Segmental and positional effects on children's coda production: comparing evidence from perceptual judgments and acoustic analysis

RACHEL M. THEODORE1, KATHERINE DEMUTH2, & STEPHANIE SHATTUCK-HUFNAGEL3

1Communication Sciences, Communication Disorders Program, University of Connecticut, Storrs, CT, USA, 2Department of Linguistics, Macquarie University, Sydney, Australia, and 3Speech Communication Group, Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, MA, USA

(Received 26 October 2011; revised 23 May 2011; accepted 04 June 2012)

Abstract
Children's early productions are highly variable. Findings from children's early productions of grammatical morphemes indicate that some of the variability is systematically related to segmental and phonological factors. Here, we extend these findings by assessing 2-year-olds' production of non-morphemic codas using both listener decisions and acoustic analyses. Results showed that utterance position and coda manner influence perception, in that more stop codas were perceived utterance-finally compared to utterance-medially but fricative codas were perceived equally across utterance positions. Acoustic analyses showed some convergence to listeners' perception in that there were more cues associated with stops utterance-finally compared to utterance-medially. However, there was some divergence between the two methods in that acoustic cues to coda segments were also present in the majority of cases where a coda was not perceived. These findings provide insight into both the nature of children's emerging phonological representations and the effectiveness of coda transcription across segment types.

Keywords: phonological acquisition, coda consonant, perception, acoustic analysis

Introduction
Children's early within-speaker productions are highly variable with individual speech segments and syllables sometimes altered or missing. This leads to difficulty in specifying a point at which individual speech segments are “acquired”. However, being able to identify when consonants and vowels are acquired is necessary for investigating the time-course of segmental acquisition. Moreover, determining appropriate ways to quantify when consonants and vowels are acquired is important for clinical practice given the ubiquitous use of metrics of consonant mastery for identifying children with specific language impairment and phonological disorders, as well as for monitoring rehabilitative outcomes. The purpose of this study is to contribute to this effort through examination
of acoustic cues present in children’s early productions. Specifically, we analyze the acoustic characteristics of coda production in 2-year-old children, focusing on individual acoustic cues that potentially mark the emergence of stop and fricative consonants in their early productions. In addition, this study examines how the acoustic cues present in children’s speech contribute to adult listeners’ perception of children’s speech.

The most pervasive method for determining consonantal mastery in both the research and clinical domains is by phonemic transcription of children’s speech. However, recent research suggests that acoustic analyses can be used to both complement and extend these perceptual analyses (Edwards & Beckman, 2008). For example, recent findings indicate that transcription is influenced by factors such as the perceived age of the speaker, the phonetic expertise of the transcriber, whether or not the speaker and transcriber share the same native language, as well as whether or not the transcriber is given one or multiple categories to transcribe (as reviewed in Munson, Edwards, Schellinger, Beckman, & Meyer, 2010). One potential way to mediate these transcriber biases, particularly in the research domain, is acoustic analysis of the speech signal, where a child’s productions are evaluated on the basis of specific acoustic information present in the utterance instead listeners’ perception of a particular utterance. We do not mean to imply that acoustic analysis is not subject to certain types of biases, particularly in terms of expertise with the acoustic signal; rather, we suggest that acoustic analysis may be one way to remove listener-dependent biases stated above and thus complement the traditional transcription approach.

Other evidence in support of a more fine-grained approach to the analysis of children’s early productions comes from the literature on covert contrast. Covert contrast is defined as acoustic distinctions that children make to signify linguistic contrasts that are not detected by adult listeners (Macken & Barton, 1980; Scobbie, 1998; Scobbie, Gibbon, Hardcastle, & Fletcher, 2000). Such covert contrasts have been documented in the speech of typically developing children as well as in children with language impairments (Forrest, Weismer, Hodge, Dinnsen, & Elbert, 1990). They have been demonstrated for stop voicing contrasts (Macken & Barton, 1980; Weismer, Dinnsen, & Elbert, 1981), stop place of articulation contrasts (White, 2001) and fricatives (Li, Edwards, & Beckman, 2009). Given that, in some cases, transcription may underestimate a child’s knowledge of language, acoustic analysis has been proposed as a potentially more fine-grained method to examine emerging linguistic knowledge.

Most of the research on covert contrasts, described above, has focused on the realization of word-initial onset consonants. Relatively less is known about the acoustic realization of word-final coda consonants and the extent to which consonants in this position are susceptible to covert contrasts. Of the work that has examined development of coda production, investigations have focused on compensatory vowel lengthening for omitted coda consonants (Song & Demuth, 2008; Weismer et al., 1981), and not on the acoustics of the coda segment per se. Examining the acoustics of coda production is a particularly interesting theoretical issue given the fact that coda consonants increase the complexity of syllable structures and are typically later acquired (Demuth, Culbertson, & Alter, 2006; Kirk & Demuth, 2006). Thus, these later-acquired forms provide a potentially rich environment in which to observe evidence of covert contrast, due to the fact that they remain omitted for longer periods of time. Moreover, languages such as English encode many inflectional morphemes as word-final consonants (e.g. plural –s, third person singular –s, past tense –ed). A better understanding of children’s production abilities at the ends of words would thus also provide a baseline against which to evaluate the production of inflectional morphemes.

One central finding from studies that examine early productions of grammatical morphemes is that these productions are highly variable, just as productions are for mono-morphemic segments (Bloom, 1970; Brown, 1973). Variability in the production of grammatical morphemes has typically been taken as evidence that children’s developing grammars are syntactically incomplete, with a syntactic construct seen as mastered only when the child produces a given morpheme consistently
Segmental and positional effects on coda production

(Marcus et al., 1992). However, recent findings suggest that within-speaker variability in the production of grammatical morphemes may not indicate impoverished syntactic representations, but rather may reflect interactions between syntactic and other levels of language representation and speech planning (Theodore, Demuth, & Shattuck-Hufnagel, 2010). This argument is based on a host of findings that indicate that morpheme production is systematically related to both semantic and phonological factors, including verb type (Bloom, Lifter, & Hafitz, 1980; Johnson & Morris, 2007), physical distinctiveness of objects (Zapf & Smith, 2008), whether or not the morpheme is part of a trochaic foot (Demuth & McCullough, 2009; Gerken, 1996), and the syllabic complexity of the coda (Marshall & van der Lely, 2007; Song, Sundara, & Demuth, 2009).

Song et al. (2009) also showed that utterance position influences production of third person singular –s, with more morphemes produced utterance-finally compared to utterance-medially. The same positional effect has also been documented in the case of plural –s, which phonologically manifests the same as third person singular –s (Theodore et al., 2010). In order to determine whether or not the positional effect is due to morpheme production proper, it is necessary to examine whether utterance position influences production of non-morphemic /s/ and /z/ as well.

