Infants’ Discrimination of Consonants: 
Interplay Between Word Position and 
Acoustic Saliency

Stephanie L. Archer\(^a\), Tania Zamuner\(^b\), Kathleen Engel\(^c\), Laurel Fais\(^d\) & Suzanne Curtin\(^c\)

\(^a\) Department of Psychology, University of Warwick
\(^b\) Department of Linguistics, University of Ottawa
\(^c\) Department of Psychology, University of Calgary
\(^d\) Department of Psychology, University of British Columbia

Published online: 26 Jan 2015.
Infants’ Discrimination of Consonants: Interplay Between Word Position and Acoustic Saliency

Stephanie L. Archer  
Department of Psychology, University of Warwick

Tania Zamuner  
Department of Linguistics, University of Ottawa

Kathleen Engel  
Department of Psychology, University of Calgary

Laurel Fais  
Department of Psychology, University of British Columbia

Suzanne Curtin  
Department of Psychology, University of Calgary

Research has shown that young infants use contrasting acoustic information to distinguish consonants. This has been used to argue that by 12 months infants have homed in on their native language sound categories. However, this ability seems to be positionally constrained, with contrasts at the beginning of words (onsets) discriminated earlier. This study explores whether English-learning 12- and 20-month-olds discriminate coda consonants in word-final and word-medial positions. The 12-month-old group successfully discriminated place of articulation contrasts for voiced stops in word-final position, though not voiceless stops in either position, while the older infants discriminated place of articulation contrasts for both voiced and voiceless stops in both positions. This indicates that voiced stops may be more acoustically salient than voiceless, and that position influences discrimination. Our findings support the claim that infants build speech sound categories starting with more salient contrasts in strong positions, which expand to other positions over the course of development.

Long before they can speak, infants are already learning the patterns of their native language, using input from their environment. Young infants demonstrate the ability to discriminate many of the speech sounds found in the world’s languages (for a review see Aslin, Jusczyk, & Pisoni, 1998), and over the course of the first year of life infants’ discrimination of speech sounds

Correspondence should be addressed to Stephanie L. Archer, Department of Psychology, University of Warwick, Coventry, West Midlands CV4 7AL, UK. E-mail: s.archer@warwick.ac.uk
reorganizes to match the phonetic constraints of their native language. Seminal studies in infant speech perception have found robust evidence that exposure to native language input shapes perception. These studies have shown that by the age of 6 months infants have narrowed their perception of vowel categories to focus on native contrasts (Kuhl, Williams, Lacerda, Stevens, & Lindblom, 1992; Polka & Werker, 1994), and that by the age of 10 months infants have also narrowed their perception of consonant contrasts to those present in the ambient language (Kuhl et al., 2006; Pegg & Werker, 1997; Werker & Tees, 1984; but see Best, McRoberts, & Sithole, 1988 for discrimination of click contrasts). While infants’ ability to discriminate nonnative contrasts changes over the first year of life, their ability to discriminate native contrasts becomes more robust. Indeed, in the case of liquid contrasts, English-learning infants’ discrimination improves over the course of development (Kuhl et al., 2006) as does their ability to discriminate /d/-/ð/ (Polka, Colantonio, & Sundara, 2001; Sundara, Polka, & Genesee, 2006), demonstrating that experience helps to sharpen perceptual abilities.

However, speech sound contrasts are not uniformly discriminated early in development. In the current study, we explore the effects of syllable position and voicing on infants’ perception of speech sound contrasts. Though some studies have provided evidence that infants undergo perceptual reorganization just before their first birthday, further studies have demonstrated that infants are able to detect some contrasts before others. Figure 1 diagrams the general trajectory of language-specific speech contrasts. We propose that the acoustic signatures of contrasting sounds significantly influence infant discrimination, and that the acoustic salience of the contrast is affected by the combination of position and phonetic properties.

Narayan, Werker, and Beddor (2010) found that English-learning infants of 4-5, 6-8, and 10-12 months have difficulty discriminating the less salient nasal contrast [na] – [N]a but are able to discriminate the more acoustically salient contrast [ma] – [na]. Here, salience was defined as the contributing factors that allow the perceiver to detect a contrast based on acoustic and perceptual studies with adults (Narayan, 2008). In this case, a contrast between speech sounds was considered salient if their acoustic signals were dissimilar enough for adult listeners to clearly discriminate the contrasting elements. The [na] – [N]a contrast in onset position was not acoustically dissimilar enough for even young Filipino infants, whose native language includes this contrast. Indeed, these infants do not show successful discrimination of [na] and [N]a until 10-12 months (Narayan et al., 2010). Even discrimination of vowel length might not be in place until 10 months of age (Sato, Sogabe, & Mazuka, 2010).

Then again, other types of language-specific acoustic information might be accessed earlier in infancy. Yeung, Chen, and Werker (2013) found that as early as 4 months Mandarin and Cantonese infants show sensitivity to language-specific perceptual patterns for tone. They speculate that the narrowing of the infant perceptual system to language-specific acoustic information in tones happens earlier than language-specific narrowing for vowels (Kuhl et al., 1992) and consonants (Werker & Tees, 1984) because the acoustic properties of tones (perceived as pitch and intensity) are likely more salient to a very young perceptual system (Yeung et al., 2013). In older infants, the salience of the syllable in which a contrast appears facilitates discrimination. Around the time of their second birthday, infants are able to detect changes in onsets if the target contrast is embedded in a stressed syllable. Twenty- to 24-month-olds were given two novel objects labeled with a pair of trochaic or iambic bisyllabic nonwords. A third object was introduced with the same label as one of the objects from the presentation phase. Infants were successful at detecting contrasting onset consonants within a stressed syllable, regardless of position (e.g.,
DISCRIMINATION OF CODA CONSONANTS

FIGURE 1 Diagram of the development of infants’ perception of language-specific contrasts.

1 Language-specific lexical tone contrasts by approximately 4 months (English, Mandarin, Cantonese learning infants; Yeung et al., 2013).

2 Language-specific vowel contrasts by approximately 6 months but vowel length contrasts by approximately 10 months (English, Swedish learning infants; Kuhl et al., 1992; English learning infants; Polka & Werker, 1994; Japanese learning infants; Sato et al., 2010).

3 Acoustically salient language-specific consonant contrasts by approximately 4 months but less salient language specific consonant contrasts by approximately 10 months or longer (English, Tagalog learning infants; Narayan et al., 2010; English, Japanese learning infants; Kuhl et al., 2006; English, French learning infants; Polka et al., 2001).

4 Acoustically salient language specific coda contrasts approximately by 2 months but less salient by 16 months or longer (English learning infants; Jusczyk 1977; Dutch learning infants; Zamuner, 2006).

