Subjective Awareness during Cross-Language Speech Perception Attending Unattended Regions of an Acoustic Continuum

Jordan Schoenherr (psychophysics.lab@gmail.com)
John Logan (john_logan@carleton.ca)
Department of Psychology, Carleton University
1125 Colonel By Drive, Ottawa, ON K1S5B6 Canada

Abstract

Linguistic experience attenuates adult listeners' attention to acoustic differences that are not phonemic in the listener's language. In the present study we found that acoustic information is available to listeners after acoustic cues have been processed to identify phonemic categories. Moreover, we also found that listeners maintained an awareness of these differences by comparing the identification function to typicality ratings and confidence reports.

Keywords: speech perception, category boundaries, confidence processing

Introduction

When adult listeners are presented a continuum of speech stimuli varying along an acoustic dimension, they divide the continuum into distinct categories defined by sharp boundaries (Liberman, Harris, Hoffman, & Griffith, 1957), a phenomenon known as categorical perception (Liberman, Cooper, Shankweiler, & Studdert-Kennedy, 1967). Categorical perception was originally taken as evidence that adult listeners are no longer capable of detecting stimulus differences in within-category regions of a speech continuum (Eimas, 1975). A developmental account of categorical perception assumes that when infants acquire phonemes they learn to segment acoustic information into discrete categories. Consistent with this view, research indicates that by the end of their first year, infants develop a reduced sensitivity to differences between stimuli within a given phoneme category (Werker, 1989), with category boundaries becoming more distinct as they become older (e.g., Hazan & Barrett, 2000). Rather than processing all stimulus dimensions, infants learn to attend to specific acoustic cues that determine category membership (Jusczyk, 2000). The question that the present study attempts to answer is whether dults are subjectively aware of attended and unattended acoustic properties.

Extending the concept of categorical perception to the perception of non-native phonemes implies that adult listeners might experience difficulties in perceiving these speech sounds due to a desensitization to acoustic information. In effect, non-native speech sounds are thought to be perceptually assimilated into existing native phonemic categorical representations (cf. Best, 1995, Flege, 1995). Although previous research indicates that listeners can learn to perceive non-native phoneme categories (e.g., Pisoni, Aslin, Perey, & Hennessy, 1982), limited work has been done to examine the metacognitive awareness associated with learning a non-native phoneme category. In the present study, we examined whether feedback could be used to allocate attention to a previously unattended region of an acoustic continuum that corresponded to a non-native phoneme category and whether attention was accompanied by awareness.

Categorical Perception of Speech Sounds

A variety of acoustic cues are used by listeners to define phoneme categories. Voice-onset time (VOT), the interval between aspiration and the vibration of the vocal cords is one such cue. Lisker and Abramson (1967) presented English and Thai listeners with synthetic speech sounds ranging from -150 VOT (prevoiced) to +150 VOT (voiced). The Thai listeners' identification performance resulted in two category boundaries corresponding to /p/, /b/, and /p^h/ whereas English listeners' identification performance yielded only one category boundary corresponding to the phonemes /b/ and /p/. These findings illustrate how non-native speech sounds are assimilated into existing phonemic categorical structure.

Learning to perceive non-native speech sounds

When categorizing stimuli, participants must attend to stimulus properties along the physical continuum that defines the stimuli (Nosofsky, 1986). With training, participants become sensitized to specific regions along the continuum that can affect performance in other tasks (e.g., discrimination tasks; Goldstone, 1994). Thus, once a phonemic category has been acquired listeners will limit their attention to only those psychophysical characteristics where attention has been directed (Jusczyk, 1992).

In order to promote attention to previously unattended acoustic characteristics, listeners typically require some form of training. For example, in order to examine the effectiveness of feedback, Pisoni et al. (1982) provided English listeners with three exemplars of speech sounds from the three regions of the VOT continuum corresponding to voiceless unaspirated, voiced aspirated, and voiceless aspirated stops (i.e, /p/, /b/, and /p^h/, respectively). Following a short period of laboratory training, listeners were capable of identifying and discriminating speech sounds from the non-native /p^h/ category.