The goal of the current work was to examine production of word-final, non-morphemic segments in 2-year-old children’s speech. In this study, we directly compared the results of acoustic analysis to those of standard transcription-based analysis of children’s coda productions. Specifically, we examined the influences of coda manner (fricative vs. stop), coda voicing (voiced vs. voiceless) and utterance position (medial vs. final) on the production of coda consonants. Our focus on these factors is motivated by recent findings. Specifically, utterance position is of interest because recent findings show that production of grammatical morphemes is systematically related to utterance position (Theodore et al. 2010). We consider coda manner in light of findings that have shown that stops and fricatives generally follow different developmental trajectories, with stops acquired earlier than fricatives, and thus may show differential effects in their acoustic manifestations (Kehoe & Stoel-Gammon, 2001). In addition, coda manner is considered given that some grammatical morphemes are marked by stops (i.e., past tense –ed) while others are marked by fricatives (e.g., plural –s); thus, including manner of articulation as a factor in this study will provide a point of comparison for the emergence of different grammatical morphemes. The inclusion of coda voicing as a factor is motivated by earlier work that has shown that some of the acoustic cues we examine here are systematically related to coda voicing (Song, Shattuck-Hufnagel, & Demuth, 2011; Song, Demuth, & Shattuck-Hufnagel, 2012).

In Experiment 1, phonetically trained adults were asked to listen to children’s utterances and indicate whether a coda was present using a two-alternative forced-choice paradigm. We examined the degree to which coda voicing, coda manner and utterance position influenced the listener-based decisions. In Experiment 2, we again quantified the presence or absence of coda consonants in children’s productions, but now the metric of coda production was based on acoustic analyses in order to detect individual cues to segmentational feature contrasts (Shattuck-Hufnagel, Demuth, Hanson, & Stevens, in press). Such a design allowed us to directly compare consonantal mastery as a function of the two different metrics of listener decision vs. acoustic analysis. It also allowed us to identify potential acoustic characteristics of the speech signal that are not detected by the adult listener. Accordingly, this design affords the possibility of providing insight into the nature of children’s emerging phonological representations that would not be obtained with each metric of coda production considered independently.

Globally, we predict that if the effect of utterance position observed on the production of morphemic /s/ and /z/ holds for non-morphemic segments, then more codas will be present in words that are produced utterance-finally compared to utterance-medially. Moreover, given that fricatives are acquired later than stops (Kehoe & Stoel-Gammon, 2001), we predict that we will observe more robust production of stop codas compared to fricative codas. We predict that effects of coda...
voicing will differentially manifest for particular acoustic cues, given earlier research showing that some acoustic cues are more prevalent for voiced compared to voiceless segments. Finally, acoustic analysis will reveal whether some of the tokens perceived by the adult listeners as having no coda nevertheless contain acoustic evidence of a coda representation.

**General Methods**

The focus of this study concerns child productions of eight target nouns that were elicited using an imitation task. In Experiment 1, these utterances were presented to adult listeners who were asked to indicate, for each target, whether they perceived a coda consonant. In Experiment 2, acoustic analyses were performed on the utterances in order to identify acoustic cues to feature contrasts. Details on the subject population and analyses are provided for each experiment below; here we describe the methods used to obtain the children's productions.

**Stimuli**

Eight monosyllabic consonant-vowel-consonant (CVC) target nouns were selected for use in the experiment. Half contained a fricative coda (e.g. *bus*) and half contained a stop coda (e.g. *foot*). Across the eight targets, all codas had an alveolar place of articulation, with equal numbers of voiced and voiceless targets for both coda manners. Thus, the segments examined are the same as the inflections used to mark grammatical morphemes; /s/ and /z/ in the case of plural –s and third person singular –s, and /t/ and /d/ in the case of past tense –ed. The eight targets are listed in Table I. Examining this set of speech segments provides baseline information for comparing coda production in terms of the manner of the segment to be produced.

In light of findings showing that lexical factors such as phonotactic probability (Zamuner, 2009) and segmental characteristics including sonority (Kirk & Demuth, 2006) and place of articulation (Kirk, 2008) may influence children's productions, an attempt was made to control the lexical characteristics of the target nouns. Thus, the onset consonant varied across the target words, but in such a way that variability was equally present for stop and fricative codas. The eight monosyllabic CVC target words were also selected to be frequent, familiar nouns that were easy to picture. The target nouns did exhibit some variability in both lexical frequency and familiarity, a consequence deemed acceptable given the otherwise precise phonological control of the targets. Eight pictures (one for each target noun) were selected to serve as visual prompts during the experiment. All pictures were selected to be realistic representations of the nouns (as opposed to cartoon style) and to be of similar size and interest. To offset potential influences of noun familiarity, the pictures of the

<table>
<thead>
<tr>
<th>Coda</th>
<th>Target</th>
<th>Utterance position</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Medial</td>
</tr>
<tr>
<td>Fricative</td>
<td></td>
<td>His nose dripped</td>
</tr>
<tr>
<td></td>
<td><em>nose</em></td>
<td>My fuzz blew</td>
</tr>
<tr>
<td></td>
<td><em>fuzz</em></td>
<td>Her juice dropped</td>
</tr>
<tr>
<td></td>
<td><em>juice</em></td>
<td>My bus crashed</td>
</tr>
<tr>
<td></td>
<td><em>bus</em></td>
<td></td>
</tr>
<tr>
<td>Stop</td>
<td></td>
<td>My food burned</td>
</tr>
<tr>
<td></td>
<td><em>food</em></td>
<td>Her bed came</td>
</tr>
<tr>
<td></td>
<td><em>bed</em></td>
<td>My boat broke</td>
</tr>
<tr>
<td></td>
<td><em>boat</em></td>
<td>His foot kicked</td>
</tr>
<tr>
<td></td>
<td><em>foot</em></td>
<td></td>
</tr>
</tbody>
</table>
target items were shown to each child prior to the onset of the experiment and a name for each object was provided.

As shown in Table I, each target stimulus noun was embedded in two sentences, one in which it appeared utterance-medially and one in which it appeared utterance-finally. To control for utterance length, all stimulus sentences consisted of three monosyllabic words. To control for potential articulatory and phonological influences at the phrase level, the utterance-medial sentences were constructed such that the target noun was always followed by a word that began with a stop consonant. This medial context was selected because (1) it is a difficult context from an articulatory perspective (compared to a following vowel, for example) and thus increased the opportunity to measure potential effects of utterance position and (2) it provides more direct comparison to earlier work (Theodore et al., 2010) showing positional influences on morpheme production that used the same medial context.