‘pu.fajo versus ‘tu.fajo or fajo.’pu versus fajo.’tu) but not in an unstressed syllable (e.g., da.’bi versus ga.’bi or ‘le.pi versus ‘le.bi) (Floccia, Nazzi, Austin, Arreckx, & Goslin, 2011). This suggests that experience, physiology, and the acoustic salience of the contrast influence discrimination. Infants’ sensitivity to certain speech sounds may be enhanced by contributing factors in addition to exposure to their native language, such as frequency, co-occurrence of cues (such as trading relations; Repp, 1982), context (Werker & Curtin, 2005), and/or changes in auditory capacities (Werner, Fay, & Popper, 2012).

Infants’ sensitivity to speech sounds also appears to be positionally constrained. Research has shown different patterns of results for consonants in word-initial position compared to word-final position. For example, infants at 9 months are most attentive to information presented at the beginning of words (Jusczyk, Goodman, & Baumann, 1999). English-learning infants show...
a significant listening preference for lists of words that contain the same consonant-vowel (CV) portion of a CVC syllable (e.g., *fet* and *fen*), the same initial consonant of a CVC syllable (e.g., *mod* and *mib*), or a common manner of articulation (e.g., *bip* and *pok*) for the initial consonant of a CVC syllable. They, however, show no preference for CVC syllables that contain a shared VC (rhyme; e.g., *bad* and *pad*), or a shared vowel nucleus (e.g., *nid* and *sib*), indicating that 9-month-old infants are better at attending to word-onset segments than to segments in other word positions. Swingley (2005) examined the perception of segments in word-initial versus word-final position in Dutch-learning 11-month-old infants. Infants in this study preferred correctly pronounced words to mispronounced words in which the word-initial segment changed in place of articulation (POA), treating the mispronounced words as nonwords. When POA was changed for a word-final segment, the infants did not prefer the correct pronunciation to the mispronunciation, suggesting better specification of consonants in word-initial position.

Discrimination studies show mixed results. In some cases, infants also show better discrimination of consonant contrasts in word-onset position than in word-final position. Ten-month-old Dutch infants tested on their discrimination of voicing contrasts (/t/-/d/) and POA contrasts (/p/-/t/ and /p/-/k/) in both word-initial and word-final positions show discrimination of all contrasts in word-initial, but not word-final position. It is not until 16 months that infants demonstrate the ability to discriminate word-final POA contrasts (/p/-/t/), and infants at this age are still unable to discriminate word-final voicing contrasts (/t/-/d/) (Zamuner, 2006). Note though that this research was done with infants learning Dutch, in which word-final voicing contrasts are not found in the ambient language. On the other hand, other studies have demonstrated that young infants are able to discriminate certain consonant contrasts in word-final position, given acoustic differences that are robust. In a high amplitude sucking paradigm, 2-month-old English-learning infants discriminate word-final contrasts that differ on multiple dimensions such as /d/-/g/ and /m/-/g/ (Jusczyk, 1977). Additionally, 6- to 8-month-old infants demonstrate the ability, in VC syllables, to discriminate syllable-final voicing contrasts /t/-/d/ and /s/-/z/, but only when vowel length cues (one of the primary cues for English voicing) and voicing cues are both present (Eilers, Wilson, & Moore, 1977). Work by Fais and colleagues (2009) showed that 6-, 12-, and 18-month-old infants discriminate word-final singleton consonants (neek) from word-final consonant clusters (neeks). Studies with older children have shown that French-learning 20-month-olds are able to distinguish novel words differing in segmental contrasts in both word-initial and word-final position (Nazzi & Bertoncini, 2009). Together these findings suggest that, in word-final position, more robust acoustic or segmental cues can facilitate discrimination for young infants.

In a recent study testing infants of ages 8, 12 and 15 months, Wang and Seidl (2014) explicitly investigated whether sensitivity to contrastive information was dependent on syllable position. They posited two hypotheses: the Symmetrical Hypothesis, in which speech sounds in onset and coda position are equally learnable (if both are phonotactically allowable), and the Asymmetrical Hypothesis, in which syllable position influences the learnability of onsets and codas. That is, infants might process speech information in onset and coda positions at different times of their perceptual development. Eight- and 12-month-olds were randomly assigned to two groups. One group was familiarized to bisyllabic nonwords in which fricatives were embedded in the onset position of the second syllable: CVC.FVC (e.g., [bug.zid]). The other group was familiarized to nonwords with a fricative in coda position: CVF.CVC (e.g., [ruv.mag]). Both groups were tested with “legal” and “illegal” words (e.g., for the onset group, [nid.var] is legal and [buz.gan])
is illegal). Eight-month-olds in both conditions failed to learn the familiarized phonotactic pattern. However, infants of 12 months that were familiarized to the fricative in onset condition successfully learned the CVC.FVC pattern, but those familiarized to the fricative in coda position (CVF.CVC) did not learn the pattern. In a follow-up experiment, 15-month-olds demonstrated success in learning the coda pattern. These results suggest that syllable position does influence learnability as shown by the different patterns of success at the ages of 8 and 15 months, supporting Wang and Seidl’s (2014) Asymmetrical Hypothesis. These results, together with the findings of previous discrimination studies, indicate that the incremental learning of phonotactic patterns seems to be driven by the development of infants’ perceptual sensitivity for contrasts appearing in different syllable positions.

One type of information that might account for these differences in discrimination abilities is the phonetic/acoustic signature of the segments, especially when combined with syllable position (Zamuner, 2006). This information might boost (or inhibit) the salience of the contrast determining, along with the perceptual ability of the listener, whether it is discriminated word-finally (Fais, Kajikawa, Amano, & Werker, 2009; Zamuner, 2006). Segment-based differences are found in the realization of consonants, specifically for POA contrasts on stops with different voicing specifications. Consider formant transitions,¹ which provide a cue to the POA of a stop following a vowel. Generally speaking, the second formant of a vowel leading into a velar stop (such as /k/ or /g/)

![Sample spectrograms of syllables /æk/ and /æb/ illustrating F1, F2, transition, closure duration and burst.](image)

---

¹ Formants are specific frequency ranges with increased intensity that contribute to the identity of a particular vowel (i.e., /a/ has a different formant signature than /i/). Formant transitions are frequency ranges that differ at the transition point where a consonant leads into or out of a vowel. These transitions are a cue to the POA of the consonant.
rises, while it lowers when releasing into a bilabial stop (such as /p/ or /b/; see Figure 2). These transitions help the listener to determine the POA of the following stop. In the case of vowels that release into voiced stops the formant transitions are maintained longer than when the same vowels release into a voiceless stop, resulting in a potentially more sustained cue as to the POA (see Appendix A for acoustic measurements and statistical comparisons of the stimuli used in this study). This cue might allow infants to differentiate two voiced stops differing only by POA even though they are unable to make the same POA discrimination with voiceless stops.

