <http://doi.org/10.1017/S0305000925100214>.

**Competing interest.** The authors declare none.

## References

- Abdoh, E. M. A. (2011). *A study of the phonological structure and representation of first words in Arabic* [Doctoral dissertation]. University of Leicester.
- Abudalbh, M. (2010). *Effects of gender on the production of emphasis in Jordanian Arabic: A sociophonetic study* [Doctoral dissertation]. University of Kansas.
- Al Malwi, I. M. (2017). *The effects of gender and age on voice onset time by Abah Arabic speakers* [Doctoral dissertation]. Long Beach: California State University.
- Alajroush, N. A. (2019). *Phonological development in typically developing Najdi Arabic-speaking children aged 1–4 years* [Doctoral dissertation]. Newcastle University.
- Al-Ani, S. H. (1970). *Arabic phonology: An acoustical and physiological investigation* (Vol. 61). Walter de Gruyter.
- Al-Awaji, N. N. (2014). *Speech production in Arabic speaking children with operated cleft palate* [Doctoral dissertation]. University of Sheffield.
- Al-Buainain, H., Shain, K., Al-Timimy, F., & Khattab, G. (2012). Baseline data for Arabic acquisition with clinical applications: Some phonological processes in Qatari children's speech. *International Journal of Business and Social Research (IJBSR)*, 2. <https://doi.org/10.18533/IJBSR.V2I6.131>.
- Algryani, A. (2014). Emphasis spread in Libyan Arabic. *International Journal on Studies in English Language and Literature (IJSELL)*, 2(12), 30–38. [www.arcjournals.org](http://www.arcjournals.org)
- Alhammad, R. (2014). *Emphasis spread in Najdi Arabic* [Doctoral dissertation]. Fresno: California State University.
- Aljenaie, K. (2010). Verbal inflection in the acquisition of Kuwaiti Arabic. *Journal of Child Language*, 37 (4), 841–863. Cambridge University Press. <https://doi.org/10.1017/S0305000909990031>.
- Al-Khairi, M. A. (2005). *Acoustic characteristics of Arabic fricatives* [Doctoral dissertation]. University of Florida.
- Al-Masri, M., & Jongman, A. (2004). Acoustic correlates of emphasis in Jordanian Arabic: Preliminary results. In A. Agwuele, W. Warren, & S.-H. Park (Eds.), *Proceedings of the 2003 Texas Linguistics society conference* (pp. 96–106). Cascadia Proceedings Project.
- Almbark, R. (2008). *A sociophonetic study of emphasis in Syrian Arabic* [Unpublished MA thesis]. The University of York.
- Almuhaimed, A. A. S. (2021). *Emphasis spread in the Najdi Arabic dialect* [Doctoral dissertation]. University of Manchester.
- Alqattan, S. (2015). *Early phonological acquisition by Kuwaiti Arabic children* [Doctoral dissertation]. Newcastle University.
- Al-Tamimi, F., Alzoubi, F., & Tarawnah, R. (2009). A Videofluoroscopic study of the emphatic consonants in Jordanian Arabic. *Folia Phoniatrica et Logopaedica*, 61 (4), 247–253. <https://doi.org/10.1159/000235644>.
- Al-Tamimi, F., & Heselwood, B. (2011). Nasoendoscopic, videofluoroscopic and acoustic study of plain and emphatic coronals in Jordanian Arabic. In B. Heselwood & Z. M. Hassan (Eds.), *Instrumental studies in Arabic phonetics* (Vol. 319, pp. 165–191). John Benjamins. <https://doi.org/10.1075/cilt.319>.
- Al-Tamimi, J. (2015). Spectral tilt as an acoustic correlate to Pharyngealisation in Jordanian and Moroccan Arabic. In *The Scottish consortium for ICPHS 2015 (Ed.), proceedings of the 18th international congress of phonetic sciences*. The University of Glasgow.
- Al-Tamimi, J. (2017). Revisiting acoustic correlates of pharyngealization in Jordanian and Moroccan Arabic: Implications for formal representations. *Laboratory Phonology*, 8 (1), 28. <https://doi.org/10.5334/lab-phon.19>.
- Al-Tamimi, J., Schiel, F., Khattab, G., Sokhey, N., Amazouz, D., Dallak, A., & Moussa, H. (2022). A Romanization system and WebMAUS aligner for Arabic varieties. In *Proceedings of the 13th conference on language resources and evaluation (LREC 2022)* (pp. 7269–7276). European Language Resources Association (ELRA), Marseille, France.