Children's productions of these 16 sentences were elicited by presenting an auditory prompt for imitation. To create the prompts, a female native speaker of American English was recorded producing the 16 sentences in an infant-directed speech register. The speaker had extensive experience in research in infant speech perception and she attained the infant-directed register by using enhanced pitch contours and a hyper-articulated speaking style. The recording session took place in a sound-attenuated booth. Speech was recorded via microphone connected to a pre-amplifier directly to computer (sampling frequency of 44.1 KHz, 16-bit quantization). All sentences were excised from this long sound file into separate files using the Praat software (Boersma & Weenink, 2010).

In order to ensure that the prompts presented to the children were consistent previous findings showing positional effects on word duration (Song et al., 2009) and that the overall duration of the utterance-medial and utterance-final sentences was equivalent, two acoustic measurements were performed for each sentence using Praat (Boersma & Weenink, 2010). First, we measured the duration of the entire sentence, calculated by the latency between the onset and offset of vocal energy present in the utterance. Second, we measured the duration of the target noun as the difference between the onset of closure (for the initial stop) or the onset of high frequency aperiodic

---

Table II. Target and sentence duration in milliseconds (ms) for the fricative and stop coda prompts in medial and final utterance position.

<table>
<thead>
<tr>
<th>Coda</th>
<th>Utterance position</th>
<th>Prompt</th>
<th>Duration (ms)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Target</td>
</tr>
<tr>
<td>Fricative</td>
<td>Medial</td>
<td>His nose dripped</td>
<td>540</td>
</tr>
<tr>
<td></td>
<td></td>
<td>My fuzz blew</td>
<td>600</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Her juice dropped</td>
<td>690</td>
</tr>
<tr>
<td></td>
<td></td>
<td>My bus crashed</td>
<td>620</td>
</tr>
<tr>
<td></td>
<td>Final</td>
<td>That's his nose</td>
<td>1230</td>
</tr>
<tr>
<td></td>
<td></td>
<td>That's my fuzz</td>
<td>960</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Here's her juice</td>
<td>770</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Here's my bus</td>
<td>820</td>
</tr>
<tr>
<td>Stop</td>
<td>Medial</td>
<td>My food burned</td>
<td>660</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Her bed came</td>
<td>450</td>
</tr>
<tr>
<td></td>
<td></td>
<td>My boat broke</td>
<td>600</td>
</tr>
<tr>
<td></td>
<td></td>
<td>His foot kicked</td>
<td>490</td>
</tr>
<tr>
<td></td>
<td>Final</td>
<td>That's my food</td>
<td>920</td>
</tr>
<tr>
<td></td>
<td></td>
<td>That's her bed</td>
<td>800</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Here's my boat</td>
<td>900</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Here's his foot</td>
<td>700</td>
</tr>
</tbody>
</table>
energy (for the initial fricative) and the offset of energy associated with the coda consonant. These measurements are shown in Table II. Each set of measurements was submitted to an ANOVA with coda manner (fricative vs. stop) as a between-subjects factor and utterance position (medial vs. final) as a within-subjects factor. In terms of sentence duration, neither an effect of coda manner \([F(1,6) = 1.20, p = 0.316]\) nor position \([F(1,6) = 0.30, p = 0.869]\), nor an interaction between these factors \([F(1,6) = 0.26, p = 0.628]\), was observed, indicating that, as expected, the overall duration of the sentences was approximately the same. In terms of duration of the target noun, the expected effect of position was observed in that targets in utterance-final position were longer than targets in utterance-medial position \([F(1,6) = 20.94, p = 0.004]\) due to phrase-final lengthening (Turk & Shattuck-Hufnagel, 2007). Neither an effect of coda manner \([F(1,6) = 0.95, p = 0.368]\) nor any interaction between utterance position and manner was observed \([F(1,6) = 0.14, p = 0.725]\). To sum, these analyses indicate that the overall duration of the stimuli was comparable for both fricative and stop coda targets in both utterance positions. Moreover, the duration of the target noun was equivalent across fricative and stop codas, but longer in utterance-final compared to utterance-medial position.

Subjects

The subjects were 11 children from the Providence, RI community. All children were full-term 2-year-olds (7 female and 4 male) from monolingual English-speaking homes. The children ranged in age from 2;2 to 2;6, with a mean of age of 2;4. All children were healthy on the day of testing and had typically developing speech and language skills according to parental report. MacArthur Communicative Development Inventory (CDI) percentile scores on vocabulary size ranged from 10 to 90 with a mean of 58 (Fenson et al., 2000). An additional 19 children participated but were not included in the analyses (12 did not speak and 7 did not meet the criterion for inclusion – i.e. repetition of at least 12 of the 16 prompts). This attrition rate is consistent with previous experiments using similar tasks with children in this age group (Song et al., 2009), and does not seem to be influenced by the age of the subjects in that the mean age (2;3) and age range (2;2–2;6) of those excluded is similar to those who were included in the analyses.

Procedure

Procedural reliability was ensured by having the same two experimenters (the first author and a highly trained research assistant) adhere to the following protocol for each subject. After a brief familiarization period with the experimenters (often involving engaging the child with a picture book), the child was invited into a sound-attenuated room with a parent to “play a game” with one experimenter. The room contained a child-sized table and chairs, with a computer monitor and speakers on top of the table. The room was equipped with two lavaliere microphones (Audio-technica 700 Series) connected to a computer in an adjoining room via the MBox 2 Audio Interface (Digidesign). In most cases, the microphones were placed on the table near the child in order to best capture his or her speech; in a few cases, a microphone was attached to the collar of the child’s shirt. Following a brief warm-up period where the child practiced repeating what the computer said, and thus was exposed to both the task and the particular target nouns used in the present experiment, the child was directed to face the computer in order for the game to begin.

The entire procedure took approximately 30 min. On each trial, a picture of the target noun appeared on the monitor along with the auditory prompt, and the child was directed to repeat the entire sentence provided in the prompt. Five attempts were allowed before moving to the next trial. The child was encouraged with praise and, in some cases, stickers, for successful trials.
completion of the 16 trials, parents completed a brief demographic survey and the MacArthur CDI short form to estimate the child's vocabulary size.

The tokens elicited from the children were then presented to adult listeners who were asked to indicate whether or not a coda was perceived (Experiment 1), and were then analyzed acoustically for cues to coda production (Experiment 2). For both the listener-based and acoustic metrics of coda production, our goal was to examine the extent to which coda production is influenced by coda manner, voicing and utterance position.

**Experiment 1**

As described above, perceptual transcription is a commonly used method for analyzing children's speech, both in the language research and clinical domains. The goal of Experiment 1 was to obtain further information regarding how coda voicing, coda manner and utterance position influence listener perception of a coda consonant.