A possibility that has not been explored in the literature is whether the position of the coda contrast shows developmental patterns in perception. Segments in syllable- and word-final position have traditionally been considered coda consonants, as they are the final segment in the unit that forms the rhyme of a syllable along with the vowel (or nucleus). Coda consonants occur not only in word-final position (e.g., the final /p/ and /k/ in tap and tack, respectively), but also word-medially, for example, as the first member of a word-internal consonant sequence that crosses a syllable boundary (e.g., the medial /p/ in captain and /k/ in doctor). While the first member of the internal consonant sequence is still a coda, contrasts that occur here have different acoustic properties than those that surface in word-final position. As with POA and voicing, there are position-based differences in the realization of segments in word-medial compared to word-final position. In a C1C2 sequence, POA cues for C1 are likely to be overwhelmed by C2 POA cues, resulting in the perception of a single place of articulation (Fujimura, Macchi, & Streeter, 1978; Ohala, 1990). For example, in a sequence such as akta, the release burst for /k/ tends to be obscured by the closure of /t/, reducing the POA cues for the /k/. In word-final position it is also possible to get utterance-final lengthening of stops (Campbell & Isard, 1991). Longer stops might include longer cues for POA and allow infants to differentiate stops in word-final position more easily than those in word-medial position. Though there is ample evidence that infants have difficulty perceiving contrasts in coda position, we know of no research that specifically examines the acoustic cues of voicing and POA in word-medial and word-final coda contrasts. Here we investigate infants’ discrimination of coda contrasts and whether the position within a word combined with voicing creates a salient contrast that can be perceived by a young perceptual system.

Across these studies, the picture that emerges is that not all sound contrasts have the same acoustic properties in all word positions. What does an infant language learner make of these differences? Growing evidence for gradience in contrast perception has implications for how we think about the acquisition of a language’s sound system and how we define “phonemes,” particularly in early language development. A phoneme is typically characterized as a contrastive, abstract element of speech in which phonetic variability within the phonemic category, for example, the differences in pronunciation associated with different word positions, are ignored (see Ladd, 2011). Equating the acquisition of a speech sound with the acquisition of a phoneme, in the strict sense, implies that the learner has a representation of the speech sound that is independent of phonetic variation. This view is consistent with Wang and Seidl’s (2014) Symmetrical Hypothesis.

---

2 The steepness of the rise or fall of the formant transition is dependent on the frequency of the second formant of the vowel itself.

3 Word-medial consonant sequences (CVC1.C2VC) in which C1 is in coda position and C2 is in onset are determined based on language-specific phonotactics. For example, doctor [dət.ər] cannot be separated into the syllables do.ctor because /kt/ cannot be an onset in English. The cluster C1.C2 must cross a syllable boundary. Our word-medial stimuli conform to this stipulation.
and predicts that infants’ ability to discriminate a sound contrast in one position carries over to other positions. However, this does not seem to be the case given the position-based differences highlighted above. How do infant learners build knowledge of the sound structures of their languages without having determined the sound categories of their native language? Research into speech sound categorization suggests that the perception of speech is derived from the acoustic signal and that a listener must make use of the variability of acoustic cues to categorize speech sounds (see Diehl, Lotto, & Holt, 2004, for review, but see Liberman & Mattingly, 1985, for an alternative view). One possibility is that at the early stages of development, learners have phonetic representations that are bound to positions (such as word-initial or word-final position), but that they do not have abstract, position-independent, phonemic representations (Pierrehumbert, 2003; Sosa & Stoel-Gammon, 2005; Werker & Curtin, 2005). The salience of the acoustic cues that underlie positionally bound phonetic representations may evolve as the perceptual system develops (Pierrehumbert, 2003).

Here we explore whether different stop contrasts in coda position are equally discriminated (POA and voicing) and whether the position of the coda, word-medial, or word-final affects discrimination. The results will help to elucidate the factors that influence discrimination of speech sound categories over the course of development.

EXPERIMENT 1

Prior studies examining discrimination of contrasts in codas have focused on word-final position and have generally found success with young infants only if those contrasts are robust. It might be the case that around one year, infants are capable of discriminating contrasts in coda position only if those contrasts are inherently more salient. It is around this age that infants begin producing their first words and have homed in on their native-language sound categories (Werker & Tees, 1984), and their perception of native-language speech sounds is enhanced (Kuhl et al., 2006).

In Experiment 1, we examine whether the voice specification of the stop contrast affects 12-month-old infants’ ability to discriminate a difference in POA in word-final position.

Method

Participants. Thirty-two participants (females = 13) 12 months of age (M = 12.48, SD = .29, range = 12–12.95) were randomly assigned to either the voiced stop condition (n = 16; 7 females) or the voiceless condition (n = 16; 6 females). All infants were from monolingual English homes. An additional nine infants were tested but excluded from the analysis due to fussiness (n = 5) or not attending to the posttest (n = 4).

Stimuli. Acoustic measurements of intensity and duration, including statistical comparisons of all stimuli used in this study, are reported in Appendix A. Auditory stimuli were produced by a female native speaker of English and recorded in a sound-attenuated booth using infant-directed speech. Tokens used were ap /æp/ and ak /æk/ (voiceless condition) and ab /æb/ and ag /æg/ (voiced condition) with 12 tokens in every trial (6 unique tokens repeated 2 times each in a semirandom order); see Appendix B for formant, VOT, and duration means of the stimuli. Tokens were paired with an unbounded checkerboard during habituation and test trials. Each
habituation trial was infant-controlled, lasting a maximum of 14 seconds. Test trials were fixed at 14 seconds in length. Tokens in each trial were separated by a 500-msec pause. Posttest consisted of an image of a waterwheel paired with the pseudoword di (/dI/).

**Apparatus.** Testing took place in a 2.74 × 1.82 m quiet, dimly lit room. Infants sat on their parent’s lap facing a screen (122 cm wide, 91.5 cm high) approximately 1.5 m away. Parents wore Bose True Noise-Canceling Headphones, over which music was played. Auditory stimuli were delivered at 65 ± 5 dB over a Bose 101 speaker located directly below the screen. The checkerboard image was projected onto the screen via a NEC LT245 projector. Infants’ looking behavior was recorded using a Sony DCRDVD92 digital video camera located directly below the screen.