Pisoni et al.'s (1982) results have several implications. First, acoustic information must be accessible to listeners in order to classify stimuli into a non-native phoneme category. Further support for the availability of acoustic information also comes from studies using typicality ratings wherein gradedness is exhibited in response functions (e.g., Miller & Volaitis, 1989). Second, when compared to previous studies (e.g., Mackain et al., 1981) the amount of exposure to non-native phonemes is less important than the method used to present the non-native phonemes. Finally, methods that allocate selective attention to previously unattended regions can facilitate acquisition of these phonemic categories. Although attention is drawn to these stimulus features, it remains unclear whether participants have a subjective awareness of these acoustic properties.

Confidence reports and subjective awareness

One method frequently used to quantify subjective awareness in a perceptual discrimination task is to have participants provide a subjective probability that they have provided a correct answer. The use of numeric response scales is particularly useful given that a direct comparison can be made with the proportion of correct responses obtained from the primary task. In this case, 50% represents a response associated with a guess whereas 100% represents absolute certainty in a response. Thus, listeners are well calibrated when they assign a level of confidence (e.g., 60%) that corresponds to their obtained accuracy (e.g., p(cor) = 0.6). Listeners frequently demonstrate failures in assigned appropriate subjective probabilities, evidencing miscalibration.

Rather then miscalibration being random, systematic deviations have been observed which some argue represent a differential ability to assess the performance of specific cognitive operations (Dawes, 1980), with still others arguing that these measures represent contributions of implicit and explicit knowledge (e.g., Dienes & Berry, 1997; cf. Lichtenstein & Fischhoff, 1977). For instance, underconfidence has been observed in perceptual tasks (e.g., Bjorkman, Juslin, & Winman, 1993) whereas overconfidence is typically observed in tasks requiring assess to general knowledge (e.g., Gigerenzer, Hoffrage, & Kleinbolting, 1991). Dawes (1980) has argued that these findings are the result of participants' uncertainty in their perceptual processes and certainty about information stored in long-term memory. This suggests that confidence processing uses a second set of operations to assess the content and output of the primary decision process (e.g., stimulus classification).

An alternative class of confidence models assumes an additional set of operations wherein primary decision information is rescaled (e.g., for a review see, Baranski & Petrusic, 1998). These accounts are supported by the observation that the requirement of confidence reports increase primary decision response time (Baranski & Petrusic, 1998) and the dissociable effects of nondiagnostic information on the primary decision response selection and confidence reports (e.g., Schoenherr, Leth-Steensen, & Petrusic, 2010). In the present study, we assume that both acoustic and phonemic information will affect confidence reports and that the correspondence of identification accuracy and mean confidence will demonstrate whether participants have an awareness of acoustic information when identifying phonemes.

Present Study

In the present study, English listeners were asked to classify stimuli from a VOT continuum into phoneme categories corresponding to either two categories, /p/ and /b/ (both found in English), or three categories, /p/, b/, plus the prevoiced category /p^h/ (found in Thai). If listeners are aware of acoustic differences, their subjective confidence should differ across regions of the VOT continuum as evidenced by miscalibration. If underconfidence is evidenced, it suggests that listeners did not have subjective awareness of a well-defined phonemic category whereas if overconfidence is evidenced, it suggests that listeners believed that they had a better understanding of the phonemic category then they in fact did.

Experiment 1

Previously, Schoenherr, Logan, and Winchester (2012) observed slight overconfidence for responses made for /b/ stimuli located at the category boundary on the VOT continuum. In Experiment 1, we sought to see whether we could increase uncertainty in this region by presenting stimuli from the prevoiced region. The addition of stimuli with these novel acoustic properties should increase uncertainty for stimuli from the /p^h-b/ portion of the continuum if acoustic properties are being used to identify stimuli whereas the participants should remain certain in their judgments of stimuli associated with the /p/ category.

Method

Participants

Nine Carleton University students participated in the study for course credit; all were native speakers of English or had extensive experience with English and reported normal hearing and no speech pathologies.

Materials

Fifteen synthetic speech stimuli were used, obtained from the Haskins Laboratories website (HL, 2011; Lisker & Abramson, 1967). These stimuli varied along the VOT continuum from -70 to 70 ms VOT. As per the method used by Pisoni et al. (1982), listeners were presented with stimuli which corresponded to the voiceless unaspirated, voiced aspirated, and voiceless aspirated stops. The latter categories are present in English while the former is not. The sounds were originally recorded on reel-to-reel tape and later converted into AIFF format at Haskins Laboratories. Stimuli were pre-processed using a DC offset correction to eliminate clicks present in the AIFF versions and then converted into WAV files.

