Journal of Child Language (2021), **48**, 1262–1280
doi:10.1017/S0305000920000781

CAMBRIDGE
UNIVERSITY PRESS

ARTICLE

Acoustic cues to coda stop voicing contrasts in Australian English-speaking children

Julien MILLASSEAU[1]*, Ivan YUEN[1], Laurence BRUGGEMAN[1,2], and Katherine DEMUTH[1]

[1]Department of Linguistics, Macquarie University, Australia 16 University Avenue, Australian Hearing Hub, North Ryde, NSW 2109 Australia and [2]The MARCS Institute & ARC Centre of Excellence for the Dynamics of Language, Western Sydney University, Australia
*Address for correspondence: Julien Millasseau, Department of Linguistics, Macquarie University, 16 University Avenue, Australian Hearing Hub, North Ryde, NSW 2109 Australia E-mail julien.millasseau@mq.edu.au

(Received 25 May 2020; revised 20 November 2020; accepted 23 November 2020; first published online 10 February 2021)

Abstract
While voicing contrasts in word-onset position are acquired relatively early, much less is known about how and when they are acquired in word-coda position, where accurate production of these contrasts is also critical for distinguishing words (e.g., *dog* vs. *dock*). This study examined how the acoustic cues to coda voicing contrasts are realized in the speech of 4-year-old Australian English-speaking children. The results showed that children used similar acoustic cues to those of adults, including longer vowel duration and more frequent voice bar for voiced stops, and longer closure and burst durations for voiceless stops along with more frequent irregular pitch periods. This suggests that 4-year-olds have acquired productive use of the acoustic cues to coda voicing contrasts, though implementations are not yet fully adult-like. The findings have implications for understanding the development of phonological contrasts in populations for whom these may be challenging, such as children with hearing loss.

Keywords: Coda stop voicing; Australian English; Children; Adults

Introduction

Children first learn to produce voicing contrasts in word-onset position, eventually learning how to produce these in word-coda position as well (Jakobson, 1968; Stoel-Gammon and Buder, 1999). But due to their syllabic complexity, coda consonants often take some time to acquire (Demuth et al., 2006). This raises questions about WHEN children learn to produce the acoustic cues needed to convey voicing contrasts in coda position, critical for distinguishing words in English such as *dog* vs. *dock*. Quantifying these cues in typically-developing children is also necessary for assessing language-delayed child populations where these contrasts may present long-term challenges (e.g., Baudonck et al., 2011; Markides, 1970; Miles et al., 2012; Xu Rattanasone and Demuth, 2014).

Most of our knowledge about child phonological acquisition comes from perceptual coding, where adults listen and transcribe children's speech phonemically. These studies have shown that children's early word productions vary substantially from those of adults, and that phonological contrasts, such as voicing, may not be systematically conveyed (e.g., Smith, 1979, 1973). Transcriptions of child speech can thus be difficult to interpret with respect to actual phonological knowledge because children sometimes produce systematic acoustic distinctions that adult listeners cannot hear (Li et al., 2009; Macken and Barton, 1980; Scobbie et al., 2000). These 'covert contrasts' have been documented for phonological contrasts such as voicing in the speech of both typically-developing children (Macken & Barton, 1980; Weismer et al., 1981) and those with language delays (Forrest et al., 1990). Systematic acoustic analysis is therefore needed for a better understanding of children's developing phonological knowledge (Munson et al., 2005; Theodore et al., 2012).

Relatively little is known about the acquisition of acoustic cues to coda voicing contrasts, as acoustic investigations of the development of voicing contrasts have mostly focused on word-initial consonants and voice onset time (VOT). Between 2 and 4 years of age, children go through several stages of development with respect to these voicing contrasts (Imbrie, 2005; Kewley-Port & Preston, 1974; Koenig, 2001; Macken & Barton, 1980; Zlatin & Koenigsknecht, 1976). Around the age of 2, children first go through a period where their VOTs for voiced and voiceless stops are not distinct from one another, both falling within the adult range of voiced stops. During the second stage, their VOTs for voiced and voiceless stops systematically differ from one another but are still perceived as the same category (i.e., voiced) by adults. Finally, around the age of 3–4, children's VOTs begin to approximate adult acoustic values, with short VOTs for voiced stops and long VOTs for voiceless stops, though children's VOTs tend to be longer than those of adults. Thus, it takes several years for children to acquire adult-like voicing contrasts in onset position (e.g., Yu et al., 2015; Zlatin and Koenigsknecht, 1976). Since coda consonants are typically acquired later than onsets, we might expect children to take even longer to reach adult-like acoustic realizations for voicing contrasts in coda position.

Coda consonants are often omitted or variably realized in children's early speech (e.g., Demuth et al., 2006; Kirk and Demuth, 2006). Despite being later acquired, children seem to exhibit a coda consonant representation at an early age, even when the target coda itself is not actually produced. For instance, in spontaneous speech, 2-year-olds lengthen the duration of the vowel when the coda consonant is missing (e.g., */dɔ**g**/ → /d**ɔː**/;* Song and Demuth, 2008). This raises questions about the acoustic cues children use in producing coda voicing contrasts, and how these become more adult-like over time.

There are multiple acoustic cues to oral stop coda voicing contrasts in English. The primary cue is vowel duration, where vowels preceding voiced codas are longer than those preceding voiceless codas (Fowler, 1992; de Jong, 2004; Penney et al., 2018; Raphael, 1972). Closure duration (e.g., Cox and Palethorpe, 2011; Lisker, 1957; Luce and Charles-Luce, 1985; Penney et al., 2018) and burst duration (Song et al., 2012) also vary as a function of coda voicing, with longer durations for voiceless than for voiced codas. Additional cues, such as the presence of a voice bar (VB) during closure preceding a voiced coda, and the presence of irregular pitch periods (IPP) at the end of the vowel preceding a voiceless coda, provide supplementary information to coda voicing (British English: Docherty and Foulkes, 1999; American English: Redi and Shattuck-Hufnagel, 2001; Australian English: Penney et al., 2018). IPP

https://doi.org/10.1017/S0305000920000781 Published online by Cambridge University Press

appears to be a strong correlate to coda voicelessness in Australian English (Penney et al., 2018), while it is less systematic in American English (Redi & Shattuck-Hufnagel, 2001; Song et al., 2012).

While many of the cues to coda voicing contrasts have been investigated in adult speech, relatively little is known about these cues in child speech, though some suggest early sensitivity. Acoustic analysis of the longitudinal spontaneous speech of four 1–3-year-olds from the Providence Corpus (Demuth et al., 2006) revealed that these American English-speaking children produced longer vowel durations before voiced than voiceless codas (Ko, 2007). Using the same corpus, Song et al. (2012) compared the speech of three mother-child dyads and found that 1;6-2;6-year-olds were already using most of the acoustic correlates to coda voicing contrasts (i.e., vowel duration, burst duration, IPP and VB). For example, they produced longer vowel durations before voiced than voiceless codas, without any developmental trend. They also started to produce longer burst durations for voiceless stops than for voiced stops around the age of 2;0, matching the adult pattern. By 2;6 years, burst durations for both voicing categories decreased in duration, approximating adult values. These children did not differ in their use of IPP as a function of coda voicing, whereas their mothers tended to produce more IPP before voiceless than voiced codas. Finally, both children and mothers showed similar patterns for voice bar (VB), with more use before voiced codas than voiceless codas.

