Six- and ten-month-old infants' perception of non-contrastive variation

Dena Krieger, and Elizabeth K. Johnson

Citation: Proc. Mtgs. Acoust. **19**, 060072 (2013); doi: 10.1121/1.4799205
View online: https://doi.org/10.1121/1.4799205
View Table of Contents: https://asa.scitation.org/toc/pma/19/1
Published by the Acoustical Society of America

Proceedings of Meetings on Acoustics

Volume 19, 2013 http://acousticalsociety.org/

ICA 2013 Montreal
Montreal, Canada
2 - 7 June 2013

Speech Communication
Session 2aSC: Linking Perception and Production (Poster Session)

2aSC28. Six- and ten-month-old infants' perception of non-contrastive variation

Dena Krieger* and Elizabeth K. Johnson

***Corresponding author's address: Psychology, University of Toronto, CCT 4019, Mississauga, L5L 1C6, Ontario, Canada, dena.krieger@mail.utoronto.ca**

Recent evidence suggests that infants do not perceive all existing speech sounds from birth. For example, the velar and alveolar nasal place contrasts are so subtle that infants require experience to perceive it (e.g., Narayan et al., 2012). Here, we examine English-learning infants' perception of another subtle contrast: pre-voicing on stop consonants. Six- and ten- month-olds' ability to discriminate between voiced and voiceless stops (phonemically contrastive in English) as well as voiced and pre-voiced stops (allophonic in English, but contrastive in other languages such as Dutch) was tested using a variant of the Stimulus Alternation Paradigm (SAPP). Six-month-olds (N=34) distinguished between voiced and voiceless stops ($p < .05$), but not between voiced and pre-voiced stops. Ten-month-olds (N=32) failed to discriminate either contrast. We conclude that 1) English pre-voicing may be a subtle contrast requiring experience to perceive, and 2) this version of the SAPP might not be an ideal methodology to examine discrimination abilities in 10-month-olds. Overall, our findings thus far fit well with the notion that some contrasts require experience to perceive, as well as with past studies reporting mixed results regarding English-learning infants' ability to perceive pre-voicing contrasts (e.g. Aslin et al., 1981; Lasky et al., 1975).

Published by the Acoustical Society of America through the American Institute of Physics

 [DOI: 10.1121/1.4799205]
Received 22 Jan 2013; published 2 Jun 2013

INTRODUCTION

How do infants acquire the phoneme inventory of their native language? This is perhaps one of the most well-studied questions in the field of infant speech perception. Over the course of the past four decades, numerous studies have examined how language experience sculpts the speech perception capabilities of young infants to best suit the language they are learning (e.g. Kuhl, Conboy, Coffey-Corina, Padden, Rivera-Gaxola, & Nelson, 2008). These studies clearly show that by the time a child reaches their first birthday, they have tuned their attention to the contrasts that signal meaningful differences in the native language and no longer attend to the contrasts that do not occur in the native language (e.g. Mattock, Molnar, Polka & Burnham, 2006; Tsao, Liu, & Kuhl, 2006; Werker & Tees, 1984). However, as much as we know about infants' perception of phonological contrasts that either do or not occur in the native language, we have little understanding of infants' perception of segments that occur in the native language, but do not signal a change in word meaning. That is, we know far less about infants' perception of common allophonic variants, such as pre-voiced stop consonants in English. We know that infants sometimes attend to these allophones in a strategic fashion. For example, 2-month-olds use allophonic information to distinguish between a phrase like 'night rates' and a word like 'nitrates' (Hohne & Jusczyk, 1994). However, we also know that infants can sometimes ignore allophonic variation. For example, infants as young as 4 months of age appear to treat oral and nasal allophones of the same vowel as two members of the same category (Seidl, Cristià, Bernard, & Onishi, 2009). But when and how do infants come to understand the role of allophones in the native language? In the current study, we add to the body of literature on infants' perception of allophonic segments by examining English learning infants' perception of pre-voiced stop consonants.