- Al-Thubaity, A. O. (2015). A 700M+ Arabic corpus: KACST Arabic corpus design and construction. *Language Resources and Evaluation*, **49** (3), 721–751. <https://doi.org/10.1007/s10579-014-9284-1>.
- Alzaidi, M. S., Xu, Y., & Xu, A. (2019). Prosodic encoding of focus in Hijazi Arabic. *Speech Communication*, **106**, 127–149. <https://doi.org/10.1016/j.specom.2018.12.006>.
- Alzoubi, A. A. S. (2017). *The effect of social factors on emphatic-plain contrast in Jordan: A Sociophonetic study of Arabic in Amman City* [Doctoral dissertation]. The University of Utah.
- Amayreh, M. M. (2003). Completion of the consonant inventory of Arabic. *Journal of Speech, Language, and Hearing Research*, **46** (3), 517–529. [https://doi.org/10.1044/1092-4388\(2003\)042](https://doi.org/10.1044/1092-4388(2003)042).
- Amayreh, M. M., & Dyson, A. T. (1998). The acquisition of Arabic consonants. *Journal of Speech, Language, and Hearing Research*, **41** (3), 642–653. <https://doi.org/10.1044/jslhr.4103.642>.
- Ambridge, B. E. N., Kidd, E., Rowland, C. F., & Theakston, A. L. (2015). The ubiquity of frequency effects in first language acquisition. *Journal of Child Language*, **42** (2), 239–273. <https://doi.org/10.1017/S030500091400049X>.
- Ammar, W., & Morsi, R. (2018). Phonological development and disorders: Colloquial Egyptian Arabic. *Phonological Development and Disorders in Children*, 204–232; Multilingual Matters. <https://doi.org/10.21832/9781853598906-011>.
- Ayyad, H. S. (2011). *Phonological development of typically developing Kuwaiti Arabic-speaking preschoolers* [Doctoral dissertation]. <https://open.library.ubc.ca/collections/24/items/1.0103320>
- Barbier, G., Boë, L.-J., Captier, G., & Laboissière, R. (2015, September). Human vocal tract growth: A longitudinal study of the development of various anatomical structures. *Interspeech 2015 -16th Annual Conference of the International Speech Communication Association, International Speech Communication Association*.
- Barton, D., & Macken, M. A. (1980). An instrumental analysis of the voicing contrast in word-initial stops in the speech of four-year-old English-speaking children. *Language and Speech*, **23** (2), 159–169. <https://doi.org/10.1177/002383098002300203>.
- Bates, D., Kliegl, R., Vasishth, S., & Baayen, H. (2015). *Parsimonious mixed models*. arXiv (Cornell University). <https://doi.org/10.48550/arxiv.1506.04967>.
- Bates, D., Mächler, M., Bolker, B. M., & Walker, S. C. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, **67** (1), 1–48. <https://doi.org/10.18637/jss.v067.i01>.
- Baum, S. R., & Waldstein, R. S. (1991). Perseveratory Coarticulation in the speech of profoundly hearing-impaired and normally hearing children. *Journal of Speech and Hearing Research*, **34** (6), 1286–1292. <https://doi.org/10.1044/jshr.3406.1286>.
- Beckman, J. N. (1998). *Positional faithfulness* [Doctoral dissertation]. University of Massachusetts.
- Bin-Muqbil, M. S. (2006). *Phonetic and phonological aspects of Arabic emphatics and gutturals* [Doctoral dissertation, The University of Wisconsin-Madison]. <https://www.proquest.com/openview/e2caaf10616f46602568ecc23ad47bbe/1?pq-origsite=gscholar&cbl=18750&diss=y>
- Boersma, P., & Weenink, D. (2023). *Praat: Doing phonetics by computer* [Computer program] (6.3.10). <http://www.praat.org/>
- Bukshaisha, F. A. M. (1986). *An experimental phonetic study of some aspects of Qatari Arabic* [Doctoral dissertation, University of Edinburgh]. <http://hdl.handle.net/1842/26361>
- Card, E. A. (1983). *A phonetic and phonological study of Arabic emphasis* [Doctoral dissertation, Cornell University]. <https://www.proquest.com/openview/48be86af87495ac91ef382c9ca2cc70/1?pq-origsite=gscholar&cbl=18750&diss=y>
- Contreras Kallens, P., Elmlinger, S., Wang, K., Goldstein, M., McLeod, S., Crowe, K., & Christiansen, M. (2023). Statistical learning or phonological universals? Ambient language statistics guide consonant acquisition in four languages. *Proceedings of the Annual Meeting of the Cognitive Science Society*, **45**(45).