Habit X 1.0 (Cohen, Atkinson, & Chaput, 2004), run on a Macintosh Power PC G5, was used to order the presentation of stimuli and collect looking time data. Both the visual and auditory stimuli were played from digitized files on the computer. The experimenter, who was blind to the auditory stimuli and type of trial, monitored the infants’ looking behavior via a closed-circuit television system from an adjacent testing room. A designated key was pressed on the computer keyboard to record infant looks.

**Procedure.** Following Werker, Cohen, Lloyd, Casasola, and Stager (1998), we used an infant-controlled visual fixation procedure. First, infants were habituated to one of the VC tokens and then were tested on their ability to discriminate that token from a different one at test. Each trial began when the infant fixated on the screen displaying a rotating blue animated flower. During habituation, infants heard one of the VC tokens paired with the unbounded checkerboard. The trial ended immediately when the infant looked away. Infants were either habituated to ap or ak (voiceless condition) or to ab or ag (voiced condition). Habituation looking times were calculated online (by the Habit software; Cohen et al., 2004) and set at a habituation criterion of 60%. Infants received a minimum of six and maximum of 24 habituation trials. Once the habituation criterion was reached (i.e., looking time drops to 60% of the longest previous looking time over three trial blocks), infants were presented two test trials, one in which they heard the same stimuli that were presented during habituation (same trial), and one in which they heard the syllable not presented during habituation (switch trial). Test trial order was randomized. In a post-test trial, infants were presented with di and the waterwheel.

**Results**

The results of a 2 (trial: same, switch) by 2 (condition: voiced, voiceless) mixed model ANOVA resulted in no significant main effect of trial, $F(1,30) = .14, p = .78$, but a significant trial by condition interaction, $F(1,30) = 4.7, p = .038, \eta^2_p = .136$. To explore the source of the interaction we ran planned pairwise comparisons (two-tailed) and found that infants showed discrimination of the contrast (same: $M = 4.35$, $SD = 2.56$; switch: $M = 5.72$, $SD = 2.99$) in the voiced condition, $t(15) = -2.87, p = .012, d = .492$, but not in the voiceless condition (same: $M = 5.1$, $SD = 2.93$; switch: $M = 4.13$, $SD = 3.14$), $t(15) = 1.00, p = .333$ (Figure 3).

These findings suggest that in word-final position, voicing supports the detection of a place of articulation contrast. Previous research found coda discrimination only by 16 months of age.
FIGURE 3 12-month-olds’ discrimination of word-final (Experiment 1) and word-medial (Experiment 2) codas. Asterisk denotes $p < .05$. Bars denote standard error (SE).

(Zamuner, 2006). However, we found that 12-month-old infants are able to detect a contrast in coda position for voiced stops. The added saliency of the voiced consonants may allow infants of this age to discriminate contrasts in coda position. To explore whether this is the case, we examined the formation transitions in our stimuli to determine whether or not the transitions were more robust in the voiced stops. Our acoustic analyses (see Appendix A) show that the formant transitions in the stimuli containing voiced word-final coda stops ($ab$, $ag$) have longer duration than their voiceless counterparts ($ap$, $ak$). Statistical analysis using a one-way ANOVA with coda ($ab$, $ag$, $ap$, $ak$) as an independent variable showed a significant difference in formant transition duration (in seconds ($s$)), $F(3,16) = 17.12, p < .001, \eta^2 = .85$. Planned comparisons between voiced and voiceless codas were significant, with longer formant transitions for voiced word-final coda stops ($ab: M = .037s, SD = .012; ag: M = .062s, SD = .012$) than voiceless ($ap: M = .016s, SD = .009; ak: M = .030s, SD = .010), $ps < .05$. A difference was also found between the voiced VC syllables $ab$-$ag$ ($p < .05$), but not the voiceless syllables $ap$-$ak$ (see Appendix A). Additionally, intensity of the formant transition drops more quickly in voiceless codas than in voiced (where intensity is measured in decibels (dB) from the onset of the formant transition to the burst). A similar analysis using the mean of the intensity of this portion of the syllable, with coda remaining as an independent variable, showed a main effect, $F(3,16) = 19.899, p < .001, \eta^2 = .79$. Planned comparisons showed lower intensity (dB) for voiceless codas ($ap: M =$
FIGURE 4 Representative samples of intensity contours for voiced (ab, ag – solid lines) and voiceless (ap, ak – dotted lines) word-final codas (Experiment 1) and voiced word-final (ab, ag – solid lines) and word-medial (abta, agta – dashed lines) codas (Experiment 2). Measurements of intensity begin 60 ms before the onset of the closure duration to the burst. Black lines denote labial stops (/p/, /b/). Grey lines denote velar stops (/k/, /g/).

42.1 dB, $SD = 2.40$; ak: $M = 42.6$ dB, $SD = 1.77$) than voiced (ab: $M = 51.9$ dB, $SD = 3.64$; ag: $M = 53.8$ dB, $SD = 3.94$) ($ps < .05$), illustrated in Figure 4 (Experiment 1; see also Appendix A). This is consistent with the suggestion that word-final voiced codas are more salient than word-final voiceless codas due to the increased duration and intensity of the formant transitions for the voiced stops.
DISCRIMINATION OF CODA CONSONANTS

EXPERIMENT 2

In this experiment, we presented infants with word-medial POA contrasts using voiceless consonants /p/ and /k/ and voiced consonants /b/ and /g/ in the same discrimination task. Our comparison of voiced and voiceless stops in coda position (Appendix A, and see discussion below) showed that voiced stops, in word-medial coda position, are not more salient than voiceless stops occurring in either word-medial or word-final position. Further, voiced stops in word-final position are more salient than voiceless stops in word-final position. Thus, we might expect that, because both voiced and voiceless stops are less salient in medial coda position, infants at this age will have difficulty detecting a contrast in this position.

Method

Participants. A total of 32 (females = 17) 12-month-old infants (\(M = 12.92, SD = 0.26\), range = 11.97–12.92) participated in this study. All infants were from monolingual English homes. Infants were randomly assigned to one of two groups (voiced, \(n = 16, 10\) females; voiceless \(n = 16, 7\) females) for counterbalancing purposes. Eight additional infants were tested but were removed from analysis due to fussiness (\(n = 4\)), or looks of less than a second during the test trials (\(n = 4\)).

Stimuli. The /æbta/ and /ægta/ (voiced condition) and /æpta/ and /ækta/ (voiceless condition) stimuli were produced in infant-directed speech by a female native-English speaker and recorded in a sound-attenuated booth. Main stress was produced on the first syllable. Each trial contained 10 tokens (5 unique tokens each repeated once semirandomly) for a maximum of 14 seconds. See Appendix B for the acoustic measurements of the stimuli (calculated using the 5 unique tokens).