Figure 1a. ID Function without the Requirement of Confidence

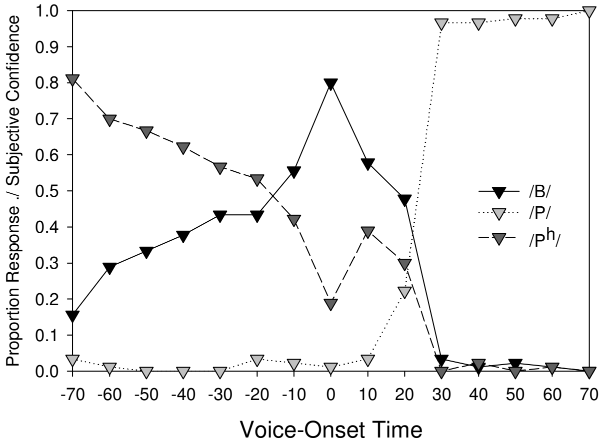
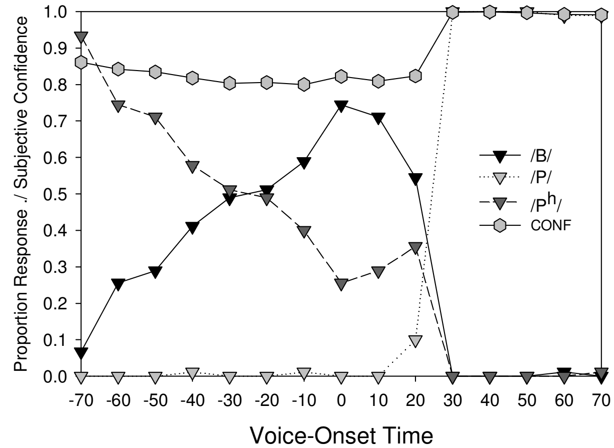


Figure 1b. Identification Function with the Requirement of Confidence



Procedure

Modelled after the procedure used by Pisoni et al. (1982), listeners were presented with a brief training block in which they heard three stimuli prior to the identification tasks, one from each region of the VOT continuum (-70, 0, and 70 ms VOT, corresponding to the /p^h/, /b/, and /p/ categories). Ten replications of these stimuli were presented in the order /p^h/, /b/, and /p/. Following this initial training, listeners also were trained using an identification (ID) task with feedback. They were presented with a stimulus and then reported whether it was a /p^h/, /b/, or /p/ using the 'V', 'B', or 'N' keys, labeled as '_B', 'B', and 'P', respectively. After they had indicated their response, 'Correct' or 'Incorrect' was presented visually on the screen. Listeners completed a total of 80 trials in the training task.

Following training, listeners again identified the stimulus presented as a /p^h/, /b/, or /p/ using the keyboard but no feedback was provided. In the first block, after they completed each ID trial they also indicated their level of confidence in their response using the 'E' through 'I' keys, on a 6-point scale with 50% representing a guess and 100% representing certainty. In the second block, confidence was not reported. Each block was composed of a total of 150 trials.

The duration of the experiment was approximately 30 minutes. Replicating the procedure of Pisoni et al. (1982), stimuli were presented randomly except for the sequential presentation of the training stimuli. Listeners were presented with the stimuli over headphones at a comfortable listening level using PsychoPy software (Peirce, 2007).

Results

Proportion Identification. Unlike studies that have examined two-category identification performance using confidence reports (e.g., Schoenherr et al., 2012), only the /p/ phoneme category showed a sharp identification function (Figures 1a and 1b). In general, however, listeners could consistently identify stimuli associated with the /p^h/ and /b/ category with greater than chance accuracy (i.e., stimuli with VOTs of -70, -60, -50, 0, and 10), indicating that even with a brief period of training, listeners can begin to acquire a non-native speech category. Supporting this, we obtained a significant effect for VOT stimulus, $F(14,112) = 7.389$, $MSE = .435$, $p = .001$, $\eta^2 = .480$. Given that we did not obtain a main effect or interaction of confidence reports, it suggests that confidence reports did not significantly affect ID performance thereby permitting a straightforward interpretation of the remaining results.