In another, elicited speech study of two American children aged 2;5 and 3;2 years, Shattuck-Hufnagel et al. (2011) also found that children produced more IPP and post-release noise (bursts) when the coda was voiceless (though Song et al., 2012 had found no difference in slightly younger children). This suggests that the use of IPP as a cue to coda voicelessness in American English may emerge around 3 years, with children still refining control of the larynx and the degree of vocal fold tension until age 3;6 years (Imbrie, 2005). Finally, children as young as 2 years systematically produced more instances of VB before voiced than before voiceless codas (Shattuck-Hufnagel et al., 2011; Song et al., 2012). This suggests that the use of VB as a cue to voicing is in place relatively early.

The studies above suggest that children under 4 years of age can make coda voicing contrasts, but that their acoustic realizations remain less systematic and more variable than those of adults. Most of the above studies also contained data from only a few children, and examined only some places of articulation (PoA). To determine how children's use of cues to coda voicing contrasts comes to approximate that of adults, a larger, more systematic investigation of coda voicing contrasts at all PoAs is needed.

The present study therefore set out to examine the acoustic correlates to coda voicing contrasts in Australian English-speaking children aged 4–5 years and compared these with those of adults. Using an elicited imitation task with systematically controlled stimuli at all PoAs, the present study thus aimed to provide a normative Australian English baseline/reference of the acoustics of coda voicing contrasts in both children and adults that could then be used for future work examining early L2 learners and language-delayed populations.

Based on previous findings from both children and adults, we hypothesized that 4-year-olds would use systematic durational differences to coda voicing contrasts. These might include longer vowel durations before voiced codas, and longer closure and burst durations for voiceless codas. We also expected that children would produce overall longer durations than adults, with more inter- and intra-speaker variation. Given that IPP is a strong correlate to coda voicelessness in Australian

Table I. List of CVC stimuli.

	/b/	/d/	/g/	/p/	/t/	/k/
/ɪ/	bib	kid	pig	tip	pit	tick
/ɐ/	tub	bud	bug	cup	gut	duck
/ɔ/	cob	god	dog	top	pot	dock

English (Penney et al., 2018), we also predicted that children might produce higher rates of IPP before voiceless codas than before voiced codas, though perhaps lower than adults. Finally, we hypothesized that children would produce more instances of VB before voiced than before voiceless codas, though again, perhaps less often than adults.

Methods

Participants

A total of 20 pre-schoolers (aged 4;1-5;8 years, *M* = 4;10; 12 females, 8 males) were recruited, along with twenty adult controls (aged 20–35 years, *M* = 28; 15 females, 5 males). All participants were monolingual speakers of Australian English, born in Australia and brought up in Sydney. No participants reported any speech, hearing or cognitive difficulties. The study was approved by the Macquarie University's Human Ethics Panel. Children received a $20 voucher and stickers for their participation; adults received course credit.

Stimuli

A total of 18 CVC picturable minimal pair words (see Table I) were selected by crossing word-final voicing (voiced vs. voiceless), PoA (bilabials vs. alveolars vs. velars) and three short–lax vowels (i.e. /ɪ/, / ɐ / and / ɔ /; Cox and Palethorpe, 2007). All stimuli were high-frequency words, with a mean Zipf frequency of 4.5 in the Subtlex-UK CBeebies pre-schooler corpus (van Heuven et al., 2014). This is a corpus of subtitles taken from the BBC channel CBeebies, which is aimed at pre-school-aged children. Target words were embedded utterance-finally in the sentence "See this XXX". All sentences were recorded by a 25-year-old female native speaker of Australian English in a sound-attenuated room (sampling rate: 44.1Khz with 16-bit quantization). Three additional sentences with non-target CVC words were recorded to serve as practice trials. To make the task engaging for children, all recorded stimuli were then paired with a cartoon-like picture and presented as an interactive game on an iPad Air using the Keynote presentation software.

Procedure

All participants were recorded in a sound-attenuated room at Macquarie University, Sydney. Participants sat at a table in front of the iPad and 30 cm away from an AKG C535 EB microphone. The microphone was connected to a pre-amplifier (Sound Devices, USBPre2) and recordings were captured and encoded with Audacity (mono WAV files: 44.1 kHz sampling rate, 16-bit quantization). Pictures and paired audio sentences were presented one at a time on the iPad, starting with the three practice

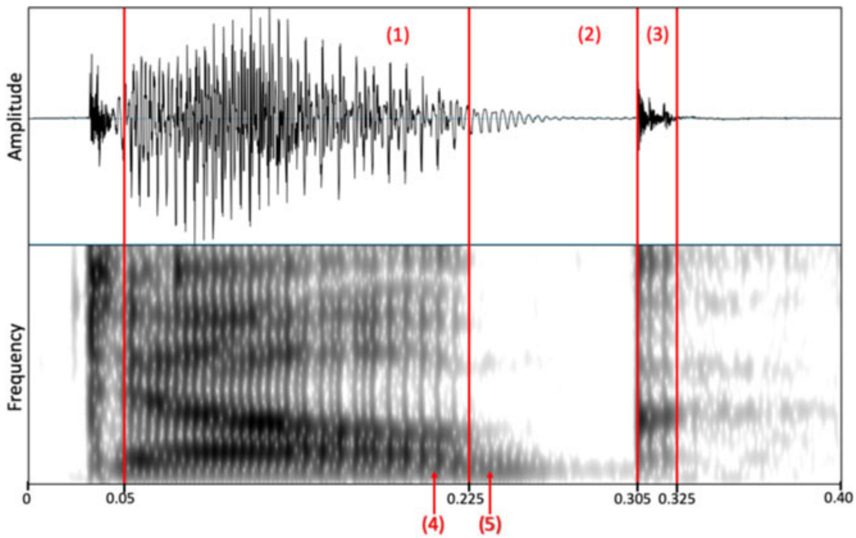

Figure 1. Representative waveform and spectrogram of the word "dog" as produced by a child in the carrier sentence "See this dog". (1) corresponds to vowel duration, (2) closure duration, (3) burst duration, (4) irregular pitch periods and (5) voice bar.

trials. When participants touched the screen, the audio file linked to the picture was played. Participants repeated the sentence, touched the iPad again to move on to the next picture, tapped on the screen to hear the next sentence, and so forth until all the stimuli had been repeated. All participants completed a total of five blocks, with all stimuli presented once in each block, in pseudo-randomized order. Each participant thus produced 90 tokens (18 target words x 5 repetitions), completing the task in 30 minutes for children and 10 minutes for adults.