Previous research has found that production of third person singular –s is more robust in utterance-final compared to utterance-medial position (Song et al., 2009; Theodore et al., 2010). If production of non-morphemic codas is like the production of morphemic codas, then listeners should report more codas in utterance-finally compared to utterance-medially. Previous research on word-initial segments has also shown that stops are acquired earlier than fricatives (Smit, 1993). If production of coda consonants is like the production of onset consonants, then listeners should report more stops than fricatives in coda position (Kehoe & Stoel-Gammon, 2001; Stites, Demuth, & Kirk, 2004).

**Methods**

**Subjects.** The listeners were six undergraduate students drawn from the Brown University community. All were native-English speakers with normal speech, language and hearing abilities according to self-report. All listeners had experience with phonetic transcription and acoustic analysis, which was acquired in their formal coursework concentrating in Cognitive Science and in their experiences as research assistants in the second author's laboratory.

**Stimuli.** The stimuli presented to listeners consisted of 149 utterances produced by 11 children, obtained using the procedures described above. (These same utterances were also analyzed acoustically, as presented in Experiment 2.)

**Procedure.** Listeners were directed to listen to each token and indicate whether or not a coda segment was produced in the target word. They indicated their responses by writing either “YES” to indicate coda presence or “NO” to indicate coda absence on an appropriately labeled sheet of paper. For each token, they were told both the target word and the position it occurred within the utterance. Listeners were directed to accept sound substitutions in making their coda decision. All listening took place in a quiet room, and stimuli were presented over headphones. Each listener controlled the pace of the experiment, with the constraint that each token be played no more than five times. The task took about 1 h to complete. The presentation of the stimuli was blocked by child, with the order of the children randomized across the six listeners.

**Results**

A majority agreement of four identical responses across the six listeners was used in order to determine coda presence and absence for the 149 tokens presented. A majority agreement was reached for 137 tokens (92%), with only 12 tokens (8%) split between three “coda present” and three “coda
absent” judgments. The 12 tokens that did not reach majority agreement were excluded the analyses presented below, but are considered in the Discussion section.

For the 137 tokens where there was a majority percept, mean percent codas perceived was calculated for each subject separately for utterance position, coda manner and coda voicing. Figure 1 shows mean percent codas perceived across the six listeners according for each of the three factors under consideration. Mean percent codas perceived was submitted to repeated-measures ANOVA with the factors of coda manner (stop vs. fricative), coda voicing (voiced vs. voiceless) and utterance position (medial vs. final). Results of the ANOVA showed a main effect of position $[F(1,5) = 26.90, p = 0.004; \eta^2_p = 0.842]$, with fewer codas perceived in utterance-medial (68%) compared to utterance-final position (82%). Results also showed a main effect of manner $[F(1,5) = 10.62, p = 0.022; \eta^2_p = 0.680]$, with more codas perceived for fricative targets (90%) compared to stop targets (61%). There was no main effect of voicing $[F(1,5) = 2.93, p = 0.148; \eta^2_p = 0.379]$, with coda perception roughly equivalent across voiced targets (74%) and voiceless targets (77%).

![Image of Figure 1](image.png)

Figure 1. Mean percent codas perceived according to utterance position and voicing for the fricative targets (top panel) and the stop targets (bottom panel). Error bars indicate standard deviation.
However, the ANOVA also revealed two interactions. The first was a reliable interaction between coda manner and utterance position \[ F(1,5) = 14.49, \, p = 0.013; \, \eta^2_p = 0.744 \] and the second was between coda manner and coda voicing \[ F(1,5) = 14.79, \, p = 0.012; \, \eta^2_p = 0.744 \]. To explicate the nature of these interactions, four paired comparisons were performed as post-hoc test. Following the Bonferroni adjustment to control family-wise error rate in these four post-hoc tests, the criterion for statistical significance for the paired comparisons was set to \( \alpha = 0.0125 \). Paired \( t \)-tests revealed that the effect of position is not significant for the fricative targets \[ t(5) = 0.87, \, p = 0.423 \], which were near ceiling in both utterance positions, but is significant for the stop targets \[ t(5) = -4.45, \, p = 0.007 \], with more stop codas perceived in utterance-final compared to utterance-medial position. Paired \( t \)-tests showed that the interaction between coda manner and coda voicing reflects the fact that the effect of voicing was marginal for the fricatives \[ t(5) = -3.54, \, p = 0.016 \], but not for the stop codas \[ t(5) = 1.88, \, p = 0.119 \]. Neither the interaction between position and voicing \[ F(1,5) = 1.69, \, p = 0.251; \, \eta^2_p = 0.243 \], nor the three-way interaction between manner, position and voicing was significant \[ F(1,5) = 0.833, \, p = 0.403; \, \eta^2_p = 0.143 \].

Summary

The goal of this experiment was to determine if perceptual judgments of coda presence and absence varied according to coda manner, coda voicing and utterance position. In terms of coda manner, more fricative codas were perceived compared to stop codas. This finding is opposite our prediction, in that it is generally reported that children acquire the stop class of consonants prior to fricatives. One possibility that may explain this discrepancy is that this study examines production of these sounds in coda position embedded in the context of an utterance, and developmental benchmarks are generally reported for onset position of single words. For the stop codas, more were perceived in utterance-final compared to utterance-medial position. This result does match our prediction, given that the cues to stop segments are generally more robust utterance-finally compared to utterance-medially. Furthermore, more codas were perceived for voiceless fricatives compared to voiced fricatives.

Recall that the results reported above concern the influences of position, manner and voicing on coda perception for those tokens where a consensus agreement was received with the adult listeners. One potentially interesting compliment to these data would be careful analysis of those tokens where listeners failed to agree as to the presence or absence of a coda consonant. Given the small number of qualifying tokens in the present data set, it is not possible to do such an analysis presently. However, we descriptively note the target was a stop coda for 11 of the 12 tokens where listeners failed to reach agreement. Fifty percent of the 12 tokens were voiced, and 75% occurred in medial position. None of the 12 tokens contained the primary acoustic cues predicted to influence coda perception, as described below.

In the following experiment, we examine the acoustic cues to coda production in order to explore the relationship between the acoustic information present in the children's speech and listeners' perception of coda segments. If listeners are sensitive to all of the acoustic information present in the children's utterances, then the same patterns should be observed in the acoustic analysis of the children's speech. However, if the children were implementing cues to feature contrasts that are undetected by adult listeners, then we would potentially observe differences between the acoustic cues present in the utterances and adult listeners' coda perception.

Experiment 2

The goal of Experiment 2 was to examine children's production of fricative and stop codas using acoustic analysis in order to detect cues to segmental feature contrasts. This method provides a
very fine-grained means of determining coda presence and absence, and thus provides a baseline for comparison to the data derived from the two-alternative forced-choice perception task reported in Experiment 1.