(In an alternate version of Experiment 1 the order of confidence and no confidence blocks was reversed [see Schoenherr and Logan, in preparation]. Under these conditions, participants' (n = 12) identification performance for the non-native phoneme category was reduced in the first block when confidence was required. Participants generally paired neighbouring stimuli together in alternating clusters of 2-4 speech sounds [e.g., VOTs 10 and 0 ms were identified as /b/, VOTs -10 and -20 ms identified as /p^h/, and VOTs -30 to -60 ms identified as /b/]. Such a finding suggests that the requirement of producing confidence reports can interfere with the primary task.)

Table 1. Mean identification RTs (ms) along the critical regions of the VOT continuum with standard error reported in parentheses.

Condition	/p^h/	/p^h-b/	/b/	/b-p/	/p/
No Conf.	888	910	893	933	796
	(45)	(51)	(54)	(75)	(43)
Conf.	1,009	1,137	1,072	992	855
	(47)	(124)	(113)	(83)	(41)

Identification Response Time. Prior to conducting an analysis of the response time (RT) data, we collapsed stimuli into regions five regions along the VOT continuum corresponding to two category boundaries (CBs) /p^h-b/ and /b-p/ corresponding to CB_1 (-30, -20) and CB_2 (20, 30), respectively, and equivalent within-category pairs corresponding to /p^h/ (-70, -60), /b/ (0,10), and /p/ (60, 70), respectively. Using the criterion of 3 standard deviations, 4.3% of the responses were identified as outliers and removed from the final analysis.

An analysis of the remaining responses times revealed a main effect of VOT region, $F(4,32) = 4.45$, $MSE = .041$, $p = .025$, $\eta^2 = .357$. Table 1 indicates that response latencies were longer at category boundaries as well as for the non-native (/p^h/) and modified native (/b/) categories relative to the native /p/ category. A main effect of the requirement of confidence report was also obtained, $F(1,18) = 14.55$, MSE

= .026, p = .005, η^2 = .645. Again, Table 1 demonstrates longer latencies with the requirement of confidence relative to the no confidence condition. Given that the confidence block always followed the no confidence block, this finding cannot be attributed to automaticity. The interaction of confidence condition and VOT region was only marginally significant, $F(4,32) = 2.724$, $MSE = .019$, $p = .099$, η^2 = .254.

Confidence Reports. Figure 1a and 1b also demonstrate the effect of confidence measures. Listeners expressed less confidence in their responses to stimuli located within the /p^h/ and /b/ categories. As was the case with ID accuracy, we observed a main effect of the stimulus location along the VOT continuum on mean confidence, $F(14,112) = 6.931$, $MSE = 1011.371$, $p = .018$, $\eta^2 = .464$. Our comparison of over/underconfidence bias did not reveal any significant effects, $F(14,112) = 2.146$, $MSE = .0354$, $p = .133$, η^2 = .212. Unlike previous studies that have observed graded responses indicative of perception of acoustic properties (e.g., McMurray, Tanenhaus, & Aslin, 2002), our findings suggest that listeners are not fully aware of the processes allowing them to identify stimuli.

Discussion

An instructive comparison can also be made between the results of the present study and conditions in which participants identify two native phonemes. Schoenherr et al. (2012) observed a small decrease in confidence around the /b-p/ category boundary when only the voiced and voiceless portions of the continuum were presented to listeners. When the prevoiced portion of the continuum was additionally presented, Schoenherr et al. (2012) observed lower confidence in the /p^h-b/ portion of the continuum. When compared to identification accuracy, this pattern of responses leads to underconfidence in comparison to the overconfidence observed in the present study. Taken together with our results, this suggests that training does result in the allocation of attention to newly relevant acoustic properties of the stimuli in this region of the VOT continuum, thereby reducing certainty. Although listeners are somewhat more conservative in their confidence reports, they are still overconfident suggesting that the phonemic representations that they are subjectively aware of are less accurate than the acoustic information necessary to identify the stimuli.