- Daniloff, R. G., & Hammarberg, R. E. (1973). On defining coarticulation. *Journal of Phonetics*, **1** (3), 239–248. [https://doi.org/10.1016/S0095-4470\(19\)31388-9](https://doi.org/10.1016/S0095-4470(19)31388-9).
- Davis, K. (1994). Stop voicing in Hindi. *Journal of Phonetics*, **22** (2), 177–193. [https://doi.org/10.1016/S0095-4470\(19\)30192-5](https://doi.org/10.1016/S0095-4470(19)30192-5).
- Davis, S. (1995). Emphasis spread in Arabic and grounded phonology. *Linguistic Inquiry*, **26**(3), 465–498.
- Dyson, A. T., & Amayreh, M. M. (2000). Phonological errors and sound changes in Arabic-speaking children. *Clinical Linguistics & Phonetics*, **14** (2), 79–109. <https://doi.org/10.1080/026992000298850>.

- Echols, C. H., & Newport, E. L. (1992). The role of stress and position in determining first words. *Language Acquisition*, 2 (3), 189–220. [https://doi.org/10.1207/s15327817la0203\\_1](https://doi.org/10.1207/s15327817la0203_1).
- Fabiano-Smith, L., & Bunta, F. (2012). Voice onset time of voiceless bilabial and velar stops in 3-year-old bilingual children and their age-matched monolingual peers. *Clinical Linguistics & Phonetics*, 26 (2), 148–163. <https://doi.org/10.3109/02699206.2011.595526>.
- Ford, C., Tabain, M., & Docherty, G. (2018). Gender differences in the spectral characteristics of voiceless sibilants produced by Australian English-speaking children. In *Proc.17th Australasian international conference on speech science and technology* (pp. 97–100). Australasian Speech Science and Technology Association (ASSTA), Sydney, Australia.
- Foulkes, P., Docherty, G., & Watt, D. (2005). Phonological variation in child-directed speech. *Language (Baltimore)*, 81 (1), 177–206. <https://doi.org/10.1353/lan.2005.0018>.
- Ghazeli, S. (1977). *Back consonants and backing coarticulation in Arabic* [Doctoral dissertation]. <https://www.proquest.com/openview/fd8f59a741fe953b19990b6efe638eef/1?pq-origsite=gscholar&cbl=18750&diss=y>
- Gildersleeve-Neumann, C. E., Davis, B. L., & Macneilage, P. F. (2000). Contingencies governing the production of fricatives, affricates, and liquids in babbling. *Applied PsychoLinguistics*, 21 (3), 341–363. <https://doi.org/10.1017/S0142716400003039>.
- Hamdan, J. M., & Amayreh, M. M. (2007). Consonant profile of Arabic-speaking school-age children in Jordan. *Folia Phoniatrica et Logopaedica*, 59 (2), 55–64. <https://doi.org/10.1159/000098338>.
- Hassan, Z. M., & Esling, J. H. (2011). Investigating the emphatic feature in Iraqi Arabic. In *Instrumental studies in Arabic phonetics* (pp. 217–234). John Benjamins Publishing Company.
- Hertrich, I., & Ackermann, H. (1995). Coarticulation in slow speech: Durational and spectral analysis. *Language and Speech*, 38 (2), 159–187. <https://doi.org/10.1177/002383099503800203>.
- Howson, P. J., & Redford, M. A. (2021). The Acquisition of Articulatory Timing for liquids: Evidence from child and adult speech. *Journal of Speech, Language, and Hearing Research*, 64 (3), 734–753. [https://doi.org/10.1044/2020\\_JSLHR-20-00391](https://doi.org/10.1044/2020_JSLHR-20-00391).
- Ingram, D., & Babatsouli, E. (2024). Cross-linguistic phonological acquisition. In *The Handbook of Clinical Linguistics* (2nd edn). <https://doi.org/10.1002/9781119875949.ch29>
- Israel, A., Proctor, M., Goldstein, L., Iskarous, K., & Narayanan, S. (2012). Emphatic segments and emphasis spread in Lebanese Arabic: A real-time magnetic resonance imaging study. *Proc. Int'l Conf. on Speech Communication and Technology*, 2178–2181. <https://doi.org/10.21437/Interspeech.2012-581>.