Acoustic coding and analysis

A total of 3600 recorded tokens (1800 from children, 1800 from adults) were inspected and manually annotated by the first author using Praat (Boersma & Weenink, 2019). Five acoustic cues (see Figure 1) were annotated as follows: (1) vowel duration was measured from the beginning to the end of a strong F2, (2) closure duration was measured from the end of a strong F2 to the first peak of the release burst of the following stop, (3) burst duration was measured from the first peak of the release burst to the end of strong energy on the spectrogram, (4) the presence of irregular pitch periods (IPP) was identified by the presence of irregularly spaced glottal pulses at the end of the vowel, (5) the presence of voice bar (VB) was identified by the presence of a low frequency, low amplitude signal following the sudden drop of amplitude at the end of the vowel. Thus, vowel, closure and burst durations were durational cues, while the presence or absence of IPP and VB were binary cues. A phonetically trained research assistant independently coded 10% of the child and adult data ($n = 360$) to check reliability. Pearson correlations were high for each of the five cues (vowel duration: $r = 0.93$, $p < .001$; closure duration: $r = 0.90$, $p < .001$; burst duration: $r = 0.88$, $p < .001$; presence of IPP: $r = 0.91$, $p < .001$ and presence of VB: $r = 0.91$, $p < .001$).

https://doi.org/10.1017/S0305000920000781 Published online by Cambridge University Press

Table II. Number of unreleased codas per PoA for both participant groups.

	Children	Adults
Bilabial	78	46
Alveolar	39	56
Velar	18	16

Table III. Number of outliers (proportion) that were removed, by durational measure and participant group.

	Children	Adults
Vowel duration	82 (4.6%)	76 (4.2%)
Closure duration	58 (3.2%)	48 (2.7%)
Burst duration	104 (5.7%)	147 (8.2%)

Statistical analysis

Some of the produced tokens were not released (especially bilabials; see Table II), which meant that burst and closure durations could not be measured. Only vowel duration, IPP and VB were therefore measured for those tokens. For all durational cues, tokens falling beyond two standard deviations from the mean were excluded as outliers (see Table III). The proportion of outliers per cue was similar for both participant groups. For children, a total of 1718 tokens were analysed for vowel duration, 1607 for closure duration and 1561 for burst duration. For adults, a total of 1724 tokens were analysed for vowel duration, 1752 for closure duration and 1653 for burst duration. For both groups, all tokens for IPP and VB were analysed, as these cues were binary, so outlier removal was not applicable.

Separate linear mixed-effects models were fitted with the *lme4* package (Bates et al., 2015) in R (R Core Team, 2013) to each of the three durational measures. For the binary cues separate generalized mixed-effects models were fitted using the same R package. Each model had the same fixed structure, which included all main effects of and interactions between Group (Children vs. Adults), Voicing (Voiced vs. Voiceless) and PoAs (Bilabial vs. Alveolar vs. Velar). The random structure included by-subject and by-item intercepts, as well as by-subject random slopes for the effects of Voicing and PoA. Fixed factors were contrast-coded for Group (Children as -1 and Adults as 1) and Voicing (Voiced stops as 1 and Voiceless stops as -1) and Helmert-coded for the three-level factor PoA. The first contrast for this factor (i.e., PoA-1) corresponded to the difference between bilabials on one hand and the combined mean of alveolars and velars on the other. This was motivated by the fact that alveolar and velar stops have a different primary articulator (i.e., the tongue) than bilabials (i.e., the lips), and that bilabials tend to be less often released. The second contrast (i.e., PoA-2) corresponded to the difference between alveolars and velars. Post-hoc Tukey tests were performed using the least-squares means method of the *lsmeans* package (Kuznetsova et al., 2017).

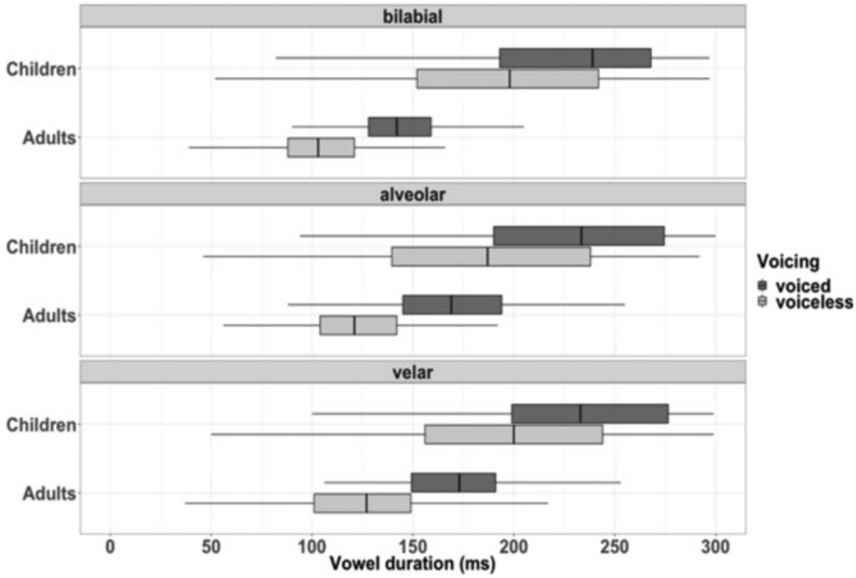

Figure 2. Absolute vowel duration (in ms) by voicing category, PoA and participant group. The middle line of each box corresponds to the median.

Results

Vowel duration

Figure 2 shows the vowel duration by voicing category, PoA and participant group (means are shown in Table IVa). The model fit showed significant effects of Voicing and PoA-1 (bilabials vs. non-bilabials) along with a two-way interaction between Group and Voicing, and a three-way interaction between Group, Voicing and PoA-1 (see Table V). This suggests that, in line with our predictions, vowel duration varied as a function of coda voicing in both children and adults, with longer vowel durations before voiced than before voiceless codas. In addition, it suggests that children had a larger difference between voiced and voiceless categories than the adults, and that this voicing difference was greater in non-bilabials than in bilabials.

Closure duration

Figure 3 shows the closure duration, and Table IVb presents the means for both participant groups. Results of the model fit showed significant effects of Voicing, PoA-1 and PoA-2. There were also three two-way interactions, between Voicing and Group, between Voicing and PoA-1 and between Group and PoA-1, along with two three-way interactions between Voicing, Group and PoA-1, and between Voicing, Group and PoA-2 (see Table VI). This indicates that, as predicted, children and adults exhibited longer closure durations for voiceless than voiced stops and that children had longer closure duration than adults. Children also had a larger difference between voiced and voiceless categories than the adults, and this voicing difference was greater in non-bilabials than in bilabials. Finally, in children, the difference between voiced and voiceless stops was larger for alveolars than for velars.

https://doi.org/10.1017/S0305000920000781 Published online by Cambridge University Press

Table IV. Mean vowel (a), closure (b) and burst (b) durations in milliseconds (SD) by voicing category, PoA and participant group.

	a) Vowel duration					
	Children			Adults		
	Bilabial	Alveolar	Velar	Bilabial	Alveolar	Velar
Voiced	380 (141)	419 (159)	411 (150)	147 (29)	170 (36)	173 (34)
Voiceless	313 (142)	310 (151)	307 (146)	108 (30)	127 (37)	130 (39)
	b) Closure duration					
	Children			Adults		
	Bilabial	Alveolar	Velar	Bilabial	Alveolar	Velar
Voiced	148 (70)	123 (63)	146 (65)	90 (23)	69 (22)	84 (21)
Voiceless	229 (97)	239 (102)	246 (92)	116 (27)	102 (35)	120 (34)
	c) Burst duration					
	Children			Adults		
	Bilabial	Alveolar	Velar	Bilabial	Alveolar	Velar
Voiced	44 (36)	78 (48)	63 (44)	11 (15)	31 (33)	20 (38)
Voiceless	66 (50)	120 (61)	96 (61)	31 (38)	89 (33)	81 (45)

Table V. Results of the linear mixed-effects model for vowel duration.