Experiment 2

The pattern of overconfidence observed in Experiment 1 suggests two features of a listener's awareness of acoustic properties. First, the observation that mean confidence exceeded identification accuracy in the /p^h-b/ region of the VOT continuum indicates that listeners did not have complete access to acoustic properties on a trial-to-trial basis. Second, mean confidence in identification responses to stimuli within the /p^h-b/ region was lower relative to stimuli from the /p/ region. These findings suggest that while participants confidence reports are influenced by their two native phoneme categories, they might have some awareness of the acoustic properties of stimuli within the /p^h-b/ region which cause uncertainty in their responses. To assess the extent to which acoustic and phonemic properties are available to subjective awareness, Experiment 2 required participants to additionally provide typicality ratings that have previously been shown to reflect acoustic properties (Miller & Volaitis, 1989).

Method

Participants, Materials, and Procedure

Fifteen Carleton University students participated in the study. Other participant and stimulus characteristics were the same as in Experiment 1.

We replicated the methods of identification block in Experiment 1 and required post-decisional confidence reports. In a subsequent another block, listeners rated the typicality of each stimuli on a scale of 1 through 9 where '9' represented highly typical of a category and '1' represented a highly atypical member of a category. Listeners were presented with the same number of trials as the identification block.

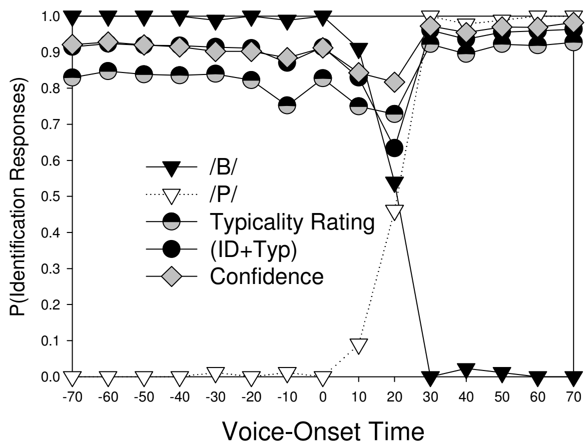
Results

Proportion Identification. Replicating the general results of Experiment 1, the location of the stimuli along the VOT continuum significantly affected identification performance, $F(14,112) = 9.149$, $MSE = .124$, $p = .005$, η^2 = .533. Figure 1b demonstrates, participants had a sharp category boundary between stimuli for the /b/ and /p/ categories. A noticeable difference was evident in the location of the boundary. Whereas in Experiment 1 the boundary was located between VOT 20 ms and 30 ms, a shift such that the boundary was now located at VOT 20 ms with a resulting decrement in performance for VOT 10 ms stimuli. We assume that these results are due to range effects. In general, these findings permit a straightforward interpretation of the remaining results.

Confidence Reports. Figure 2 also demonstrates the effect of confidence measures. Like ID accuracy, we found that subjective confidence varied along the VOT continuum, $F(1,14) = 6.55$, $MSE = 44.11$, $p = .008$, $\eta^2 = .319$. Relative to Experiment 1, we did observe greater underconfidence in the negative portion of the VOT continuum.

Typicality Task. Typicality ratings also varied significantly as a function of stimulus location along the VOT continuum, $F(14,112) = 5.820$, $MSE = .3.295$, $p = .009$, η^2 = .421. Unlike accuracy, but like mean confidence, typicality ratings appeared to be more responsive to the acoustic properties of the stimuli. Participants considered stimuli in the /b/ and /p^h/ range as less typical then stimuli in the /p/ range even though they exhibited equal accuracy. Moreover, within-category ratings exhibited more graded responses.

Figure 2. ID, Subjective Ratings, and Interpolated Function



Interpolated Function. The similarities in patterns observed in confidence and typicality suggested a potential relationship between these two functions. As Figure 2 indicates, mean confidence ratings are situated between accuracy in the identification task and typicality ratings in the typicality task. Pearson's correlations revealed the strongest relationship between confidence and typicality ratings, $r^2 = .960$, $p < .001$. The correlation between identification responses and mean confidence was also significant, $r^2 = .446$, $p = .007$, although the correlation between identification and typicality was only marginally significant, $r^2 = .261$, $p = .051$. These findings suggest that confidence is associated with both identification accuracy and typicality ratings but that identification accuracy and typicality ratings are only weakly related.