English contains pre-voiced stops, but English listeners consider pre-voiced stops to be allophonic variants of voiced stops. In other words, in contrast to other languages in the world (e.g. Dutch), pre-voicing a stop consonant does not result in the production of a different phoneme (or a change to word meaning). Also, only some English speakers produce pre-voiced stops, and within this group of speakers, there is great variability in how often pre-voiced stops are produced. As a result, the distribution of pre-voicing in a child's language input must be relatively unsystematic (e.g. Lisker & Abramson, 1967). This leads to an interesting question: how do infants' perceive pre-voiced stops? Do they ignore them, like they do segments that do not occur in the native language? Or do they attend to them, like they do contrastive segments that occur in the native language? Moreover, what is the developmental trajectory of infants' attention to allophonic segments such as pre-voiced stops? Work that has investigated whether infants are able to perceive the contrast between pre-voiced and voiced stops are inconclusive, with some studies showing that English-learning infants are unable to perceive the pre-voiced-voiced stop contrast, and some studies showing that 6- to 12-month-olds are able to perceive this contrast (Lasky, Sydral-Lasky, & Klein, 1975; Aslin, Pisoni, Hennessey, & Perey, 1980).

The goal of the current study was to re-visit the question of when and how infants distinguish between pre-voiced and voiced stop consonants. Experiment 1 was conducted with 6-month-olds (infants who have not yet acquired the consonant inventory of their native language) and Experiment 2 was conducted with 10-month-olds (infants who have learned to attend preferentially to consonants that are phonologically distinctive in the native language) In Experiment lA and 2A, 6- and 10- month-old infants were tested on their ability to distinguish between a bilabial stop consonant with negative VOT (a pre-voiced stop) and a bilabial stop consonant with zero VOT (a voiced stop consonant), which is a non-contrastive distinction in English. In Experiment 1B and 2B, a different group of 6- and 10-month-old infants were tested on their ability to distinguish between a bilabial stop consonant with positive VOT and aspiration (a voiceless stop consonant) and a bilabial stop consonant with zero VOT (a voiced stop consonant), which is a contrastive distinction in English.

METHOD

Participants

Thirty-eight 6-month-olds (5 months 15 days to 6 months 28 days; 18 females) participated in Experiment 1. Sixteen of them participated in Experiment 1A and 22 participated in Experiment 1B. Thirty-two 10-month-old infants (10 months, 1 day to 10 months 30 days; 12 females) participated in Experiment 2. Sixteen of these infants participated in Experiment 2A, and 16 participated in Condition 2B. An additional 18 infants were tested but were excluded from analyses due to fussiness (n=10), experimenter error (n=6), not fitting the age range (n=1), and failure to meet language requirements (n=1). All infants were born at 37 weeks gestation, heard English at least 90% of the time, and had at least one parent who learned English in Canada before age 5. Participants were screened to ensure that they had not had an ear infection in the month preceding testing.

Stimuli

All stimuli used in Experiments 1 and 2 were recorded by a young female from Southern Ontario who was naïve to the purpose of the study. This speaker was chosen because she naturally produced pre-voiced bilabial stops without being explicitly instructed to do so. Twelve tokens were recorded: Six tokens of the pre-voiced variant of the syllable /ba/ and six tokens of the syllable /pa/.

In order to minimize acoustic distinctions between stimuli, the two contrast pairs were derived from the same tokens. In Experiment 1A and 2A, voiced stop tokens were created by splicing off the pre-voiced portion of the negative VOT tokens (as in van Alphen & McQueen, 2006). In Experiment 1B and 2B, voiced stop tokens were created by removing the aspiration from the syllable. An example of one of each token, both in its full and manipulated form, is shown in Figures 1 and 2.

The pre-voiced portion of the negative VOT stimuli had an average duration of 165.8 msec (range: 121-212 msec). The aspirated portion of the positive VOT stimuli had an average of 116.83 msec (range: 105-130 msec). Four tokens were used for the habituation phase, and the remaining two were used for the test phase. The stimuli for the habituation phase consisted of tokens of the same stimulus type (e.g. either pre-voiced or voiced stops if in Experiment A), counterbalanced across participants. The Non-Alternating Trials consisted of the same type of stimuli (e.g. pre-voiced stops) as participants heard in habituation. For the Alternating test trials, the same token was repeated, and alternated between its full and manipulated form, with the novel token (a full token if the participant was habituated to a manipulated token, and vise versa) presented first. In order to maintain perceptual similarity between the spacing of syllables in the Non-Alternating and Alternating test trials, the vocalic portions of the syllables were all separated by 750 msec.