Apparatus. Same as Experiment 1.

Procedure. Same as Experiment 1.

Results

The results of a 2 (trial: same, switch) by 2 (condition: voiced, voiceless) mixed model ANOVA resulted in no significant main effect of trial, \(F(1,30) = .079, p = .78\), nor any significant trial by condition interaction, \(F(1,30) = .182, p = .672\). These findings suggest that infants of 12 months cannot discriminate word-medial coda contrasts regardless of their voicing specification (voiced: same \(M = 4.88, SD = 1.91\); switch \(M = 4.79, SD = 3.02\), voiceless: same \(M = 5.09, SD = 3.39\); switch \(M = 5.56, SD = 2.33\); see Figure 3).

Comparisons of the acoustic properties of word-final and word-medial codas reveal differences similar to those between the voiced and voiceless word-final coda stimuli. Measurements and results of statistical comparisons of all of the stimuli (shown in Appendix A) demonstrate that formant transitions for voiced labial and voiced velar codas are longer in duration when in word-final position than word-medial position, though voiceless word-medial and word-final codas are
of relatively equal duration. Word-medial codas, analyzed using a one-way ANOVA with coda (abta, agta, apta, akta) as an independent variable, showed no differences in formant transition duration when voiced, \( F(3,16) = 1.808, p = .186 \) (abta: \( M = .022s, SD = .013 \); agta: \( M = .033s, SD = .013 \)) or voiceless (apta: \( M = .021s, SD = .009 \); akta: \( M = .018s, SD = .010 \)). Additionally, in measurements from the onset of the formant transition to the end of the burst, intensity drops off earlier in word-medial voiced codas than word-final, as shown in Figure 4. A one-way ANOVA of mean intensity showed that the difference in this drop-off between voiced and voiceless medial codas is not significant, \( F(3, 16) = 1.475, p = .259 \) (apta: \( M = 42.2 \text{ dB}, SD = 2.40 \); akta: \( M = 44.9 \text{ dB}, SD = 2.44 \); abta: \( M = 44.4 \text{ dB}, SD = 2.73 \); agta: \( M = 41.7 \text{ dB}, SD = 3.88 \)); however, word-final and word-medial coda comparisons do show significant differences but in voiced coda pairs only (see Appendix A for individual analyses of coda pairs). This is consistent with the suggestion that word-final codas are more salient than word-medial codas, based on the duration and intensity of the formant transitions.

We next examined whether infants of 20 months can discriminate the voiced word-medial POA contrast. Since infants of 12 months succeed with only the voiced contrast, and only in word-final position, we examined whether older infants could detect the voiced contrast in word-medial position as well.

**EXPERIMENT 3**

We tested 20-month-olds on their ability to discriminate a voiced contrast in word-medial position. Infants of 12 months are unable to discriminate a voiced coda contrast in word-medial position, though they are successful at discriminating the same contrast word-finally. Could difficulty detecting contrasts in word-medial coda position be overcome by infants who have had a few more months of experience with the native language?

**Method**

**Participants.** Sixteen 20-month-old infants (\( M = 20.49, SD = .35, \text{ range} = 20–20.92 \)) participated in this study (females = 9). All infants were from monolingual English homes. Eleven additional infants were tested but were removed from analysis due to technical error (\( n = 4 \)), looking times more than 2 SDs from the mean (\( n = 2 \)), possible language delay (\( n = 1 \)), and fussiness (\( n = 4 \)).

**Stimuli.** We used the same stimuli as in the voiced condition in Experiment 2 (abta /æbta/ and agta /ægta/).

**Apparatus.** Same as Experiments 1 and 2.

**Procedure.** Same as Experiments 1 and 2, except for a difference in the habituation phase for this age group. In order to maintain the 20-month-olds’ attention, we used a fixed habituation (see Werker et al., 1998, for similar habituation settings between experiments).
DISCRIMINATION OF CODA CONSONANTS

8.0000
6.0000
4.0000
2.0000
0.0000

Mean Looking Time (seconds)

Experiment 3
Experiment 4

FIGURE 5 Twenty-month-olds' discrimination of voiced word-medial (Experiment 3) and voiceless word-medial and word-final (Experiment 4) codas. Asterisk denotes $p < .05$. Bars denote standard error (SE).

Results

The results of a two-tailed t-test revealed that infants looked significantly longer during the switch trial ($M = 6.79$, $SD = 3.91$) than during the same trial ($M = 4.43$, $SD = 1.6$), $t(15) = 2.32$, $p = .035$, $d = .790$. Thus, by 20 months, infants are able to detect a voiced POA contrast in word-medial coda position (see Figure 5).

Given that by 20 months infants can succeed in the voiced condition, we next explored whether they could also now detect a voiceless POA contrast in word-final and word-medial coda position.

EXPERIMENT 4

Twenty-month-old infants were tested on their discrimination of the POA /p/-/k/ contrast in both word-medial (apta-akta) and word-final (ap-ak) coda positions. Because 20-month-old infants could succeed in discriminating voiced stops in the less salient word-medial coda position, it may be the case that they rely less on acoustic salience than did the younger infants in Experiments 1 and 2. If that is the case, 20-month-old infants may be able to succeed in discriminating voiceless stops in both positions.
Method

Participants. Sixteen 20-month-old infants ($M = 20$ months, $SD = 0.29$, range $= 19.52–21.03$) participated in this study (females $= 8$). All infants were from monolingual English homes. Infants were randomly assigned to one of two groups for counterbalancing purposes (medial, $n = 8$, 4 females; final, $n = 8$, 4 females). Twenty-one additional infants were tested but were removed from analysis due to: not completing the experiment ($n = 5$), experiment error ($n = 1$), fussiness ($n = 4$), no habituation ($n = 3$), parental interference ($n = 5$), out of screen view ($n = 2$), and distraction ($n = 1$).

Stimuli. The stimuli were the same as in Experiments 1 and 2. However, only the voiceless stops were used.

Apparatus. In this experiment, testing took place in a different laboratory with an equipment configuration functionally identical to that used in Experiments 1 to 3, though the actual components were different. Testing took place in a $2.8 \times 2.3$ m quiet, dimly lit room. Infants sat on their parent’s lap facing a 32-in. NEC MultiSync V321 LCD monitor approximately 1.2 m away. Parents wore Peltor HTB 79A-02 headphones, over which music was played. Auditory stimuli were delivered at $65 \pm 5$ dB over two Bose Roommate speakers located on either side of the monitor. Infants’ looking behavior was recorded using a Canon ZR950 digital video camera located directly below the screen. Habit 2000 (Cohen, Atkinson, & Chaput, 2000) was used to order the presentation of stimuli and collect looking time data.