	Estimate	Std. Error	*t* value
(Intercept)	250.63	14.18	17.67 ***
Voicing	−33.83	4.18	−8.09 ***
Group	107.46	13.60	7.90 ***
PoA-1 (bilabials vs. non-bilabials)	−20.92	8.87	−2.36 *
PoA-2 (alveolars vs. velars)	0.88	10.25	0.09
Voicing * Group	−12.45	1.13	−11.02 ***
Voicing * PoA-1	8.34	8.87	0.94
Voicing * PoA-2	−3.95	10.25	−0.39
Group * PoA-1	2.46	2.38	1.03
Group * PoA-2	2.97	2.78	1.07
Voicing * Group * PoA-1	5.42	2.38	2.27 *
Voicing * Group * PoA-2	−2.96	2.78	−1.07

* = $p < .05$; ** = $p < .01$; *** = $p < .001$

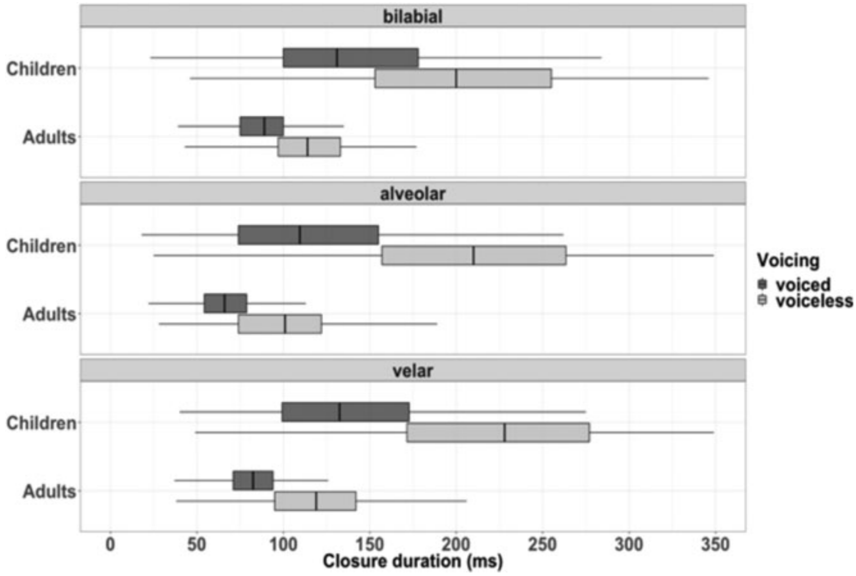

Figure 3. Absolute closure duration (in ms) by voicing category, PoA and participant group. The middle line of each box corresponds to the median.

Table VI. Results of the linear mixed-effects model for closure duration.

	Estimate	Std. Error	*t* value
(Intercept)	141.94	6.77	20.98 ***
Voicing	32.29	0.86	37.71 ***
Group	45.11	6.77	6.67 ***
PoA-1 (bilabials vs. non-bilabials)	3.93	1.83	2.15 *
PoA-2 (alveolars vs. velars)	−16.17	2.08	−7.76 ***
Voicing * Group	16.82	0.81	20.78 ***
Voicing * PoA-1	−8.99	1.82	−4.93 ***
Voicing * PoA-2	3.90	2.08	1.87
Group * PoA-1	−6.21	1.73	−3.59 ***
Group * PoA-2	0.37	1.97	0.19
Voicing * Group * PoA-1	−4.30	1.72	−2.49 *
Voicing * Group * PoA-2	5.76	1.97	2.93 **

* = $p < .05$; ** = $p < .01$; *** = $p < .001$

Burst duration

The burst durations are displayed in Figure 4 by voicing category, PoA and participant group (means are shown in Table IVc). The results of the model fit (see Table VII) revealed significant effects of Voicing, PoA-1 and PoA-2 along with three two-way

https://doi.org/10.1017/S0305000920000781 Published online by Cambridge University Press

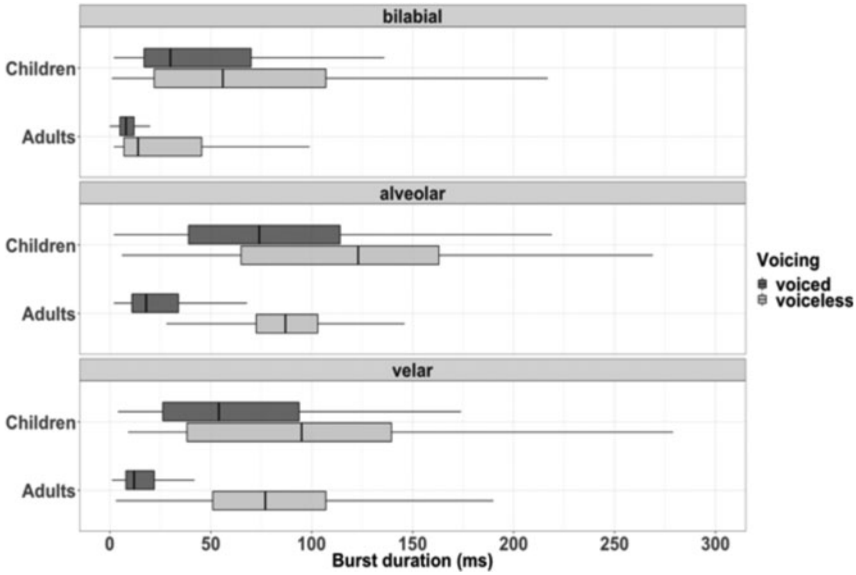

Figure 4. Absolute burst duration (in ms) by voicing category, PoA and participant group. The middle line of each box corresponds to the median.

Table VII. Results of the linear mixed-effects model for burst duration.

	Estimate	Std. Error	*t* value
(Intercept)	60.88	3.39	17.95 ***
Voicing	19.32	0.69	28.07 ***
Group	17.27	3.39	5.09 ***
PoA-1 (bilabials vs. non-bilabials)	−33.95	1.50	−22.69 ***
PoA-2 (alveolars vs. velars)	13.21	1.65	7.98 ***
Voicing * Group	−3.81	0.69	−5.54 ***
Voicing * PoA-1	−13.75	1.48	−9.26 ***
Voicing * PoA-2	1.13	1.65	0.68
Group * PoA-1	0.04	1.50	0.02
Group * PoA-2	4.27	1.66	2.58 *
Voicing * Group * PoA-1	5.96	1.48	4.01 ***
Voicing * Group * PoA-2	2.10	1.65	1.27

* = $p < .05$; ** = $p < .01$; *** = $p < .001$

interactions between Voicing and Group, Voicing and PoA-1, and Group and PoA-2. In addition, there was one three-way interaction between Voicing, Group and PoA-1. In line with our predictions, children exhibited longer burst durations for voiceless than for voiced codas, as did adults, and children had overall longer burst durations than

https://doi.org/10.1017/S0305000920000781 Published online by Cambridge University Press