- Jakobson, R. (1968). *Child language, aphasia and phonological universals*. De Gruyter Mouton. <https://doi.org/10.1515/9783111353562>.
- Jongman, A., Herd, W., Al-Masri, M., Sereno, J., & Combest, S. (2011). Acoustics and perception of emphasis in urban Jordanian Arabic. *Journal of Phonetics*, 39 (1), 85–95. <https://doi.org/10.1016/j.wocn.2010.11.007>.
- Kahn, M. (1975). Arabic emphatics: The evidence for cultural determinants of phonetic sex-typing. *Phonetica*, 31 (1), 38–50. <https://doi.org/10.1159/000259648>.
- Kalaldehy, R., & Al-Shdaifat, A. A. (2019). The influence of emphatic /d<sup>ɛ</sup>/ on modern standard Arabic vowels: An acoustic analysis. *Lingua Posnaniensis*, 61 (1), 43–61. <https://doi.org/10.2478/linpo-2019-0003>.
- Khattab, G., Al-Tamimi, F., & Heselwood, B. (2006). Acoustic and auditory differences in the /t/–/T/ opposition in male and female speakers of Jordanian Arabic. In *Perspectives on Arabic Linguistics XVI: Papers from the Sixteenth Annual Symposium on Arabic Linguistics* (pp. 131–160). John Benjamins Publishing Company, based in Amsterdam & Philadelphia.
- Kisler, T., Reichel, U., & Schiel, F. (2017). Multilingual processing of speech via web services. *Computer Speech & Language*, 45, 326–347. <https://doi.org/10.1016/j.csl.2017.01.005>.
- Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, 82 (13), 1–26. <https://doi.org/10.18637/jss.v082.i13>.
- Ladefoged, P., & Maddieson, I. (1996). *The sounds of the world's languages*. Blackwell Publishers.
- Laufer, A., & Baer, T. (1988). The emphatic and pharyngeal sounds in Hebrew and in Arabic. *Language and Speech*, 31 (2), 181–205. <https://doi.org/10.1177/002383098803100205>.
- Lehn, W. (1963). Emphasis in Cairo Arabic. *Language (Baltimore)*, 39 (1), 29–39. <https://doi.org/10.2307/410760>.
- Levelt, C. C., & de Vijver, R. (2004). Syllable types in cross-linguistic and developmental grammars. In *Constraints in phonological acquisition* (pp. 204–218). Cambridge University Press.



- Levelt, C. C., Schiller, N. O., & Levelt, W. J. (2000). The acquisition of Syllable types. *Language Acquisition*, 8(3), 237–264. [https://doi.org/10.1207/S15327817LA0803\\_2](https://doi.org/10.1207/S15327817LA0803_2).
- Ludecke, D., Bartel, A., Schwemmer, C., Powell, C., Djalovski, A., & Titz, J. (2015). *Data visualization for statistics in social science. Comprehensive R Archive Network (CRAN)*.
- Macken, M. A., & Barton, D. (1980). The acquisition of the voicing contrast in English: A study of voice onset time in word-initial stop consonants. *Journal of Child Language*, 7 (1), 41–74. <https://doi.org/10.1017/S0305000900007029>.
- MacNeilage, P. F., & DeClerk, J. L. (1969). On the motor control of coarticulation in CVC monosyllables. *The Journal of the Acoustical Society of America*, 45 (5), 1217–1233. <https://doi.org/10.1121/1.1911593>.
- Mashaqba, B., Daoud, A., Zuraq, W., & Huneety, A. (2022). Acquisition of emphatic consonants by Ammani Arabic-speaking children. *Language Acquisition*, 29 (4), 441–456. <https://doi.org/10.1080/10489223.2022.2049600>.
- Morsi, R. (2003). Phonological acquisition of normal Egyptian children from the age of two and a half to five years. In *The 15th international congress of phonetic science, Barcelona, Spain* (pp. 3–9). The Proceedings of the 15th International Congress of Phonetic Sciences (ICPhS), Universitat Autònoma de Barcelona.
- Munson, B., Crocker, L., Pierrehumbert, J. B., Owen-Anderson, A., & Zucker, K. J. (2015). Gender typicality in children's speech: A comparison of boys with and without gender identity disorder. *The Journal of the Acoustical Society of America*, 137 (4), 1995–2003. <https://doi.org/10.1121/1.4916202>.