Procedure and Appuratus

The infant testing methodology used in the current study combined the modified variant of the SAPP used by Houston et al. (2007) and the Headturn Preference Procedure used by Saffran et al. (1996). Children were seated on their parents lap in the middle of a 3-sided booth. A light was mounted at eye level on the front and side walls, and a speaker was hidden behind the lights mounted on the side walls. The testing booth was located within a sound-attenuating chamber. Parents listened to masking music over closed headphones so they could not hear the experimental stimuli presented to their child. The experimenter, who was located outside of the sound attenuating chamber, watched the child on a muted video monitor and conveyed the child’s looking behavior to the computer via a button box. Customized software controlled the presentation of sound files and the activity of the lights in the booth.

The exposure phase began with the center light blinking. Once the child oriented toward the center blinking light, it would immediately stop blinking and a side light would begin blinking. At the same time, one of the 59-second exposure sound files described in the stimulus section would begin to play. As was the case in Saffran et al. (1996), the infant’s looking behavior controlled when and how long the lights in the booth flashed, but the exposure sound file continued to play until its completion regardless of where the child looked. The test phase began immediately after the exposure phase ended. Each of the 6 trials began with the flashing of the center light. Once the infant oriented to the center light, it would immediately stop blinking and one of the two side lights would begin blinking. Once the infant oriented toward the side light a test sound file would begin to play, and continue for a maximum of 16 repetitions (~20 seconds) or until the infant looked away for more than two seconds. Looking time to the side lights was recorded for each trial, with looks away from the light that lasted less than two seconds subtracted from the total looking time. The test trials were presented in blocks of two trials each. Attention appeared to drop off substantially in blocks 2 and 3, so only the first block of trials are reported here.

During each block of test trials, infants were presented with one Alternating and one Non-alternating Trial (see description in stimulus section above). We predicted that if infants perceived the difference between the two syllable types presented during Alternating Trials, then they should listen longer on Alternating than Non-alternating Trials. The logic behind this prediction was two fold. First, the Alternating Trials contained syllable types that did not occur during the exposure phase, and therefore they should draw infants’ attention if perceived. Second, past studies using variants of the SAPP have shown that infants find Alternating Trials more interesting than Non-Alternating Trials

(Best & Jones, 1998; Houston, Horn, Qi, Ting & Gao, 2007; Maye, Werker & Gerken, 2002; Bentocini, Nazzi, Cabrera, & Lorenzi, 2010).

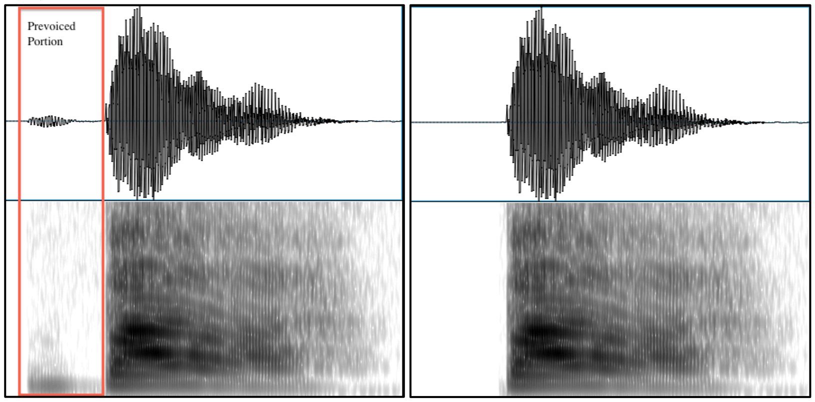

FIGURE 1. Sample stimuli used in Experiment 1A and 2A. As indicated by the red box, the /ba/ on the left begins with a pre-voiced stop (negative VOT). The /ba/ on the right lacks pre-voicing (zero VOT). The latter syllable was created by splicing the pre-voicing off of the syllable on the left. In other words, the voiced and pre-voiced stimuli were acoustically identical in all respects besides the presence or absence of pre-voicing.