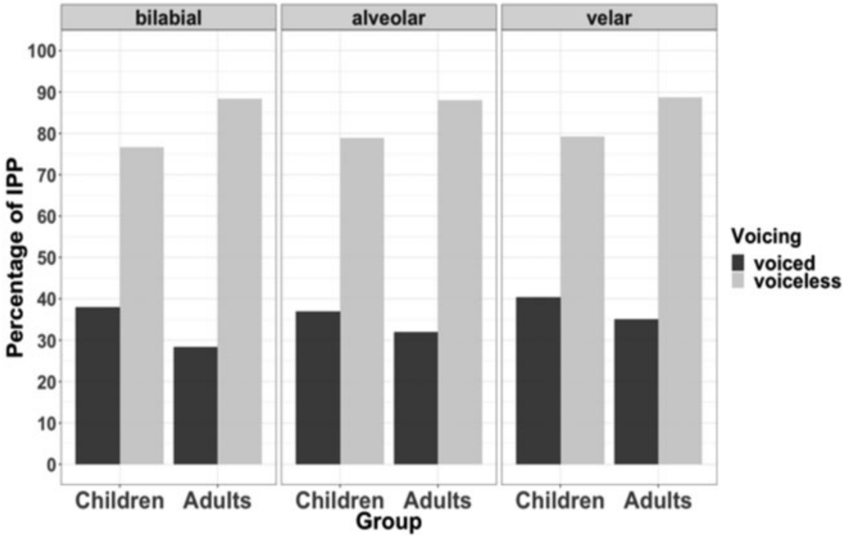

Figure 5. Percentage of irregular pitch period (IPP) by voicing category, PoA and participant group.

the adults. In addition, children had a smaller difference between voiced and voiceless non-bilabials than the adults, and both children and adults showed a pattern of longer burst durations for alveolar than for velar codas, with the durational difference more apparent in the children. Finally, the difference between voiced and voiceless non-bilabials was smaller in children than in adults.

Irregular Pitch Periods

Figure 5 and Table VIIIa show the proportion of IPP at the end of the vowel preceding the coda consonant, by voicing category, PoA and participant group. Results of the model fit (see Table IX) showed a significant effect of Voicing, and a two-way interaction between Voicing and Group. As expected, both children and adults used more IPP before voiceless than voiced codas, though children produced more IPP before voiced codas than adults, and less IPP before voiceless codas than adults.

Voice Bar

The percentage of VB during the closure of the coda consonant, by voicing category, PoA and participant group are presented in Figure 6 and Table VIIIb. The results of the model fit (see Table X) revealed a significant effect of Voicing and a two-way interaction between Voicing and Group. As predicted, both children and adults produced more VB before voiced than voiceless codas. However, children produced fewer instances of VB than adults before voiced codas.

Discussion

The present study investigated the acoustic realization of coda voicing contrasts in 4-year-old speakers of Australian English to determine if these children had acquired

https://doi.org/10.1017/S0305000920000781 Published online by Cambridge University Press

Table VIII. Percentage of IPP (a) and VB (b) by voicing category, PoA and participant group.

	a) Irregular pitch period					
	Children			Adults		
	Bilabial	Alveolar	Velar	Bilabial	Alveolar	Velar
Voiced	38%	37%	40%	28%	32%	35%
Voiceless	76%	79%	79%	88%	88%	89%
	b) Voice bar					
	Children			Adults		
	Bilabial	Alveolar	Velar	Bilabial	Alveolar	Velar
Voiced	62%	65%	67%	94%	92%	91%
Voiceless	3%	2%	1%	0.7%	1%	0.3%

Table IX. Results of the generalized mixed-effects model for IPP.

	Estimate	Std. Error	*z* value
(Intercept)	0.84	0.19	4.48 ***
Voicing	1.50	0.05	28.13 ***
Group	−0.27	0.18	−1.46
PoA-1 (bilabials vs. non-bilabials)	−0.13	0.10	−1.20
PoA-2 (alveolars vs. velars)	−0.15	0.12	−1.46
Voicing * Group	−0.39	0.05	−7.43 ***
Voicing * PoA-1	0.04	0.10	0.35
Voicing * PoA-2	0.01	0.12	0.12
Group * PoA-1	0.03	0.10	0.31
Group * PoA-2	0.00	0.12	0.04
Voicing * Group * PoA-1	−0.15	0.10	−1.43
Voicing * Group * PoA-2	0.02	0.12	0.20

* = $p < .05$; ** = $p < .01$; *** = $p < .001$

adult-like acoustic cues to these contrasts, critical for distinguishing minimal pair words. In line with our predictions, the children in the present study used vowel, closure and burst durations in an adult-like manner to distinguish voiced and voiceless stop codas. Children's vowels before voiced codas were longer than those before voiceless codas, consistent with previous findings in younger American English-speaking children (Ko, 2007; Song et al., 2012), and they produced shorter closure and burst durations for voiced than voiceless codas (cf. Luce & Charles-Luce, 1985; Penney et al., 2018; Song et al., 2012). Nonetheless, children's productions at

https://doi.org/10.1017/S0305000920000781 Published online by Cambridge University Press

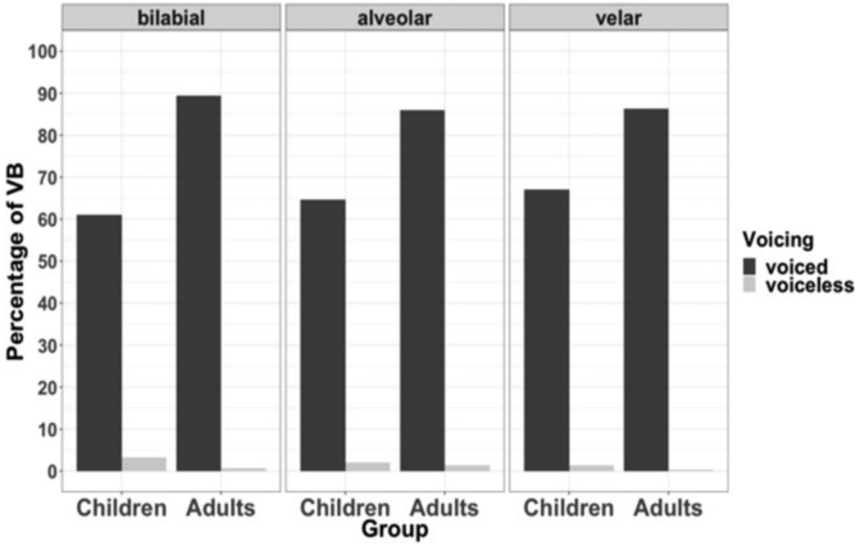

Figure 6. Percentage of voice bar (VB) by voicing category, PoA and participant group.

Table X. Results of the generalized mixed-effects model for VB.

	Estimate	Std. Error	*z* value
(Intercept)	−1.41	0.14	−10.16 ***
Voicing	−2.99	0.14	−21.44 ***
Group	−0.20	0.13	−1.56
PoA-1 (bilabials vs. non-bilabials)	0.24	0.28	0.87
PoA-2 (alveolars vs. velars)	0.45	0.35	1.29
Voicing * Group	0.77	0.13	5.83 ***
Voicing * PoA-1	0.11	0.28	0.40
Voicing * PoA-2	0.44	0.35	1.25
Group * PoA-1	0.04	0.27	0.14
Group * PoA-2	−0.30	0.33	−0.91
Voicing * Group * PoA-1	0.33	0.27	1.22
Voicing * Group * PoA-2	−0.19	0.33	−0.57

* = $p < .05$; ** = $p < .01$; *** = $p < .001$

this age were still systematically longer than those of adults. Figure 7 summarises these observations by voicing category and PoA.