In order to examine the relationship between accuracy, typicality, and confidence ratings we converted typicality to a proportion, summed it with proportion correct, and produced an interpolated function. A paired-samples t-test revealed that the mean confidence function and the interpolated function did not significantly differ from one another, $t(14) = .309$, $p = .762$. This suggests that confidence reports were closely associated with information from both identification accuracy (associated with phonemic representations) and typicality ratings (associated with acoustic information). All other paired-sample t-tests were significant (all ts > 3.283, ps < .005) indicating that different sources of information contributed to response selection for each dependent measure.

Discussion

The typicality results provide strong support for the accessibility of acoustic information during speech perception. When normalized along a common scale, typicality ratings were lower in the /p^h-b/ region of the VOT continuum relative to identification responses, suggesting that participants had an awareness of acoustic properties. This finding replicates previous studies (e.g., Miller & Volatis, 1989). A similar pattern was again observed in confidence ratings. More importantly, when typicality and identification functions were interpolated, a near perfect fit was obtained with the confidence function. This finding is of particular interest as it suggests that confidence is affected by both phonemic representations stored in long-term memory and acoustic properties available in short-term memory.

Conclusion

The findings of the present study have several implications. First, we replicated results from early studies (e.g., Pisoni et al., 1982) showing that participants can learn a non-native speech category using only three exemplars selected from the VOT continuum. Second, our results indicate that the acquisition of non-native phonemes likely resulted from the allocation of attention to regions along the VOT continuum that were previously unattended (e.g., Jusczyk, 1992). In comparison to tasks where participants must identify two categories, such as in Experiment 2 and Schoenherr et al. (2012), participants exhibited well-defined /b/ and /p/ phonemic categories. The induction of a novel category boundary requires attention to acoustic properties. In the present study, support for the detection of acoustic properties of stimuli from the same phonemic category comes from both confidence reports and typicality ratings. Both subjective confidence and typicality decreased in the /p^h-b/ region of the VOT continuum relative to ratings within the /b-p/ region.

As briefly noted above in reference to Experiment 1, this pattern appears to hold provided that participants do not need to concurrently monitor their performance (i.e., provide confidence reports). Thus, although feedback directs attention toward previously unattended regions of the VOT continuum, additional top-down monitoring might emphasize differences between exemplars to the detriment of the formation of novel phonemic categories. Similar findings of top-down interference have been observed in visual search tasks wherein activation of object-level representations in long-term memory creates interference in detection of stimuli based on visual features (e.g., Zhaoping & Firth, 2011).

Finally, our study also has implications for models of confidence. The identification function and typicality ratings are believed to reflect the detection of phonemic and acoustic information, respectively. When interpolating between these two functions, we obtained a function that approximated that obtained from confidence reports. This suggests that additional processing is required to create a confidence report and that such a process can integrate information from short- and long-term memory.

References

Abramson, A., & Lisker, L. (1965). Voice onset time in stop consonants: Acoustic analysis and synthesis. Proceedings *of the Fifth International Congress on Acoustics*, Liege, A51.

Baranski, J. V., & Petrusic, W. M. (1994). The calibration and resolution of confidence in perceptual judgements. *Perception & Psychophysics*, *55*, 412-428.

Baranski, J. V., & Petrusic, W. M. (1998). Probing the locus of confidence judgments: Experiments on the time to determine confidence. *Journal of Experimental Psychology: Human Perception and Performance*, *24*, 929-945.

Best, C. T. (1995). A direct realist view of cross-language speech perception: New Directions in Research and Theory. In W. Strange (Ed.) *Speech perception and linguistic experience: Theoretical and methodological issues* (pp. 171–204). Baltimore: York Press.

Dawes, R. M (1980). Confidence in intellectual vs. confidence in perceptual judgments. In E. D. Lantermann & H. Feger (Eds.) *Similarity and Choice: Papers in Honor of Clyde Coombs* (pp. 327-345). Bern Hans Huber.

Eimas, P. D. (1975). Auditory and phonetic coding of the cues for speech: Discrimination of the [r-1] distinction by young infants. *Perception & Psychophysics, 18,* 341-347.

Flege, J., (1995). Second language speech learning: Theory, findings and problems. In W. Strange (Ed.). *Speech Perception and Linguistic Experience: Theoretical and Methodological Issues* (pp. 233–277). Baltimore: York Press.