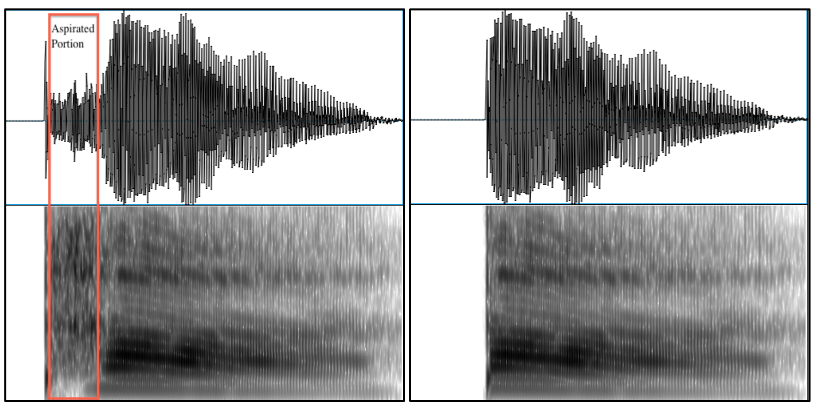

FIGURE 2. Sample stimuli used in Experiment 1B and 2B. As indicated by the red box, the ‘pa’ syllable on the left contains an aspirated portion that makes this token have positive VOT. The ‘ba’ syllable on the right lacks this aspiration, which makes it a zero VOT token. The latter syllable was created by splicing the aspirated portion off of the syllable on the left. In other words, the voiced and pre-voiced stimuli were acoustically identical in all respects besides the presence or absence of aspiration.

RESULTS

As there were no significant differences between trial types in the second and third blocks of either experiment, the results from only the first block of Experiment 1 and 2 are reported here. In Experiment 1A (the pre-voiced-voiced contrast), 6-month-old infants did not show any preference for the Alternating trials over the Non-Alternating trials (M_{alt} = 8.205, *SD* = 5.528; M_{nonalt} = 8.777, *SD* = 5.272; *t(15)* = 0.412, *p* = .686). Thus, we found no evidence that 6-month-old English-learners perceive the difference between pre-voiced and voiced bilabial stops. However, in Experiment 1B (the voiced-voiceless contrast), 6-month-old infants listened significantly longer to the alternating trials over the non-alternating trials (M_{alt} = 13.328, *SD* = 5.918; M_{nonalt} = 10.876, *SD* = 5.503; *t(21)* = -2.114, *p* = .047). This latter finding replicates past studies demonstrating that infants attend to phonologically meaningful contrasts in the native language. Ten-month-old infants did not listen significantly longer to alternating trials over non-alternating trials in Experiment 2A (M_{alt} = 8.814, *SD* = 6.079; M_{nonalt} = 7.449, *SD* = 5.136; *t(15)* = -0.672, *p* = .512) or in Experiment 2B (M_{alt} = 8.099, *SD* = 5.004; M_{nonalt} = 7.994, *SD* = 5.283; *t(15)* = -0.058, *p=.955)*. Results are shown in Figure 3. Thus, we observed no evidence that the 10-month-olds in our study could distinguish between either a pre-voiced stop and voiced stop or a voiced and voiceless stop. This last finding is very surprising as numerous past research has shown that 10-month-olds should be very sensitive to the difference between /pa/ and /ba/.