Methods

Subjects. The subjects were 11 children from the Providence, RI community, and their demographics are described above in the General Methods section.

Acoustic analysis. Each utterance was excised using Praat (Boersma & Weenink, 2010) and saved to an individual file for subsequent acoustic coding. Following conventions established by Shattuck-Hufnagel, Demuth, Hanson, and Stevens (in press), each utterance was coded for an assortment of acoustic cues to the distinctive features of the coda segments, six of which are reported in this study. The particular conventions we adopted are described below for each acoustic cue. This method is based on Stevens’ (2002) feature-cue-based model, which proposes that a given feature contrast may be signaled by many acoustic cues, and that the precise set of cues that a speaker employs may vary depending on the features in the phonemic segment, as well as on the context (e.g. segmental, structural) in which the feature occurs. This model separates the acoustic cues in the speech signal into acoustic landmarks (Halle, 1992; Stevens & Hanson, 2010) and an additional set of cues to articulator-bound features (such as voicing and place of articulation) that are found near the landmarks (Keyser & Stevens, 2006; Stevens & Keyser, 2010). For expository convenience, in this study we refer to the manner classes of our contrasting segments (i.e. stops vs. fricatives), rather than to the specific distinctive features that distinguish them.

All cues were identified by a combined visual inspection of the waveform and spectrogram, which were generated using Praat (Boersma & Weenink, 2010). The visual inspections followed highly specific criteria outlined in a coding manual developed in our lab specifically for the detection of acoustic landmarks in children’s speech. Though visual identification was the primary means of identifying the acoustic cues, listening to particular aspects of the signal was used in some cases. For example, one criterion for identifying a voice bar (described below) includes listening to the portion of the signal in question to insure that it does not have a vowel-like perceptual quality. One coder performed all acoustic measurements, and a second coder re-measured 25% of the utterances. Reliability between the two coders was 92%. Figure 2 shows representative waveforms and spectrograms illustrating the six acoustic cues of interest. The first cue, which served as the acoustic metric of fricative coda production, was the presence of high frequency, aperiodic noise following periodicity for the vowel. This cue can be seen in panel (d) of Figure 2. We expected to observe increased presence of high frequency, aperiodic noise for stop targets with fricative compared to stop codas. In fact, the presence of fricative noise for stop targets would indicate that a stop target was produced in a fricative manner.

The second and third cues, release burst and voice bar, have been established to be associated with stop consonant production and thus pertain to the targets with stop codas. A release burst results from the release of pressure generated by occluding the vocal tract for stop production while air continues to flow through the vocal folds into the mouth, and acoustically manifests as a sudden spike of transient energy. Representative examples are shown in panels (a) and (b) of Figure 2. A voice bar occurs when the oral and nasal tracts are closed for stop consonant production though the vocal folds continue to vibrate, and manifests acoustically as a simple waveform of low amplitude with little to no energy in the higher frequencies during the consonant closure (Shattuck-Hufnagel et al., in press). Panel (c) of Figure 2 displays a voice bar. Thus, both voice bar and burst are evidence that the coda stop was produced.
The fourth cue, post-release noise (also referred to as aspiration), is also an indicator of stop coda production, and results from air passing through the glottis for a short period of time following the release burst. Examples are shown in panels (a) and (b) of Figure 2. The mechanism of post-release noise is noise generated at the point of stop constriction proper and noise generated as air passes through the glottis. This presence of this cue can lead to the perception of a “released” stop, and the absence of this acoustic cue can lead to the perception of an “unreleased” stop.

The fifth cue, vowel-final noise, is the presence of noise at the end of the vowel before vocal tract closure into the coda consonant, and is shown in panel (b) of Figure 2. This cue, also called pre-aspiration, has been previously shown to be associated with stop consonant production, particularly for voiceless stop codas. The sixth cue was the presence of irregular pitch periods (IPPs), or
glottalization, at the end of the vowel. These creaky sounding pitch periods have been shown to be more frequent in utterance-final compared to utterance-medial position, which may reflect a decline in respiratory reserve at the ends of utterances (Turk & Shattuck-Hufnagel, 2007). IPPs are shown in panels (a) and (b) of Figure 2. Taken together, these six cues provide important (if not exhaustive) acoustic information regarding the possible presence of absence of a coda consonant.

Results

Across the utterances examined, frication noise was present in 44% of the tokens, with voice bar the second most prevalent cue, occurring in 23% of the tokens. Release bursts and post-release noise occurred in 14% and 15% of the tokens, respectively. The other two cues, vowel-final noise and IPPs at the end of the vowel, also occurred with relatively low incidence (11% and 14%, respectively). This general pattern of occurrence in children's speech is consistent with recent examination of the prevalence of these cues in spontaneous speech production data from children the same age (Song et al., 2011, 2012). As stated previously, our primary question is not solely the extent to which these cues are present in children’s speech, but rather the degree to which coda manner, coda voicing and utterance position influence the presence of these cues.

To this end, for each of the six acoustic cues examined, mean proportion of occurrence was calculated for each child in each of the six conditions created by crossing the independent variables of utterance position (medial vs. final), coda manner and coda voicing. A repeated-measures ANOVA with the factors of utterance position (medial vs. final), coda manner (stop vs. fricative) and coda voicing (voiced vs. voiceless) was performed on mean proportion of occurrence separately for each acoustic cue. Results for each acoustic cue are addressed in turn. The results from the ANOVAs performed for each acoustic cue are summarized in Table III. As shown in Table III, across the set of analyses there were two significant interactions that required post-hoc testing in the form of four paired comparisons; accordingly alpha for all paired comparisons was set to 0.0125 to adjust family-wise error rate across the set of analyses.