The finding that children's durational cues are longer in absolute terms than those of adults raises the possibility that this might be due to children's slower speaking rate (e.g., Nip and Green, 2013), resulting in longer acoustic durational measures (Green

https://doi.org/10.1017/S0305000920000781 Published online by Cambridge University Press

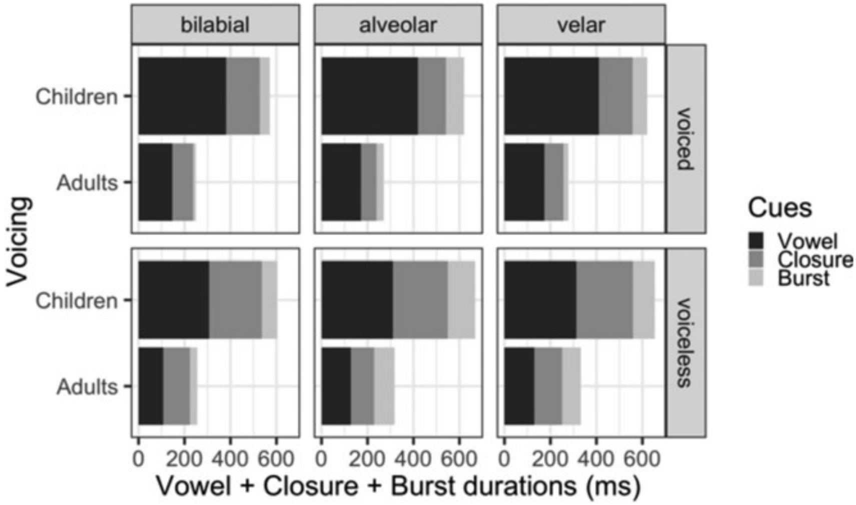

Figure 7. Total rhyme duration (vowel + closure + burst) in milliseconds by voicing category, PoA and participant group.

et al., 2000, 2002; Kowal et al., 1975). To examine this possibility, we compared the children's data to those of the adults, with durational measures mean-centered by group (Enders & Tofighi, 2007): that is, for vowel, closure and burst durations separately, we subtracted the grand mean of each participant group from the individual values within this group. We then re-fitted the original linear mixed-effects models for vowel, closure and burst durations using these mean-centered values. This showed that children's durational cues to coda voicing contrasts remained significantly longer than those of adults, even after speech rate differences were taken into consideration. It is possible that these children's longer durational values arise from a lack of articulatory mastery, with children needing more time to fine-tune the use of these temporal cues to voicing contrasts. This would corroborate previous findings regarding the lack of precision in articulatory timing of speech segments in children below the age of 6 (Green et al., 2000; Lowenstein and Nittrouer, 2008; Nittrouer, 1993; Nittrouer et al., 1996).

Interestingly, our results also showed some effect of PoA on closure and burst durations, with group differences at different PoAs whereby children had a larger difference between bilabials and non-bilabials than adults. Similar results have previously been observed for word-initial voicing contrasts, where 4-year-olds tend to show more adult-like voicing contrasts for bilabials first, followed later by alveolars and velars (Barton & Macken, 1980). This is likely due to the lips being easier to articulate than the tip and body of the tongue. Given young children's smaller oral cavities, contrasts that occur further back in the mouth may need more time to develop (Green et al., 2000).

Finally, we found that the burst durations of alveolar codas were longer than those of velars, for both children and adults. In contrast, Song et al. (2012) reported that the burst durations of alveolars were systematically SHORTER than those of velars for both children and their mothers. This suggests that burst duration may be used somewhat differently in Australian English compared to American English, and that 4-year-olds

https://doi.org/10.1017/S0305000920000781 Published online by Cambridge University Press

are attuned to the cues used in their own dialect of English. This reinforces previous claims in the literature about the importance of systematically documenting the different acoustic implementations of various phonological contrasts for different dialects of English (Chodroff & Wilson, 2017; Scobbie, 2006; Stuart-Smith et al., 2015).

With respect to the binary cues to coda voicing contrasts, the present study found more occurrence of IPP before voiceless than voiced stop codas, for both children and adults, consistent with previous findings for Australian English-speaking adults though not American English children (Song et al., 2012) where the rates of occurrence of IPP did not vary with voicing. Despite producing IPP less often than adults, children followed the adult pattern. As suggested in previous literature (Penney et al., 2018), it seems that IPP in Australian English is a strong correlate to coda voicelessness that is already in place by the age of 4.

In line with previous findings (e.g., Cole et al., 2007; Shattuck-Hufnagel et al., 2011; Song et al., 2012), the probability of VB during closure was higher for voiced than voiceless codas for both children and adults, though children produced VB less often than adults. The proportion of VB found for adults was also consistent with previous literature on both Australian English (Penney et al., 2018) and American English (Song et al., 2012).

In light of previous findings on 2-year-olds (Song et al., 2012), the results of the present study suggest that, as they become older, children refine their use of vowel duration to contrastively mark voicing toward the adult model. It is noted that in Song et al. (2012), the 2-year-olds' vowel duration before VOICELESS codas was similar to that of the adults' in the same study, estimated at ~200 ms. The vowel duration preceding VOICED codas, on the other hand, was estimated at ~375 ms for the 2-year-olds and ~275 ms for the adults. This asymmetry in adult-likeness suggests that the finding of a larger voicing contrast (i.e., a larger difference between the vowel durations preceding voiced and voiceless codas) for the 2-year-olds than for the adults was the result of the children's exaggerated lengthening of the vowel before voiced codas. In the current study, however, children had longer vowel durations than adults before both voiced AND voiceless codas: the mean vowel duration in children was about 400 ms for voiced and 300 ms for voiceless codas whereas in adults it was 160 ms for voiced and 120 ms for voiceless (see Table IV). Surprisingly, the 4-year-olds in the present study still exhibited a larger magnitude of vowel durational difference than adults for voicing contrasts. This suggests that it may take time for children to develop the fine articulatory timing control as discussed above in relation to speech rate (e.g., Green et al., 2002).

Although the burst duration of the 2-year-olds in Song et al. (2012) did not differ from that of the adults, the 4-year-olds in the current study produced longer burst durations than the adults. This difference might be related to different speech registers: lab speech in the present study vs. spontaneous speech in Song et al. (2012).

Both the 2;6-year-olds in Song et al. (2012) and the 4-year-olds in the current study produced VB for approximately 60% of all voiced codas, suggesting that the use of VB as a cue to coda voicing might have been established by around 2;6 years. Interestingly, these 2-year-olds produced more VB for voiceless codas than the 4-year-olds, which might be related to the different dialects these children were acquiring (i.e., American English vs. Australian English). The use of IPP also varies between the two studies. The (American English-speaking) 2-year-olds produced IPP for about 30% of all voiceless codas, whereas the (Australian English-speaking) 4-year-olds did so for over 70% of all voiceless codas. This difference might be due to the dialectal difference

https://doi.org/10.1017/S0305000920000781 Published online by Cambridge University Press

between Australian and American English, given that in both studies adults and children had a similar proportion of IPP. This suggests that the language-specific association of IPP with voicelessness in Australian English might limit any ambiguous use of VB to signal voicelessness.