Procedure. The same procedure was used as in Experiment 3.

Results

The results of a 2 (trial: same, switch) by 2 (position: medial, final) mixed model ANOVA showed a significant effect of trial type ($F(1,14) = 4.77$, $p = .047$, $\eta_p^2 = .25$), but no significant interaction between trial type x word position ($F(1,14) = .07$, $p = .79$) and no significant main effect of word position ($F(1,14) = 2.66$, $p = .13$). Infants looked longer to the switch trials (Word-Final: $M = 8.32$, $SD = 3.73$; Word-Medial: $M = 5.60$, $SD = 3.35$; Combined: $M = 6.96$, $SD = 3.54$) than to the same trials (Word-Final: $M = 7.13$, $SD = 4.61$; Word-Medial: $M = 4.08$, $SD = 3.14$; Combined: $M = 5.61$, $SD = 3.88$), indicating that they were able to perceive the contrasts in both word-medial and word-final position (see Figure 5). Additionally, 14 of the 16 infants looked longer during the switch trial (8 of these infants looked over 1 second longer at the switch than the same trial). This suggests that by 20 months infants can successfully discriminate POA contrasts in both word-final and word-medial coda position.

DISCUSSION

This aim of this study was to determine if and when infants can perceive POA contrasts in coda positions. Studies on the effect of syllable position have shown that infants have difficulty
learning (Wang & Seidl, 2014) and discriminating (Zamuner, 2006) consonant contrasts when target speech information is in coda position. Here, we have demonstrated that infants in the early stages of language-specific phonological understanding perceive consonant contrasts differently, depending upon the salience of both the contrast and its position in the word in which it appears. That is, 12-month-old infants can perceive voiced stops in word-final coda position but not voiceless stops and not word-medial stops regardless of voicing. Thus, it is also the case that not all contrasts in coda position are perceived equally. We argue that, given our acoustic analysis of the stimuli, the combination of position and voicing can boost the acoustic salience of some contrasts so that they are easier to perceive.

Our exploration of older infants’ ability to discriminate the voiced POA contrast word-medially and the voiceless POA contrast in both coda positions revealed that by 20 months infants can discriminate both voiced and voiceless contrasts in both word-final and word-medial coda positions. Our analyses of the position-sensitive acoustic differences in the stop consonants coupled with these findings suggest that the position and the acoustic salience of the contrast influence perception early in development. Indeed, 20- to 24-month-olds can distinguish word-initial and word-medial onsets only when their carrier syllables are stressed (Floccia et al., 2011). Further, our results are consistent with the conclusion that young infants are using phonetic perception, rather than phoneme-level perception, to discriminate contrasts. Also, in our experiments, we presented infants with pairs of stop consonants (p, b, k, g), which represent only one type of consonant, potentially creating a problem for the generalizability of our findings. Our stimuli were chosen specifically because acoustic cues to POA in stops are short in duration and thus can be difficult to detect in positions in which those cues are even further reduced (e.g., Zamuner, 2006, or see Stager & Werker, 1997). Though the contrasts presented to the infants were limited in variety, our behavioral data supports our position that acoustic salience has an influence on the development of the perception of speech. Further studies must continue to determine whether different contrasts perform similarly, the influence of other cues, and the effect of acoustic salience on a wider range of contrasts and environments.

Though the findings of this study contribute to the evidence that by 12 months infants appear to perceive speech sounds phonetically, we cannot make a specific claim regarding whether 20-month-olds perceived the contrastive information in codas at the phonetic or phonemic level. However, our results do show that the acoustic differences between contrasts do not disrupt the ability of 20-month-old infants to discriminate stops in coda position, as they do for 12-month-old infants. In addition to having more experience with native language contrasts, 20-month-olds also have linguistic knowledge beyond the sound level. Particularly, the development of a lexicon leads to a more sophisticated understanding of language-specific linguistic processes (see Swingley, 2009, for review).

Our findings demonstrate that infants are continually learning about their sound categories over the course of development. While perception of contrasts in onset position appears to be native-like by the end of the first year of life, discrimination of contrasts that are less salient by virtue of occurring in other word positions takes longer to emerge. It seems plausible that the perception of less salient contrasts should appear later in development (see Yeung et al., 2013). For instance, 11-month-olds are more sensitive to consonant switches in stressed syllables than unstressed syllables, consistent with the claim that unaccented syllables are less salient (Vihman, Nakai, DePaolis, & Hallé, 2004). Salience also appears to influence word-object mappings. In novel word-object associative learning tasks, infants are sensitive to height distinctions
but not backness in vowels at 15 months (Curtin, Fennell, & Escudero, 2009). Indeed, there is more acoustic energy in the first formant (F1), which is an indicator of vowel height in an acoustic space, than in the other formants (i.e., the second formant (F2), which corresponds with backness) (Lacerda, 1993), suggesting that salience might influence performance.

There are several reasons why young infants would be more successful at discriminating place contrasts of voiced consonants than those of voiceless ones. One possibility is that, across languages, syllable onsets with low sonority (i.e., resonance of a speech sound) segments are preferred. In coda position, higher sonority segments are preferable following a syllable nucleus (Clements, 1990). Speech sounds can be arranged on a sonority scale with vowels being the most sonorous and stop consonants being the least sonorous (Selkirk, 1984). This scale lists phonetic segments in terms of their relative resonance in relation to other speech sounds, which has phonological implications for the types of consonants preferable in onset and coda positions. Within the stop consonants, voiced stops are more sonorant than voiceless ones, so it may be that infants attend more to higher sonority coda segments than to lower sonority ones, making it easier to discriminate voiced stops than voiceless stops word-finally. This is not confounded with frequency in English, as voiced stops are less frequent than voiceless stops (Zamuner, Gerken, & Hammond, 2005).

Consistent with our results is the possibility that the ability to discriminate voiced stops in coda position can be attributed to acoustic factors rather than phonological ones. In coda position, voiced stops are typically characterized by shorter release times and longer preceding vowels than voiceless stops (Wright, 2004). While our study controlled for vowel length preceding the voiced consonant, our acoustic analysis of the tokens used demonstrates that the transitions from the vowel into the stop closure are more robust in voiced stops. These formant transitions provide information about the place of articulation of a stop. Indeed, the longer duration and greater intensity of the formant transitions contribute to making the voiced word-final codas more salient than the voiceless codas, rendering the former distinction easier for infants to perceive. This is consistent with other findings in the literature demonstrating that perceptual saliency may influence infants’ ability to discriminate contrasting acoustic information. Recall, for example, that young infants have difficulty with acoustically similar nasal contrasts such as [na]-[ŋa] even prior to the perceptual reorganization of consonants at 12 months, but can discriminate the more salient contrast [ma]-[ŋa] (Narayan et al., 2010).