Fricative noise. Results of the ANOVA revealed a main effect of coda manner, with fricative noise present in a greater proportion of fricative targets compared to stop targets \([F(1,10) = 97.93, p < 0.001, \eta^2_p = 0.907]\), indicating that most fricative targets were realized as fricatives and not as

<table>
<thead>
<tr>
<th>Acoustic cue</th>
<th>Fricative coda</th>
<th>Release burst</th>
<th>Post-release noise</th>
<th>Voice bar</th>
<th>Vowel-final noise</th>
<th>IPPs (end of vowel)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Position</td>
<td>0.863 (0.003)</td>
<td>0.000 (0.763)</td>
<td>0.002 (0.641)</td>
<td>0.003 (0.597)</td>
<td>0.432 (0.064)</td>
<td>0.015 (0.462)</td>
</tr>
<tr>
<td>Manner</td>
<td>0.000 (0.907)</td>
<td>0.001 (0.698)</td>
<td>0.004 (0.575)</td>
<td>0.013 (0.478)</td>
<td>0.002 (0.620)</td>
<td>0.810 (0.006)</td>
</tr>
<tr>
<td>Voicing</td>
<td>0.341 (0.092)</td>
<td>0.810 (0.006)</td>
<td>0.796 (0.007)</td>
<td>0.009 (0.507)</td>
<td>0.493 (0.049)</td>
<td>0.152 (0.193)</td>
</tr>
<tr>
<td>Position × manner</td>
<td>0.096 (0.253)</td>
<td>0.000 (0.763)</td>
<td>0.002 (0.641)</td>
<td>0.557 (0.036)</td>
<td>0.432 (0.064)</td>
<td>0.211 (0.151)</td>
</tr>
<tr>
<td>Position × voicing</td>
<td>0.132 (0.212)</td>
<td>0.518 (0.044)</td>
<td>0.756 (0.011)</td>
<td>0.465 (0.055)</td>
<td>0.645 (0.022)</td>
<td>0.810 (0.006)</td>
</tr>
<tr>
<td>Manner × voicing</td>
<td>0.132 (0.212)</td>
<td>0.810 (0.006)</td>
<td>0.796 (0.007)</td>
<td>0.052 (0.327)</td>
<td>0.493 (0.049)</td>
<td>0.277 (0.119)</td>
</tr>
<tr>
<td>Position × manner × voicing</td>
<td>0.756 (0.011)</td>
<td>0.518 (0.044)</td>
<td>0.756 (0.011)</td>
<td>0.706 (0.015)</td>
<td>0.796 (0.007)</td>
<td>0.465 (0.055)</td>
</tr>
</tbody>
</table>

Notes: Values in bold indicate significance \((p < 0.05)\). Values in italics indicate marginal significance \((p < 0.10)\).
stops, and that most stops were realized as stops and not as fricatives. Results of the ANOVA showed no main effect of position \(F(1,10) = 0.03, p = 0.863, \eta_p^2 = 0.003\) or voicing \(F(1,10) = 1.00, p = 0.341, \eta_p^2 = 0.092\). In addition, there was no interaction between position and manner \(F(1,10) = 3.38, p = 0.341, \eta_p^2 = 0.253\), position and voicing \(F(1,10) = 2.69, p = 0.132, \eta_p^2 = 0.212\) or manner and voicing \(F(1,10) = 2.69, p = 0.132, \eta_p^2 = 0.212\). The three-way interaction between position, manner and voicing was not statistically reliable \(F(1,10) = 0.10, p = 0.756, \eta_p^2 = 0.011\). Notably, the lack of an interaction between manner and position, as well as manner and voicing, indicates that, for the fricative targets, the presence of fricative noise was equally robust in utterance-medial and utterance-final positions, as well as in voiced and voiceless fricatives.

**Release burst.** As expected, given that there were no occurrences of release bursts for any of the fricative targets, results of the ANOVA showed a main effect of manner, with more release bursts occurring for stop targets compared to fricative targets \(F(1,10) = 23.07, p < 0.001, \eta_p^2 = 0.698\). Results also showed a main effect of position, with more release bursts occurring utterance-finally compared to utterance-medially \(F(1,10) = 32.23, p < 0.001, \eta_p^2 = 0.763\). Moreover, there was a significant interaction between manner and position \(F(1,10) = 23.07, p < 0.001, \eta_p^2 = 0.763\), arising because there was an effect of position for the stop targets with more bursts in utterance-final compared to utterance-medial position \(t(10) = 5.68, p < 0.0001\) but not for the fricative targets (as stated above, no fricative targets contained a release burst). There was no main effect of voicing \(F(1,10) = 0.06, p = 0.810, \eta_p^2 = 0.006\), no interaction between position and voicing \(F(1,10) = 0.450, p = 0.518, \eta_p^2 = 0.044\) or manner and voicing \(F(1,10) = 0.06, p = 0.810, \eta_p^2 = 0.006\) and no interaction between position, manner and voicing \(F(1,10) = 0.45, p = 0.518, \eta_p^2 = 0.044\).

**Post-release noise.** Results from the ANOVA for post-release noise paralleled those for release burst, which is not surprising given the articulatory coupling of these two acoustic cues. ANOVA revealed a main effect of manner \(F(1,10) = 13.52, p < 0.001, \eta_p^2 = 0.575\), a main effect of position \(F(1,10) = 17.84, p < 0.001, \eta_p^2 = 0.641\) and an interaction between manner and position \(F(1,10) = 17.84, p < 0.001, \eta_p^2 = 0.641\). As with release burst, no fricative targets contained post-release noise, but there was greater prevalence of post-release noise for stops in utterance-final compared to utterance-medial position \(t(10) = 4.22, p = 0.002\). There was no main effect of voicing \(F(1,10) = 0.07, p = 0.796, \eta_p^2 = 0.007\), no interaction between position and voicing \(F(1,10) = 0.10, p = 0.756, \eta_p^2 = 0.011\) or manner and voicing \(F(1,10) = 0.07, p = 0.796, \eta_p^2 = 0.007\) and no interaction between position, manner and voicing \(F(1,10) = 0.10, p = 0.756, \eta_p^2 = 0.011\). Thus, post-release noise occurred more frequently in stop targets compared to fricative targets, and, for the stop targets, occurred more frequently in utterance-final compared to utterance-medial position.

**Voice bar.** As expected based on previous literature, ANOVA revealed a significant effect of voicing, with voice bar occurring more frequently before voiced compared to voiceless targets \(F(1,10) = 10.30, p = 0.009, \eta_p^2 = 0.507\). ANOVA also showed a main effect of manner, with more voice bars occurring before stops compared to fricatives \(F(1,10) = 9.15, p = 0.013, \eta_p^2 = 0.478\), and a main effect of position, with more voice bars occurring utterance-medially compared to utterance-finally \(F(1,10) = 14.80, p = 0.003, \eta_p^2 = 0.597\). Though there was a trend for the effect of position to be greater for stop targets compared to fricative targets, the interaction between manner and voicing was marginal \(F(1,10) = 4.86, p = 0.052, \eta_p^2 = 0.327\). There was no interaction between position and manner \(F(1,10) = 0.37, p = 0.557, \eta_p^2 = 0.036\) nor position and voicing \(F(1,10) = 0.58, p = 0.465, \eta_p^2 = 0.055\). The three-way interaction was not statistically reliable \(F(1,10) = 0.15, p = 0.706, \eta_p^2 = 0.015\).
Vowel-final noise and IPPs. These two cues were of low prevalence in the children's speech compared to the other four cues examined. For vowel-final noise, there was a main effect of manner, with vowel-final noise occurring more frequently in stop targets compared to fricative targets \[F(1,10) = 16.30, p = 0.002, \eta^2_p = 0.620\]. There was no main effect of position \[F(1,10) = 0.67, p = 0.432, \eta^2_p = 0.064\] or voicing \[F(1,10) = 0.51, p = 0.493, \eta^2_p = 0.049\]. There was no interaction between position and manner \[F(1,10) = 0.67, p = 0.432, \eta^2_p = 0.064\], position and voicing \[F(1,10) = 0.23, p = 0.645, \eta^2_p = 0.022\] nor manner and voicing \[F(1,10) = 0.51, p = 0.493, \eta^2_p = 0.049\]. Moreover, the three-way interaction was not significant \[F(1,10) = 0.07, p = 0.796, \eta^2_p = 0.007\].