- 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. [https://doi.org/10.1044/2018\\_JSLHR-S-17-0148](https://doi.org/10.1044/2018_JSLHR-S-17-0148).
- Nwokah, E. E. (1986). Consonantal substitution patterns in Igbo phonological acquisition. *Language and Speech*, 29 (2), 159–176. <https://doi.org/10.1177/002383098602900204>.
- Omar, M. K. (1973). *The Acquisition of Egyptian Arabic as a Native Language*. Berlin & Boston: De Gruyter Mouton. <https://doi.org/10.1515/9783110819335>.
- Owaida, H. (2015). *Speech sound acquisition and phonological error patterns in child speakers of Syrian Arabic: A normative study* [Doctoral dissertation]. City University London.
- Pye, C., Ingram, D., & List, H. (1987). A comparison of initial consonant acquisition in English and Quiché. In *Children's language* (1st ed., pp. 175–190). Psychology Press.
- R Core Team. (2021). *A language and environment for statistical computing* (4.3.1). R foundation for statistical computing. <https://www.R-project.org/>.
- Recasens, D. (1984). V -to-C coarticulation in Catalan VCV sequences: An articulatory and acoustical study. *Journal of Phonetics*, 12 (1), 61–73. [https://doi.org/10.1016/S0095-4470\(19\)30851-4](https://doi.org/10.1016/S0095-4470(19)30851-4).
- Rose, Y. (2002). Relations between segmental and prosodic structure in first language acquisition. *Annual Review of Language Acquisition*, 2 (1), 117–155. <https://doi.org/10.1075/arla.2.06ros>.
- Rubertus, E., & Noiray, A. (2020). Vocalic activation width decreases across childhood: Evidence from carryover coarticulation. *Laboratory Phonology*, 11(1), 7. <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. [https://doi.org/10.1044/1092-4388\(2008\)06-0222](https://doi.org/10.1044/1092-4388(2008)06-0222).
- Saleh, M., Shoeib, R., Hegazi, M., & Ali, P. (2007). Early phonological development in Arabic Egyptian children: 12–30 months. *Folia Phoniatrica et Logopaedica*, 59(5), 234–240. <https://doi.org/10.1159/000104461>.
- Sander, E. K. (1972). When are speech sounds learned? *Journal of Speech and Hearing Disorders*, 37(1), 55–63.
- Sapir, S., Ramig, L. O., Spielman, J. L., & Fox, C. (2010). Formant centralization ratio: A proposal for a new acoustic measure of Dysarthric speech. *Journal of Speech, Language, and Hearing Research*, 53(1), 114–125. [https://doi.org/10.1044/1092-4388\(2009\)08-0184](https://doi.org/10.1044/1092-4388(2009)08-0184).
- Schiel, F. (1999). Automatic phonetic transcription of non-prompted speech. In *Proceedings of ICPhS, San Francisco* (pp. 607–610). Proceedings of ICPhS-14, Department of Phonetics, University of Munich.
- Scobbie, J. M. (1998). Interactions between the acquisition of phonetics and phonology. *Papers from the 34th Annual Regional Meeting of the Chicago Linguistic Society*.
- Shetewi, O., Corrigan, K. P., & Khattab, G. (2024). Children's multilectal repertoires: Diglossic style-shifting by Palestinian children and adolescents in Syria. *Language*, 9 (11). <https://doi.org/10.3390/languages9110341>.
- Shriberg, L. D. (1993). Four new speech and prosody-voice measures for genetics research and other studies in developmental phonological disorders. *Journal of Speech and Hearing Research*, 36 (1), 105–140. <https://doi.org/10.1044/jshr.3601.105>.

- Stoehr, A., Benders, T., van Hell, J. G., & Fikkert, P. (2017). Second language attainment and first language attrition: The case of VOT in immersed Dutch–German late bilinguals. *Second Language Research*, 33(4), 483–518. <https://doi.org/10.1177/0267658317704261>.
- Stoel-Gammon, C. (1985). Phonetic inventories, 15–24 months: A longitudinal study. *Journal of Speech and Hearing Research*, 28(4), 505–512. <https://doi.org/10.1044/jshr.2804.505>.