The emerging picture from studies on positional constraints, and the contribution of the findings of this study, is that infants first acquire phonetic categories of segments that differ positionally (e.g., Pierrehumbert, 2003; Saffran, 2003; Werker & Curtin, 2005; Zamuner, 2006) instead of abstract phonemic categories of segments independent of variation or position. Researchers have proposed that available statistical information in the speech signal is constrained by an infant’s perception of the phonetic properties of a contrast. For example, the Constrained Statistical Learning (CSL) hypothesis maintains that constraints that accommodate natural human auditory processes are necessarily applied to a statistical learning mechanism (Saffran, 2003). Similarly, Pierrehumbert (2003) postulates that infants build their initial categories by tracking statistical information given their current stage of perceptual development. Werker and Curtin (2005) further this argument by hypothesizing that the differences in the raw acoustic salience between contrastive speech sounds can be dampened or enhanced based on the information gathered through language experience and the development of representations. Indeed, work exploring infants’ ability to track distributional information has demonstrated
that infants’ experience with the input shapes their perception (Maye, Werker, & Gerken, 2002; Werker, Yeung, & Yoshida, 2012).

These predictions are consistent with other evidence suggesting that phonemic representations encompassing phonetic variation may not appear earlier than 18 months. By this age infants can use perceptual attunement to phonetic patterns in native speech to form a higher-order phonemic organization of those patterns and start to use these more abstract categories to identify sounds and word forms across surface variation (Best, 1994). In a study using accented speech, Best, Tyler, Gooding, Orlando, and Quann (2009) found that 15-month-olds can only recognize words that are restricted to native English pronunciations, whereas 19-month-olds can also recognize Jamaican-accented speech as possible pronunciations of familiar words, suggesting that the older infants have acquired constancy across variations in speech sounds. This ability seems to be facilitated by lexical acquisition such that increased expressive vocabulary scores are correlated with tolerance for variation in pronunciation (Best, Tyler, Kitamura, Notley, & Bundgaard-Nielsen, 2008). These findings, in addition to those indicating that perception of consonants differs with position, suggest that context-independent phonemic representations may not emerge before 19 months. And indeed, infants of 20 months had no difficulty in our study discriminating voiced and voiceless stop contrasts in word-final or word-medial coda positions. This is likely due to language experience and the ability to home in on more acoustically difficult contrasts as the infant learner gains greater native language knowledge and attends less to non-native, or non-distinctive, information. The current evidence, as a whole, suggests that experience with the ambient language, a growing lexicon, and emergent sound categories facilitate discrimination of speech sounds in less robust positions. However, it remains to be seen whether infants of this age treat word-initial, word-medial and word-final speech sounds as members of the same category.

According to the Processing Rich Information from Multidimensional Interactive Representations (PRIMIR) framework (Curtin, Byers-Heinlein, & Werker, 2011; Werker & Curtin, 2005), speech processing relies on three filters: initial biases, developmental level, and task demands, which can enhance or diminish the saliency of information in the auditory signal. These filters act upon the rich speech information that the developing infant hears and organizes it along several dimensions, resulting in differences in performance on tasks depending upon, for example, developmental level. At 12 months infants may not be at a level where they can abstract across all positions and all instances of a sound, resulting in their construction of phonetic categories instead of abstract “phonemic” ones. Over time, infants increase their tolerance for variation in speech sounds and by 18 or 19 months may begin to develop representations of sounds as emergent speech sound categories.

Overall the results of this study suggest that the perception of consonant contrasts is positionally constrained, and that there seems to be a perceptual advantage for voiced stop contrasts over voiceless stop contrasts in word-final coda position for younger infants. This advantage may result from longer and more salient formant transitions in voiced consonants, which provide better cues to difference in POA. We argue that these results suggest that at the earliest stage of development, learners do not exhibit evidence of having representations of abstract phoneme categories. If young learners had context-free or position-free representations, we would have expected them to discriminate the contrasts regardless of word position. Instead, at 12 months, infants demonstrated the ability to perceive POA contrasts only on voiced stop consonants in word-final position, and it is not until later in development that they can discriminate voiced
and voiceless stop contrasts in a variety of word positions. As infants gain experience with their language, their perception of native language contrasts is more focused (Kuhl et al., 2006) and they are able to detect finer distinctions between contrasts.

Our results provide crucial evidence for the claim that infants around 1 year old are influenced by position-varying acoustic information in their discrimination of coda stops, and thus do not discriminate these sounds phonemically. By 20 months, however, infants can draw on greater language experience and sophistication to recognize distinctions that transcend position-dependent phonetic variation. This is a part of the developmental change that supports the impressive trajectory of language learning that takes off as infants near the end of their second year of life.

ACKNOWLEDGMENTS

The authors thank the parents and infants who participated and especially thank Janet Werker for her invaluable input and support.

FUNDING

This research was supported by funding from Social Sciences and Humanities Research Council of Canada (SSHRC) awarded to S. Curtin (grant no. 435-2012-0124) and T. S. Zamuner (grant no. 410-2011-1961).

REFERENCES


Downloaded by [University of Ottawa] at 11:54 28 January 2015


APPENDIX A

TABLE A1
Means of Acoustic Measurements of Codas Used in Experiments 1-4 (Standard Deviation in Parentheses)

<table>
<thead>
<tr>
<th>Stimulus</th>
<th>Intensity of steady state of vowel (dB)</th>
<th>Duration from onset of ft to onset of burst (s)</th>
<th>Mean intensity of onset of ft to onset of burst (dB)</th>
<th>Closure duration (s)</th>
<th>Duration of ft (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Voiced word-final</td>
<td>ab</td>
<td>73.5 (1.95)</td>
<td>.096 (.015)</td>
<td>51.9 (3.64)</td>
<td>.065 (.012)</td>
</tr>
<tr>
<td></td>
<td>ag</td>
<td>71.8 (2.02)</td>
<td>.117 (.008)</td>
<td>53.8 (3.94)</td>
<td>.054 (.006)</td>
</tr>
<tr>
<td></td>
<td>Voiceless word-final</td>
<td>ap</td>
<td>63.3 (2.42)</td>
<td>.125 (.021)</td>
<td>42.1 (2.40)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ak</td>
<td>63.6 (2.92)</td>
<td>.150 (.025)</td>
<td>42.5 (1.77)</td>
</tr>
<tr>
<td></td>
<td>Voiced word-medial</td>
<td>abta</td>
<td>74.0 (3.15)</td>
<td>.092 (.012)</td>
<td>44.4 (2.73)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>agta</td>
<td>69.8 (1.43)</td>
<td>.137 (.041)</td>
<td>41.7 (3.88)</td>
</tr>
<tr>
<td></td>
<td>Voiceless word-medial</td>
<td>apta</td>
<td>69.8 (1.43)</td>
<td>.130 (.019)</td>
<td>42.2 (2.40)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>akta</td>
<td>69.2 (1.90)</td>
<td>.096 (.015)</td>
<td>44.9 (2.44)</td>
</tr>
</tbody>
</table>

Note. ft = formant transition, dB = decibels, s = seconds. All measurements based on means of five tokens.