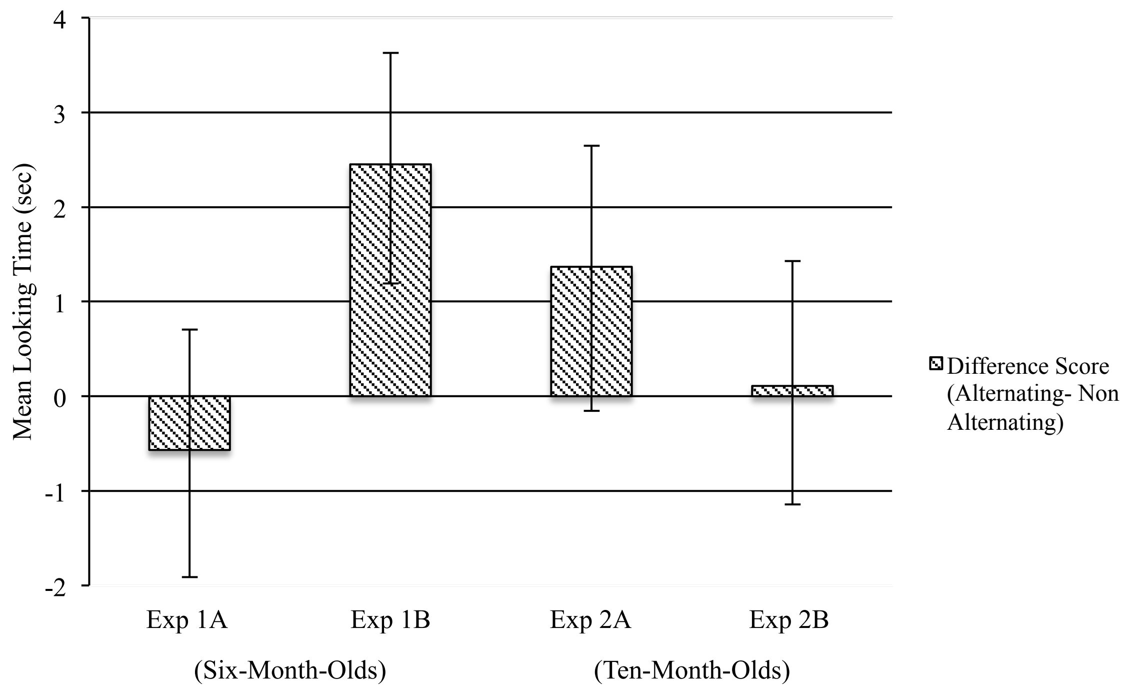

FIGURE 3. Mean looking time the first block in all Experiments. The only place we observed any evidence that infants could discriminate the presented contrast pair was in Experiment 1B, where 6-month-olds were presented with voiced and voiceless stops. Error bars denote SEM.

DISCUSSION

In this study, we examined 6- and 10-month-old infants' ability to perceive the distinction between pre-voiced stop consonants (which have negative VOT) and voiced stop consonants (with zero VOT). Six-month-old infants, who are still in the process of acquiring the consonant categories of their native language, were able to distinguish between voiced and voiceless stops (or positive VOT and zero VOT stops). However, these infants were not able to distinguish between pre-voiced and voiced stops. Ten-month-old infants, in contrast, showed no evidence of being able to distinguish either pre-voiced and voiced stops, or voiced and voiceless stops.

Why did we fail to find any evidence that infants can discriminate between voiced and pre-voiced stops? There are four likely explanations for our findings. The first is that pre-voicing may be acoustically subtle, and that infants are not born with the ability to perceive this contrast. Perhaps the distinction between pre-voiced and voiced stops requires exposure to perceive, like other acoustically subtle contrasts (e.g. Narayan, Werker, & Beddor, 2010). Thus, infants exposed to English, regardless of age, would not be able to perceive this contrast. Then again, this explanation is somewhat inconsistent with past work that suggests that English-learning infants are sensitive to variation in VOT (McMurray & Aslin, 2005). This explanation is also somewhat difficult to reconcile with reports that adult English speakers readily distinguish between zero and negative VOTs (Zlatin, 1974). A second potential explanation for our results is that the artificial nature of our stimuli (e.g. creating a voiced stop by removing the aspiration from a voiceless consonant) may have made it difficult for infants to succeed in our discrimination task. Although our stimulus creation procedures allowed us to minimize the acoustic distance between tokens of different syllable types, it could have also led to the elimination of critical voicing cues in the speech signal. Thus, had we used all naturally produced syllables, we might have observed a different result. Another related possibility is that the pre-voicing produced by the English speaker who recorded our stimuli may not have been as clear as the type of pre-voicing made by speakers of languages that use pre-voicing in a phonologically contrastive manner. Or maybe the stimuli used in the current study were too uniform. Had we used more variable tokens, we may have been more likely to tap into infants' fine-grained speech sound discrimination skills (Rost & McMurray, 2009). Finally, the variant of the SAPP we used in the current study may simply not be an effective method for testing infants' speech perception capabilities. Indeed, the fact that the 10-month-olds we tested failed to discriminate even voiced and voiceless stops suggests that there was a problem with our testing methodology. Perhaps using a habituation criteria rather than a fixed duration exposure period would have improved the sensitivity of our discrimination procedure.

In summary, the results from the current study are somewhat inconclusive. Although we have provided evidence that 6-month-old infants are more sensitive to the distinction between voiced and voiceless stops than the distinction between pre-voiced and voiced stops, we have at the same time failed to find any evidence that either 6- or 10-month-month-olds attend to the contrast between zero and negative VOT stops. In at least some sense, our findings fit well with the past literature reporting mixed results for infants' ability to perceive zero and negative VOT contrasts. A more conclusive answer to the questions we posed in the current study may await further study with a revised testing methodology that takes into account all of the potentially problematic factors mentioned above. A goal for the future will be to eventually understand how infants represent allophonic information in their emerging word form representations.