- Stokes, S. F., & Surendran, D. (2005). Articulatory complexity, ambient frequency, and functional load as predictors of consonant development in children. *Journal of Speech, Language, and Hearing Research*, 48(3), 577–591. [https://doi.org/10.1044/1092-4388\(2005/040](https://doi.org/10.1044/1092-4388(2005/040)
- Sučková, M. (2020). First language attrition in voice onset times in Anglophone expatriates residing in the Czech Republic. *Brno studies in English*, 2, 47–66. <https://doi.org/10.5817/BSE2020-2-3>.
- Surendran, D., & Niyogi, P. (2003). *Measuring the functional load of phonological contrasts*. arXiv (Cornell University). <https://doi.org/10.48550/arxiv.cs/0311036>.
- Szalay, T., Benders, T., Cox, F., & Proctor, M. (2022). Vowel merger in Australian English lateral-final rimes: /æɔ-æ/. In R. Billington (Ed.), *Proceedings of the 18th Australasian international conference on speech science and technology* (pp. 106–110). Australasian Speech Science and Technology Association (ASSTA).
- Tar, E. (2014). The acquisition of the voicing contrast in word-initial bilabial and alveolar stops - atypical data from Hungarian. *Clinical Linguistics & Phonetics*, 28 (4), 269–282. <https://doi.org/10.3109/02699206.2013.835445>.
- Turki, D. F. (2022). *Phonological development of typically developing monolingual Saudi Hejazi Arabic-speaking children aged 2;6;5;11* [Doctoral dissertation], University of Sheffield]. <https://etheses.whiterose.ac.uk/32903/>
- Vorperian, H. K., Wang, S., Schimek, E. M., Durtschi, R. B., Kent, R. D., Gentry, L. R., & Chung, M. K. (2011). Developmental sexual dimorphism of the Oral and pharyngeal portions of the vocal tract: An imaging study. *Journal of Speech, Language, and Hearing Research*, 54 (4), 995–1010. [https://doi.org/10.1044/1092-4388\(2010/10-0097](https://doi.org/10.1044/1092-4388(2010/10-0097)
- Wahba, K. (1993). *A sociolinguistic treatment of the feature of emphasis in Egypt* [Doctoral dissertation]. University of Texas at Austin.
- Waldstein, R. S., & Baum, S. R. (1991). Anticipatory Coarticulation in the speech of profoundly hearing-impaired and normally hearing children. *Journal of Speech and Hearing Research*, 34 (6), 1276–1285. <https://doi.org/10.1044/jshr.3406.1276>.
- Walsh, B., & Smith, A. (2002). Articulatory movements in adolescents: Evidence for protracted development of speech motor control processes. *Journal of Speech, Language, and Hearing Research*, 45 (6), 1119–1133. [https://doi.org/10.1044/1092-4388\(2002/090](https://doi.org/10.1044/1092-4388(2002/090)
- Watson, J. C. E. (2002). *The phonology and morphology of Arabic*. Oxford University Press.
- Wilkinson, L. (2011). ggplot2: Elegant graphics for data analysis by WICKHAM, H. *Biometrics*, 67 (2), 678–679. <https://doi.org/10.1111/j.1541-0420.2011.01616.x>.
- Yang, J. (2018). Development of stop consonants in three- to six-year-old mandarin-speaking children. *Journal of Child Language*, 45 (5), 1091–1115. <https://doi.org/10.1017/S0305000918000090>.
- Younes, M. A. (1993). Emphasis spread in two Arabic dialects. In M. Eid & C. Holes (Eds.), *Perspectives on ARABIC LINGUISTICS v: Papers from the fifth ANNUAL SYMPOSIUM on ARABIC LINGUISTICS* (Vol. 1993, pp. 119–145). John Benjamins Publishing Company.
- Zawaydeh, B. A. (1999). *The phonetics and phonology of gutturals in Arabic* [Doctoral dissertation]. Indiana University.
- Zawaydeh, B. A., & de Jong, K. (2011). The phonetics of localising uvularisation in Ammani-Jordanian Arabic. In *Instrumental studies in Arabic phonetics* (p. 257). John Benjamins Publishing.
- Zharkova, N. (2018). An ultrasound study of the development of lingual Coarticulation during childhood. *Phonetica*, 75 (3), 245–271. <https://doi.org/10.1159/000485802>.
- 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. [https://doi.org/10.1044/2014\\_JSLHR-S-11-0350](https://doi.org/10.1044/2014_JSLHR-S-11-0350).