TABLE A2
Statistical Comparisons (Independent Samples T-test) of Acoustic Analyses of Stimuli Used in Experiments 1 to 4

<table>
<thead>
<tr>
<th>Stimulus Comparison</th>
<th>Formant Transition Duration (s)</th>
<th>Mean Intensity (formant transition to burst) (dB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ap-ak</td>
<td>( t(8) = -2.341, p = .047 ) (n.s.)</td>
<td>( t(8) = -.311, p = .764 ) (n.s.)</td>
</tr>
<tr>
<td>ab-ag</td>
<td>( t(8) = -3.455, p = .009, r^2 = .6^* )</td>
<td>( t(8) = -.768, p = .465 ) (n.s.)</td>
</tr>
<tr>
<td>ap-ab</td>
<td>( t(8) = -3.204, p = .013, r^2 = .6 )</td>
<td>( t(8) = -5.031, p = .001, r^2 = .8^* )</td>
</tr>
<tr>
<td>ak-ag</td>
<td>( t(8) = -4.809, p = .001, r^2 = .7^* )</td>
<td>( t(8) = -5.812, p &lt; .001, r^2 = .8^* )</td>
</tr>
<tr>
<td>apta-akta</td>
<td>( t(8) = .506, p = .627 ) (n.s.)</td>
<td>( t(8) = -1.762, p = .116 ) (n.s.)</td>
</tr>
<tr>
<td>abta-agta</td>
<td>( t(8) = -1.367, p = .209 ) (n.s.)</td>
<td>( t(8) = 1.267, p = .241 ) (n.s.)</td>
</tr>
<tr>
<td>ap-akta</td>
<td>( t(8) = -2.000, p = .846 ) (n.s.)</td>
<td>( t(8) = -1.309, p = .227 ) (n.s.)</td>
</tr>
<tr>
<td>akta-agta</td>
<td>( t(8) = -2.155, p = .063 ) (n.s.)</td>
<td>( t(8) = 1.590, p = .151 ) (n.s.)</td>
</tr>
<tr>
<td>ap-akta</td>
<td>( t(8) = -1.370, p = .410 ) (n.s.)</td>
<td>( t(8) = .057, p = .956 ) (n.s.)</td>
</tr>
<tr>
<td>ab-abta</td>
<td>( t(8) = 1.876, p = .096 ) (n.s.)</td>
<td>( t(8) = 3.732, p = .006, r^2 = .6^* )</td>
</tr>
<tr>
<td>ak-akta</td>
<td>( t(8) = 1.902, p = .094 ) (n.s.)</td>
<td>( t(8) = -1.755, p = .117 ) (n.s.)</td>
</tr>
<tr>
<td>ag-agta</td>
<td>( t(8) = 3.798, p = .005, r^2 = .6^* )</td>
<td>( t(8) = 4.902, p = .001, r^2 = .8^* )</td>
</tr>
</tbody>
</table>

Note. Planned comparisons subject to Bonferroni correction \( (p = .0125) \), n.s. denotes non-significant results, * denotes significant results (Bonferroni correction). dB = decibels, s = seconds. Results show that intensity (dB) in voiced word-final coda syllables \((ab, ag)\) is significantly higher between the onset of the formant transition to the burst than voiceless word-final or voiced word-medial syllables.
APPENDIX B

Means of Formants, VOT, and Duration for Stimuli Used in Experiments 1 to 4 (Standard Deviation in Parentheses)

<table>
<thead>
<tr>
<th>Stimulus</th>
<th>$F1^a$ (Hz)</th>
<th>$F2^a$ (Hz)</th>
<th>$F2$ Offset$^b$ (Hz)</th>
<th>VOT (ms)$^c$</th>
<th>Duration$^d$ (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>apta</td>
<td>818 (106)</td>
<td>1545 (77)</td>
<td>1484 (155)</td>
<td>14 (8)</td>
<td>0.467 (0.103)</td>
</tr>
<tr>
<td>akta</td>
<td>800 (62)</td>
<td>1639 (99)</td>
<td>1846 (30)</td>
<td>30 (17)</td>
<td>0.419 (0.076)</td>
</tr>
<tr>
<td>ap</td>
<td>802 (46)</td>
<td>1649 (41)</td>
<td>1562 (68)</td>
<td>51 (40)</td>
<td>0.549 (0.089)</td>
</tr>
<tr>
<td>ak</td>
<td>825 (63)</td>
<td>1606 (177)</td>
<td>1902 (112)</td>
<td>126 (35)</td>
<td>0.534 (0.063)</td>
</tr>
<tr>
<td>abta</td>
<td>955 (91)</td>
<td>1554 (180)</td>
<td>1483 (167)</td>
<td>16 (8)</td>
<td>0.266 (.024)</td>
</tr>
<tr>
<td>agta</td>
<td>926 (153)</td>
<td>1734 (516)</td>
<td>1938 (697)</td>
<td>5 (3)</td>
<td>0.363 (.028)</td>
</tr>
<tr>
<td>ab</td>
<td>901 (172)</td>
<td>1749 (90)</td>
<td>1586 (110)</td>
<td>51 (19)</td>
<td>0.509 (.035)</td>
</tr>
<tr>
<td>ag</td>
<td>1071 (79)</td>
<td>2081 (103)</td>
<td>2446 (98)</td>
<td>75 (18)</td>
<td>0.461 (.037)</td>
</tr>
</tbody>
</table>

Note. $F1 = \text{first formant}$; $F2 = \text{second formant}$; $\text{VOT} = \text{voice onset time}$; Offset = position of formant transition into consonant (Hz). Analysis of 5 unique tokens in each set of stimuli.

$^a$ Formant measured at steady state of preceding vowel.

$^b$ Formant Offset measured at end of formant transition into coda consonant.

$^c$ VOT measured including release burst.

$^d$ Duration of target syllable.