For IPPs at the end of the vowel, utterance position was the only reliable effect, with increased irregularity in final compared to medial position \[F(1,10) = 8.60, p = 0.015, \eta^2_p = 0.462\]. There was no main effect of manner \[F(1,10) = 0.61, p = 0.810, \eta^2_p = 0.006\] or voicing \[F(1,10) = 2.40, p = 0.152, \eta^2_p = 0.193\]. There was no interaction between position and manner \[F(1,10) = 1.79, p = 0.211, \eta^2_p = 0.151\], position and voicing \[F(1,10) = 0.06, p = 0.810, \eta^2_p = 0.006\] nor manner and voicing \[F(1,10) = 1.32, p = 0.277, \eta^2_p = 0.119\]. Finally, the three-way interaction between position, manner and voicing was not significant for IPPs at the end of the vowel \[F(1,10) = 0.58, p = 0.465, \eta^2_p = 0.055\].

Discussion

The goal of Experiment 2 was to characterize coda production in terms of acoustic cues, and to examine whether or not the presence of these cues was systematically related to coda manner, coda voicing and utterance position. As shown in Table I, the influence of manner, voicing and position varied depending on the particular acoustic cue in question. Consider first the acoustic cues that were influenced by coda manner. Five of the six cues examined were in fact influenced by coda manner, the only exception being IPPs at the end of the vowel. For the five cues that were systematically distributed based on coda manner, fricative noise was the only cue more frequently present in fricative targets compared to stop targets, with the other four (release burst, post-release noise, voice bar and vowel-final noise) more prevalent in stop targets compared to fricative targets.

Of the six cues examined, coda voicing influenced only the occurrence of a voice bar. That this cue was more prevalent in voiced compared to voiceless codas is not surprising given previous literature (Song et al., 2011, 2012); however, the lack of a voicing effect on release burst and post-release noise is contrary to other findings indicating that these cues (in fact associated with “released stops”) are more frequent in voiceless compared to voiced stops (Demuth et al., 2009). One critical distinction between the current work and previous findings concerns the nature of the speech corpus; the current study consists of elicited imitations in an experimental setting, whereas earlier findings stemming from analysis of spontaneous speech (Song et al., 2009). Thus, this result indicating that the presence of release burst and post-release noise were equally robust across both voiced and voiceless codas may reflect more careful speech as a consequence of the elicitedimitation task.

In terms of utterance position, release burst and post-release noise occurred more frequently in utterance-final compared to utterance-medial position for the stop targets, and voice bar occurred more often utterance-medially for both stop and fricative targets. IPPs at the end of the vowel occurred more often utterance-finally than utterance-medially for both stop and fricative targets, in line with earlier research. Given previous results showing that morphemic /s/ and /z/ (used to mark the third person singular inflection) is produced more often utterance-finally compared to utterance-medially (Song et al., 2009), the lack of a positional effect on the production of non-morphemic fricatives suggests that the locus of the position effect for morphemes is syntactic in nature, and not solely a consequence of articulatory considerations.
Comparisons between coda perception and coda acoustics. Results from this experiment provide critical information on the acoustic cues present for segments produced in coda position, and thus serve as a benchmark of coda production in children. As described below, we compared these acoustic findings regarding cues to coda consonants with direct perceptual transcription in Experiment 1 in order to evaluate the perceptual salience of these acoustic cues. Some of the perceptual patterns are in line with the acoustic data presented in Experiment 2. For example, more stops were perceived in utterance-final compared to utterance-medial position (i.e. the position where the acoustic cues of release burst and post-release noise were more frequently observed). Moreover, more fricative codas were perceived compared to stops, in line with the fact that the cues most often produced for each segment (fricative noise and burst/post-release noise, respectively) were also more frequently observed in fricatives compared to stops.

However, two patterns from the perceptual data were not predicted by the acoustic data. First, more voiceless fricatives were perceived compared to voiced fricatives, which is not in accord with the fact that high frequency, aperiodic noise was equally present across both voiced and voiceless fricatives. This raises the possibility that some cue other than those measured may also play a role in the perception of fricative codas. One candidate for such a cue may be the duration of the fricative, since previous literature has shown that the duration of voiceless fricatives is longer than that of voiced fricatives, at least in adults (Stevens, Blumstein, Glicksman, Burton, & Kurowski, 1992). This longer frication duration might make voiceless fricatives easier to detect compared to their shorter voiced counterparts. To examine if this durational difference was observed in the current data set, we conducted a repeated-measures ANOVA with the factors of voicing and utterance position on mean fricative duration for those tokens that did contain fricative noise (thus analyzing utterances from 9 of the 11 children). Results of the ANOVA showed no effect of voicing on fricative duration \([F(1,8) = 0.07, p = 0.801]\), but did reveal a robust effect of utterance position \([F(1,8) = 82.57, p < 0.001]\) with frication being longer in utterance-final compared to utterance-medial position. The interaction was not reliable \([F(1,8) = 0.37, p = 0.559]\). This finding indicates that duration of the fricative alone is not guiding perceptual transcription. Specifically, although fricative durations were equal across voicing categories, the adult listeners perceived more voiceless compared to voiced fricatives, and although utterance-final fricatives were lengthened compared to utterance-medial fricatives, transcribers perceived fricatives equally as often in both utterance positions. Thus, it is likely that other acoustic characteristics (e.g. vowel duration) and their interactions with fricative duration may better explain the perceptual data than fricative duration alone.