The results of the present study thus build on previous studies of younger, American English-speaking children, showing that Australian English-speaking 4-year-olds can use adult-like acoustic cues to coda voicing contrasts, including both durational information (vowel, closure and burst durations) and binary cues (IPP and VB). However, even at this older age, children's acoustic implementation of the durational cues tends to be longer than those of adults, and they still use less IPP and VB. These findings contribute to our understanding of phonological development in typically-developing children and provide a much-needed acoustic baseline for evaluating the development of voicing contrasts in populations with language delay.

Unlike stop voicing contrasts in word onset position, coda stop voicing contrasts are still understudied. Although we have here made a start at remedying this situation by looking at the acoustic cues to coda voicing contrasts in PRODUCTION, it would be interesting in future to investigate the PERCEPTION of coda contrasts. This could elucidate, for instance, whether children rely on different acoustic cues than adults when listening to voicing distinctions in coda position, and how the weighting of various cues develops over time. This would provide a more comprehensive picture of coda development, providing a baseline for understanding the perception and production of codas in other populations, such as those with hearing difficulties.

Conclusion

The goal of the present study was to determine if pre-school-aged children had acquired adult-like phonetic implementations for coda voicing contrasts, critical for distinguishing word meanings in English. Our results provide a much-needed acoustic understanding of children's ongoing phonological development. Since most language evaluations of clinical populations are transcription-based, systematic acoustic analysis is essential for providing complementary information about HOW and WHEN voicing contrasts are acquired in these populations. The findings presented here will therefore provide a valuable baseline for contributing to our knowledge of voicing contrasts in Australian English, and for evaluating challenges faced by various populations with language delay.

Acknowledgments. We thank the members of the Child Language Lab and the Phonetics Lab at Macquarie University for helpful comments and feedback. We also thank Xin Cheng for assistance with data coding. This research was supported, in part, by a Macquarie University scholarship (2013225) to the first author, an Australian Research Council (ARC) Laureate Fellowship to the last author (#FL130100014), and by the ARC Centre of Excellence in Cognition and its Disorders (#CE110001021). The authors have no competing interests to declare.

References

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

Baudonck, N., Lierde, K. V., D'haeseleer, E., & Dhooge, I. (2011). A comparison of the perceptual evaluation of speech production between bilaterally implanted children, unilaterally implanted children, children using hearing aids, and normal-hearing children. *International Journal of Audiology*, *50*(12), 912–919. https://doi.org/10.3109/14992027.2011.605803

Boersma, P., & Weenink, D. (2019). *Praat: Doing phonetics by computer* (6.1.08) [Computer software]. http://www.praat.org/

Chodroff, E., & Wilson, C. (2017). Structure in talker-specific phonetic realization: Covariation of stop consonant VOT in American English. *Journal of Phonetics*, *61*, 30–47. https://doi.org/10.1016/j.wocn.2017.01.001

Cole, J., Kim, H., Choi, H., & Hasegawa-Johnson, M. (2007). Prosodic effects on acoustic cues to stop voicing and place of articulation: Evidence from Radio News speech. *Journal of Phonetics*, *35*(2), 180–209. https://doi.org/10.1016/j.wocn.2006.03.004

Cox, F., & Palethorpe, S. (2007). Australian English. *Journal of the International Phonetic Association*, *37*(3), 341–350. https://doi.org/10.1017/S0025100307003192

Cox, F., and Palethorpe, S. (2011). Timing differences in the VC rhyme of standard Australian English and Lebanese Australian English, in W. S. Lee & E. Zee (Eds.), *Proceedings of the ICPhS XVIIth International Congress of Phonetic Sciences*, pp. 528–531, Hong Kong.

de Jong, K. (2004). Stress, lexical focus, and segmental focus in English: Patterns of variation in vowel duration. *Journal of Phonetics*, *32*(4), 493–516. https://doi.org/10.1016/j.wocn.2004.05.002

Demuth, K., Culbertson, J., & Alter, J. (2006). Word-minimality, epenthesis and coda licensing in the early acquisition of English. *Language and Speech*, *49*(2), 137–173. https://doi.org/10.1177/00238309060490020201

Docherty, G. J., & Foulkes, P. (1999). Derby and Newcastle: Instrumental phonetics and variationist studies. In G. J. Docherty & P. Foulkes (Eds.), *Urban voices: Accent studies in the British Isles* (pp. 47–71). Arnold.

Enders, C. K., & Tofighi, D. (2007). Centering predictor variables in cross-sectional multilevel models: A new look at an old issue. *Psychological Methods*, *12*(2), 121–138. https://doi.org/10.1037/1082-989X.12.2.121

Forrest, K., Weismer, G., Hodge, M., Dinnsen, D. A., & Elbert, M. (1990). Statistical analysis of word-initial /k/ and /t/ produced by normal and phonologically disordered children. *Clinical Linguistics & Phonetics*, *4*(4), 327–340. https://doi.org/10.3109/02699209008985495

Fowler, C. A. (1992). Vowel duration and closure duration in voiced and unvoiced stops: There are no contrast effects here. *Journal of Phonetics*, *20*(1), 143–165. https://doi.org/10.1016/S0095-4470(19)30244-X

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

Green, J. R., Moore, C. A., & Reilly, K. J. (2002). The sequential development of jaw and lip control for speech. *Journal of Speech, Language, and Hearing Research*, *45*(1), 66–79. https://doi.org/10.1044/1092-4388(2002/005)

Imbrie, A. (2005). *Acoustical study of the development of stop consonants in children* [Thesis, Massachusetts Institute of Technology]. https://dspace.mit.edu/handle/1721.1/33072

Jakobson, R. (1968). *Child language, aphasia and phonological universals*. Walter de Gruyter.

Kewley-Port, D., & Preston, M. S. (1974). Early apical stop production: A voice onset time analysis. *Journal of Phonetics*, *2*(3), 195–210. https://doi.org/10.1016/S0095-4470(19)31270-7

Kirk, C., & Demuth, K. (2006). Accounting for variability in 2-year-olds' production of coda consonants. *Language Learning and Development*, *2*(2), 97–118. https://doi.org/10.1207/s15473341lld0202_2

Ko, E.-S. (2007). "Acquisition of vowel duration in children speaking American English," in Proceedings of Interspeech 2007, pp. 1881–1884, Antwerp, Belgium.

Koenig, L. L. (2001). Distributional characteristics of VOT in children's voiceless aspirated stops and interpretation of developmental trends. *Journal of Speech, Language, and Hearing Research*, *44*(5), 1058–1068. https://doi.org/10.1044/1092-4388(2001/084)

Kowal, S., O'Connell, D. C., & Sabin, E. J. (1975). Development of temporal patterning and vocal hesitations in spontaneous narratives. *Journal of Psycholinguistic Research*, *4*(3), 195–207. https://doi.org/10.1007/BF01066926

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.