Gigerenzer, G., Hoffrage, U., & Kleinbölting, H. (1991). Probabilistic mental models: A Brunswikian theory of confidence. *Psychological Review*, *98*, 506-528.

Goldstone, R. (1994). Influences of categorization on perceptual discrimination. *Journal of Experimental Psychology: General*, 123, 178-200.

Haskins Laboratories (2011). Abramson/Lisker VOT Stimuli. Retrieved 01/12/2011. From http://www.haskins.yale.edu/featured/demo-liskabram/index.html /.

Hazan, V. & Barrett, S. (2000). The development of phonemic categorization in children aged 6-12. *Journal of Phonetics,* 28, 377-396.

Jusczyk, P. (1989). Developing phonological categories from the speech signal. In C. A. Ferguson, L. Menn, & C. Stoel-Gammon (Eds.), *Phonological Development: Models, Research, Implications*. Monkton, MD: York.

Jusczyk, P.W. (2000). *The Discovery of Spoken Language*. Cambridge: MIT Press.

Kvidera, S., & Koustaal, W. (2008). Confidence and decision type under matched stimulus conditions: overconfidence in perceptual but not conceptual *Decisions. Journal of Behavioral Decision Making*, *21*, 253–281.

Liberman, A.M., Cooper, F.S., Shankweiler, D.P., & Studdert-Kennedy, M. (1967). Perception of the speech code. *Psychological Review*, *74*, 431-461.

Liberman, A. M., Harris, K. S., Hoffman, H. S., & Griffith, B. C. (1957). The discrimination of speech sounds within and across phoneme boundaries. *Journal of Experimental Psychology*, *54*, 358-368.

Lichtenstein, S., & Fischhoff, B. (1977). Do those who know more also know more about how, much they know? *Organizational Behavior and Human Performance*, *20*, 159-183.

Lisker, L, & Abramson, A. S. (1967). The voicing dimension: Some experiments in comparative phonetics. *Proceedings of the 6th International Congress of Phonetic Sciences.* Prague: Academia.

Mackain, K. S., Best, C. T., & Strange, W. (1981). Categorical perception of English /r/ and /l/ by Japanese bilinguals. *Applied Psycholinguistics*, *2*, 369-390.

McMurray, B., Tanenhaus, M., & Aslin, R. (2002). Gradient effects of within-category phonetic variation on lexical access. *Cognition*, *86*, B33-B42.

Miller, J. L., & Volaitis, L. E. (1989). Effect of speaking rate on the perceptual structure of a phonetic category. *Perception & Psychophysics*, *46*, 505-512.

Nosofsky. R. (1986). Attention, similarity, and the identification-categorization relationship. *Journal of Experimental Psychology: General*, *115*, 39-57.

Pisoni, D. B., Aslin, R. N., Percy, A. J., & Hennessy, B. L. (1982). Some effects of laboratory training on identification and discrimination of voicing contrasts in stop consonants. *Journal of Experimental Psychology: Human Perception and Performance*, *8*, 297-314.

Pisoni, D. B., & Tash, J. B. (1974) Reaction times to comparisons within and across phonetic categories. *Perception & Psychophysics, 15,* 285-290.

Schoenherr, J. R., Leth-Steensen, C., & Petrusic, W. M. (2010). Selective attention and subjective confidence calibration. *Attention, Perception & Psychophysics*, *72*, 353-368.

Schoenherr, J. R. & Logan, J. (in preparation). *Attending Unattended Regions of an Acoustic Continuum: Evidence for Acoustic and Phonemic Representations in Speech Perception.*

Schoenherr, J. R., Logan, J. & Winchester, A. (2012). Subjective confidence of acoustic and phonemic representations during speech perception. *Proceedings of the 34th Annual Meeting of the Cognitive Science Society*, Sapporo, Japan.

Vickers, D., & Packer, J. S. (1982). Effects of alternating set for speed or accuracy on response time, accuracy, and confidence in a unidimensional discrimination task. *Acta Psychologica*, 50, 179-197.

Zhaoping, L. & Firth, U. (2011). A clash of bottom-up and top-down processing visual search: The reversed letter effect revisited. *Journal of Experimental Psychology: Human Perception & Performance*, 37, 997-1006.