ACKNOWLEDGMENTS

This research was funded by SSHRC and NSERC grants awarded to EKJ. We thank Caterina Minaudo and Ysabel Domingo for assistance in running participants, and Huiwen Goy for assistance with PRAAT.

REFERENCES

Aslin, R.N. & Pisoni, D.B., Hennessy, B.L., & Perey, A.J. **(1981)**. "Discrimination of voice onset time by human infants- new findings and implications for the effects of early experience," *Child Dev*. **52**, 1135-1145.

Best, C.T., & Jones, C. **(1998)**."Stimulus alternation preference procedure to test infant speech discrimination," Infant Behav. Dev. **21**, 295.

Bertoncini, J., Nazzi, T., Cabrera, L., & Lorenzi, C. **(2011)**. "Six-month-olds discriminate voicing on the basis of temporal envelope cues," J. Acoust. Soc. Am. **129**, 2761-2764.

Hohne E.A. & Jusczyk, P.W. **(1994)**. 2 month old infants sensitivity to allophonic differences. Percept. Psychophys. , **56**, 613-623.

Houston, D.M., Horn, D.L., Qi, R., Ting, R.Y., & Gao, S.J. **(2007)**. "Assessing speech discrimination in individual infants," Infancy **12**, 119-145.

Lasky, R.E, Sydral-Lasky, A., & Klein, R.E. **(1975)**."VOT discrimination by four and six and a half month old infants from Spanish environments," Journal Exp. child Psychol. **20**, 221-225.

Lisker, L., & Abramson, A.S **(1964)**. "A cross-language study of voicing in initial stops: Acoustical measurements," Word **20**, 384-422

Lisker, L., & Abramson, A. S. **(1967)**. "Some effects of voice onset time in English stops," Lang Speech, **10**, 1-28.

Kuhl, P.K., Conboy, B.T., Coffey-Corina, S., Padden D., Rivera-Gaxiola, M., Nelson, T. **(2008)**. "Phonetic learning as a pathway to language: new data and native language magnet theory expanded (NLM-e)," Phil. Trans. R. Soc. B **363**, 979-1000.

Mattock, K., Molnar, M., Polka, L., & Burnham, D. **(2008)**. "The developmental course of lexical tone perception in the first year of life," Cognition, **106**, 1367-1381.

Maye, J, Werker, J.F., & Gerken, L. **(2002)**. "Infant sensitivity to distributional information can affect phonetic discrimination," Cognition **82**, B101-B111.

McMurray, B. & Aslin R. **(2005)** "Infants are sensitive to within-category variation in speech perception," Cognition **95**, B15-B26.

Narayan, C.R., Werker, J.F., & Beddor, P.S. **(2010)**. "The interaction between acoustic salience and language experience in developmental speech perception: evidence from nasal place discrimination," Dev. Sci. **13**, 407-220.

Rost, G., and McMurray, B. **(2009)**. "Speaker variability augments phonological processing in early word learning," Dev. Sci. 12, 339-349.

Saffran, J.R., Aslin, R.N., & Newport, E.L. **(1996)**. "Statistical learning by 8-month-olds," Science, **274**, 1926-1928.

Seidl, A., Cristià, A., Bernard, A., Onishi, K.H. **(2009)**. "Allophonic and phonemic contrasts in infants' learning of sound patterns," Lang. Learn. Dev. **5**, 191-202

Tsao, F., Liu, H., & Kuhl, P.K. **(2006)**. "Perception of native and non-native affricate-fricative contrasts: Cross-language tests on adults and infants," J. Acoust. Soc. Am. **120**, 2285-2294.

van Alphen, P.M., & McQueen, J.M. **(2006)**."The effect of voice onset time differences on lexical access in Dutch," J. Exp. Psychol. Human, **32**, 178-196.

Werker, J.F., & Tees R.C. **(1984)**. Cross language speech perception-evidence for perceptual reorganization during the first year of life. Infant Behav. Dev. **7**, 49-63.

Zlatin, M.A. **(1974)**. Voicing contrast: perceptual and productive voice onset time characteristics of adults. J. Acoust. Soc. Am. **56**, 981-994.