Li, F., Edwards, J., & Beckman, M. E. (2009). Contrast and covert contrast: The phonetic development of voiceless sibilant fricatives in English and Japanese toddlers. *Journal of Phonetics*, *37*(1), 111–124. https://doi.org/10.1016/j.wocn.2008.10.001

Lisker, L. (1957). Closure duration and the intervocalic voiced-voiceless distinction in English. *Language*, *33*(1), 42. https://doi.org/10.2307/410949

Lowenstein, J. H., & Nittrouer, S. (2008). Patterns of acquisition of native voice onset time in English-learning children. *The Journal of the Acoustical Society of America*, *124*(2), 1180–1191. https://doi.org/10.1121/1.2945118

Luce, P. A., & Charles-Luce, J. (1985). Contextual effects on vowel duration, closure duration, and the consonant/vowel ratio in speech production. *The Journal of the Acoustical Society of America*, *78*(6), 1949–1957. https://doi.org/10.1121/1.392651

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

Markides, A. (1970). The speech of deaf and partially-hearing children with special reference to factors affecting intelligibility. *International Journal of Language & Communication Disorders*, *5*(2), 126–140. https://doi.org/10.3109/13682827009011511

Miles, K., Demuth, K., & Ching, T. (2012). Acoustic analysis of the speech of an Australian English-speaking child with hearing aids. In F. Cox, K. Demuth, S. Lin, K. Miles, I. Yuen, S. Palethorpe, & J. Shaw (Eds.), *Proceedings of the 14th Australasian International Conference on Speech Science and Technology*. (pp. 97–100). Australian Speech Science and Technology Association.

Munson, B., Edwards, J., & Beckman, M. E. (2005). Phonological knowledge in typical and atypical speech–sound development. *Topics in Language Disorders*, *25*(3), 190–206. https://doi.org/10.1097/00011363-200507000-00003

Nip, I. S. B., & Green, J. R. (2013). Increases in cognitive and linguistic processing primarily account for increases in speaking rate with age. *Child Development*, *84*(4), 1324–1337. https://doi.org/10.1111/cdev.12052

Nittrouer, S. (1993). The emergence of mature gestural patterns Is not uniform: Evidence from an acoustic study. *Journal of Speech, Language, and Hearing Research*, *36*(5), 959–972. https://doi.org/10.1044/jshr.3605.959

Nittrouer, S., Studdert-Kennedy, M., & Neely, S. T. (1996). How children learn to organize their speech gestures: Further evidence from fricative-vowel syllables. *Journal of Speech, Language, and Hearing Research*, *39*(2), 379–389. https://doi.org/10.1044/jshr.3902.379

Penney, J., Cox, F., Miles, K., & Palethorpe, S. (2018). Glottalisation as a cue to coda consonant voicing in Australian English. *Journal of Phonetics*, *66*, 161–184. https://doi.org/10.1016/j.wocn.2017.10.001

Raphael, L. J. (1972). Preceding vowel duration as a cue to the perception of the voicing characteristic of word-final consonants in American English. *The Journal of the Acoustical Society of America*, *51*(4B), 1296–1303. https://doi.org/10.1121/1.1912974

R Core Team. (2013). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. http://www.R-project.org/

Redi, L., & Shattuck-Hufnagel, S. (2001). Variation in the realization of glottalization in normal speakers. *Journal of Phonetics*, *29*(4), 407–429. https://doi.org/10.1006/jpho.2001.0145

Scobbie, J. M. (2006). Flexibility in the face of incompatible English VOT systems. In L. Goldstein, D. H. Whalen and C. T. Best (Eds.), *Laboratory Phonology* Vol. 1–8, pp. 367–392. Berlin: Mouton de Gruyter.

Scobbie, J. M., Gibbon, F., Hardcastle, W., & Fletcher, P. (2000). Covert contrast as a stage in the acquisition of phonetics and phonology. In J. B. Pierrehumbert & M. Broe (Eds.), *Papers in Laboratory Phonology V: Acquisition and the Lexicon* (pp. 194–207). Cambridge University Press.

Shattuck-Hufnagel, S., Demuth, K., Hanson, H., & Stevens, K. (2011). Acoustic cues to stop-coda voicing contrasts in the speech of American English 2–3-year-olds. In G. Clements Nick & R. Ridouane (Eds.), *Where do phonological features come from? Cognitive, physical and developmental bases of distinctive speech categories* (pp. 327–341). John Benjamins.

Smith, B. L. (1979). A phonetic analysis of consonantal devoicing in children's speech. *Journal of Child Language*, *6*(1), 19–28. https://doi.org/10.1017/S0305000900007595

Smith, N. V. (1973). *The acquisition of phonology: A case study*. Cambridge University Press.

Song, J. Y., & Demuth, K. (2008). Compensatory vowel lengthening for omitted coda consonants: A phonetic investigation of children's early representations of prosodic words. *Language and Speech*, *51* (4), 385–402. https://doi.org/10.1177/0023830908099071

Song, J. Y., Demuth, K., & Shattuck-Hufnagel, S. (2012). The development of acoustic cues to coda contrasts in young children learning American English. *The Journal of the Acoustical Society of America*, *131*(4), 3036–3050. https://doi.org/10.1121/1.3687467

Stoel-Gammon, C., & Buder, E. (1999). Vowel length, postvocalic voicing and VOT in the speech of two-year olds. In K. Elenius & P. Branderud (Eds.), *Proceedings of the 13th International Congress of Phonetic Sciences* (Vol. 3, pp. 2485–2488).

Stuart-Smith, J., Sonderegger, M., Rathcke, T., & Macdonald, R. (2015). The private life of stops: VOT in a real-time corpus of spontaneous Glaswegian. *Laboratory Phonology*, *6*(3–4). https://doi.org/10.1515/lp-2015-0015

Theodore, R. M., Demuth, K., & Shattuck-Hufnagel, S. (2012). Segmental and positional effects on children's coda production: Comparing evidence from perceptual judgments and acoustic analysis. *Clinical Linguistics & Phonetics*, *26*(9), 755–773. https://doi.org/10.3109/02699206.2012.700680

van Heuven, W. J. B., Mandera, P., Keuleers, E., & Brysbaert, M. (2014). SUBTLEX-UK: A new and improved word frequency database for British English. *Quarterly Journal of Experimental Psychology*, *67*(6), 1176–1190.

Weismer, G., Dinnsen, D., & Elbert, M. (1981). A study of the voicing distinction associated with omitted, word-final stops. *Journal of Speech and Hearing Disorders*, *46*(3), 320–328. https://doi.org/10.1044/jshd.4603.320

Xu Rattanasone, N., & Demuth, K. (2014). The acquisition of coda consonants by Mandarin early child L2 learners of English. *Bilingualism: Language and Cognition*, *17*(3), 646–659. https://doi.org/10.1017/S1366728913000618

Yu, V. Y., De Nil, L. F., & Pang, E. W. (2015). Effects of age, sex and syllable number on voice onset time: Evidence from children's voiceless aspirated stops. *Language and Speech*, *58*(2), 152–167. https://doi.org/10.1177/0023830914522994

Zlatin, M. A., & Koenigsknecht, R. A. (1976). Development of the voicing contrast: A comparison of voice onset time in stop perception and production. *Journal of Speech and Hearing Research*, *19*(1), 93–111. https://doi.org/10.1044/jshr.1901.93

Cite this article: Millasseau J, Yuen I, Bruggeman L, Demuth K (2021). Acoustic cues to coda stop voicing contrasts in Australian English-speaking children. *Journal of Child Language* **48**, 1262–1280. https://doi.org/10.1017/S0305000920000781