Given the small proportion of tokens that were perceived without a coda (29/139), a comprehensive analysis of covert contrast is not possible. However, as a means of pointing to some candidates of covert contrast for future examination, we conducted one further analysis on the subset of 29 tokens that were perceived as containing no coda. Of these, only 6 were fricative targets (20%) and 23 were stop targets (80%). Half of the fricative targets perceived as not containing a coda occurred in utterance-final position, and none of these final tokens contained any cue to vocal tract closure. The other half of the fricative targets perceived as not containing a coda occurred in utterance-medial position; moreover, all were voiced targets. Interestingly, in two of these three cases, voice bar and vowel-final noise were present in the speech signal. For the 23 stop targets, 74% were in utterance-medial position and 26% were in utterance-final position; for both utterance positions, there were slightly more voiceless tokens perceived to be lacking codas than voiced tokens. For the voiced medial tokens, 57% of these contained voice bar and for the voiced utterance-final tokens, 50% contained vowel-final noise. None of the voiced utterance-final tokens contained a voice bar, which further highlights the influence of utterance position on this cue to coda production. For both voiced and voiceless utterance-final tokens, 50% contained IPPs at the end of the vowel. Given that IPPs (i.e. glottalization) are not a hallmark of coda production per se in that they can occur in the absence of a coda consonant, it is not surprising that these tokens were categorized
as without coda. This descriptive analysis suggests that voice bar and vowel-final noise, which both involve closure in the vocal tract (and therefore an intended coda), did not trigger the segmental perception in the adult listener. Thus, 55% of the tokens perceived as having no coda may actually have been cases of covert contrast, where the child had some coda consonant representation, even if this was only partially represented in their actual speech.

Conclusions

In this study, we conducted two experiments designed to examine 2-year-olds’ production of coda consonants. In the first experiment, we analyzed children’s speech by asking adult listeners to determine whether they perceived a coda consonant for particular target words. In the second experiment, we analyzed children’s speech by identifying acoustic cues associated coda production. In both experiments, we investigated the influences of coda manner, coda voicing and utterance position on the production of coda segments. Consistent with previous research on inflectional morphemes (Song et al., 2009; Theodore et al., 2010), our results indicate that many of the acoustic cues associated with coda production are influenced by utterance position. This suggests that the developmental production of cues to feature contrasts may be influenced by a host of contextual and speech planning factors. This is particularly relevant for the clinical domain in that it highlights the need to assess children’s production of speech sounds across syllable and utterance contexts. In addition, results from the adult listeners indicate that they detected many more fricative codas compared to stop codas. However, in several of the cases where listeners did not report “hearing” a coda, the child produced cues to vocal tract closure. This latter point suggests potential covert contrasts in coda consonant production, suggesting that children’s representations of syllable structure may often be more advanced than adult ears may detect.

There are two methodological considerations to note given their impact on the interpretation of the current findings. First, recall that the task presented to the adult listeners required making a two-alternative forced-choice decision regarding how they perceived the target coda. Given this binary alternative, it is possible that a different task such as open-ended transcription may have led to a different pattern of perception results. Related, the particular demographic used in the perception task may have influenced the pattern of coda perception. Though the listeners did have some experience with phonetic transcription and linguistic analysis, it was not to the same degree that is likely present in those individuals who are practicing clinicians. If the latter population were considered, it may be the case that the heightened sensitivity to acoustic characteristics of speech may have led to a better match between coda percept and the acoustic cues that were present in the children’s utterances.

Second, the particular target nouns and utterance contexts examined in this study likely play a strong role in the reported findings, particular those of utterance position. Recall that the medial context was formed such that the target noun was always followed by a stop consonant. As described previously, this frame was selected so as to provide a very difficult medial context such that positional effects would be more likely to emerge. Indeed, preliminary work has shown that the particular medial context can modulate the influence of utterance position such that plural –s production is more robust utterance-finally compared to utterance-medially when the morpheme is followed by a stop consonant (e.g. The dogs bark) (Theodore et al., 2010), but the positional effect is not observed when the morpheme is followed by a vowel (i.e. The dogs eat) (Theodore, Demuth, & Shattuck-Hufnagel, 2011). Moreover, it should be noted that the utterance contexts selected for the fricative targets consisted of coda plus following onset that occurred at the same place of articulation (e.g. His nose dripped) and some that occurred at a different place of articulation (e.g. My bus crashed). It may be the case that the homorganic contexts provided increased articulatory difficulty relative to the other context; however, to the extent that there were equal numbers of these two contexts for the voiced and

---

The image contains a page of a document, and the text is a part of an academic paper discussing the production of coda consonants in children and the perception of these consonants by adult listeners. The discussion includes experimental methods, results, and conclusions related to the study.
voiceless fricative targets, it is unlikely that this place of articulation difference is a confound to the results considered presently.

In moving forward, the results of this study provide baseline data on the development of coda production in 2-year-old children. Given that some grammatical morphemes are marked in coda position in languages such as English, these data suggest constraints on interpreting morphological competency in developmental populations. Specifically, given the robust effects of utterance position for the acoustic cues associated with stop consonants, assessing mastery of past tense –ed should consider the utterance position in which the morpheme is elicited. In addition, perception of fricative codas was influenced by utterance position, which is striking because the presence of frication noise in children’s productions was equal across medial and final utterance positions. This highlights a potential place where acoustic analysis may be more sensitive in assessing children’s productions and points to the need for further research that links segmental perception to individual acoustic cues.

Finally, this study examines the acoustic information present in children’s utterances and aims to consider how adult listeners perceive this information. The particular environment used to elicit the children’s productions was highly optimized in that they were provided with an auditory prompt to imitate; despite such an optimal environment, children reliably omitted salient acoustic cues that subsequently triggered the perception of “no coda” in adult listeners. It is well-documented in the speech production literature that adults, particularly in casual speech, often omit or “reduce” the acoustic cues of individual segments (Picheny, Durlach, & Braida, 1986). Consider the example (drawn from the current stimulus set) “My food burned”. Adults may not produce the abutting medial stops using citation cues (e.g. two separate release bursts) and in fact this context of abutting stops may be reduced completely to an elongated period of vocal tract closure. However, listeners reliably recover the segmental structure of reduced speech such that they comprehend the intended message and are not left with the perception of omitted segments. The type of reduction in children’s speech examined in this study did, in many cases, give rise to the perception of omitted segments for the adult listeners. This finding raises the possibility that the reduction observed in children’s speech is qualitatively different than reduction observed in casual adult speech, given that segmental recovery is impaired in the former but not in the latter case. Systematic examination of the acoustic information available in emerging productions vs. casually produced adult productions would be most informative.

Acknowledgments

We thank Melanie Cabral and Karen Evans for assistance with data collection and analysis, and Jae Yung Song for helpful comments on this manuscript. We sincerely thank two anonymous reviewers for their careful and thorough suggestions on the exposition of the current data and for their input on the theoretical considerations addressed in this study.

Declaration of Interest: This research was supported by NIH Grant No. R01 HD 057606.

References


White, D. (2001). *Covert contrast, merger, and substitution in children's productions of /k/ and /t/* (Unpublished master's thesis), The Ohio State University, USA.
