INTERSPEECH 2017
August 20–24, 2017, Stockholm, Sweden

MMN responses in adults after exposure to bimodal and unimodal frequency distributions of rotated speech

Ellen Marklund, Elísabet Eir Cortes and Johan Sjons

Stockholm Babylab, Phonetics Laboratory, Dept. of Linguistics, Stockholm University, Sweden
ellen|elisabet|johan@ling.su.se

Abstract

The aim of the present study is to further the understanding of the relationship between perceptual categorization and exposure to different frequency distributions of sounds. Previous studies have shown that speech sound discrimination proficiency is influenced by exposure to different distributions of speech sound continua varying along one or several acoustic dimensions, both in adults and in infants. In the current study, adults were presented with either a bimodal or a unimodal frequency distribution of spectrally rotated sounds along a continuum (a vowel continuum before rotation). Categorization of the sounds, quantified as amplitude of the event-related potential (ERP) component mismatch negativity (MMN) in response to two of the sounds, was measured before and after exposure. It was expected that the bimodal group would have a larger MMN amplitude after exposure whereas the unimodal group would have a smaller MMN amplitude after exposure. Contrary to expectations, the MMN amplitude was smaller overall after exposure, and no difference was found between groups. This suggests that either the previously reported sensitivity to frequency distributions of speech sounds is not present for non-speech sounds, or the MMN amplitude is not a sensitive enough measure of categorization to detect an influence from passive exposure, or both.
Index Terms: perceptual categorization, MMN, speech sounds, rotated speech

1. Introduction

The ability to discriminate different speech sounds undergoes a change during the first year of life [1, 2]. In short, infants go from being able to discriminate most of the speech sound contrasts used in any language to mainly being able to discriminate native contrasts. This perceptual reorganization is typically considered to be related to language exposure, as discrimination of native speech sounds is enhanced even as discrimination of non-native contrasts is attenuated [3]. Furthermore, ten- to twelve-month-old infants exposed to interpersonal interaction in a non-native language, maintain the ability to discriminate a contrast relevant to the non-native language past the age when it would otherwise be attenuated [4].

Infants are sensitive to multiple types of statistical regularities in the speech input, as for instance probabilities of sequential co-occurrence of sounds and probability distributions of sounds (for a review, see [5]). A statistical learning mechanism that likely is highly relevant to the development of speech sound categorization is sensitivity to frequency distributions of sounds. Six- and eight-month-old infants discriminate between endpoints of a syllable continuum (/da/-/ta/) after they have been exposed to the syllable continuum in a bimodal distribution, but not after they have been exposed to the continuum in a unimodal distribution [6]. In a follow-up experiment with eight-month-olds, a contrast was used that had previously been proven harder for infants to discriminate (prevoiced versus short-lag stop consonants [7]). Infants in the bimodal group discriminated between sounds after exposure, whereas the control group and the unimodal group did not [8]. Similar results were found for ten-month-old infants, although in this case more exposure was needed [9]. Results along the same lines have been found for sibilants [10]. Taken together, these findings suggest that the distributional properties of the speech input are important for the development of speech sound categorization, in terms of disregarding irrelevant acoustic variation as well as in terms of highlighting less prominent acoustic differences.

Contrastiveness between speech sounds in a given language is typically reflected in the frequency distributions of sounds along one or more acoustic dimensions [11]. In American English for example, productions of /l/ and /r/ have a bimodal frequency distribution along onset frequency of the third formant, whereas in Japanese, the flap closest corresponding to English /l/ and /r/ has a unimodal distribution along the same dimension [12]. Similarly, bimodal and trimodal frequency distributions have been found for initial stop consonants in languages with two and three phonemic categories respectively [13]. This suggests that the information needed to form perceptual categories based on frequency distributions is available to the language learning infant.

It remains unclear however, whether detection of category information from passive exposure to distributional properties in the input is restricted to speech or if it is a domain-general mechanism, such as for example sensitivity to transitional probabilities [14, 15, 16]. Although category formation of non-speech sounds has been demonstrated in adults from distributional properties of the input, it has for the most part been the results of either explicit [17], or implicit training [18, 19]. In one case however, passive exposure to non-speech sounds prior to categorization training had a positive impact on performance in the post-training categorization test [17], Experiment 4). Based on these previous studies, there is reason to believe that the sensitivity extends beyond speech, and that short passive exposure to non-speech sounds along different frequency distributions may well lead to perceptual categorization of those sounds.

1.1. Present study

In the present study, non-speech sounds along a continuum was presented to adults, and categorization was measured before and after exposure using the event-related potential (ERP) component mismatch negativity (MMN). The non-speech stimuli consisted of a spectrally rotated vowel continuum. Rotated speech [20] sounds more like croaking or buzzing noise than human speech, and was used for two reasons: first, because it has an acoustic complexity comparable to speech [21, 22]; second, because participants are highly unlikely to have been exposed to it, which allows controlling for the amount of exposure. Adults'

http://dx.doi.org/10.21437/Interspeech.2017-1110

speech sound perception is influenced by exposure to bimodal and unimodal speech sound distributions, much like infants' speech sound perception is [23, 24, 25, 26, 27, 28, 29]. Since the mechanism under investigation has been found in both adults and infants, adult subjects were used in the present study for practical reasons. The MMN amplitude was used to quantify categorization. The acoustic difference between stimuli sounds in MMN experiments is reflected in the MMN amplitude, with greater difference resulting in larger amplitude [30]. However, when the stimuli is speech, the phonemic categories of the stimuli also influences the MMN amplitude. If the two sounds belong to different categories, the MMN has a larger amplitude than if they belong to the same category, even if the acoustic difference is controlled for [31, 32]. Since the MMN amplitude thus reflects perceptually sorting two sounds into different categories, that is, categorization, it should differ before and after exposure if perceptual categories were formed.

In the current study, adults are expected to show categorization in response to different frequency distributions of rotated sounds. After exposure to a bimodal distribution the participants are expected to have a larger MMN amplitude, whereas after exposure to a unimodal distribution they are expected to have a smaller or unchanged MMN amplitude.

2. Method

2.1. Participants

Twenty adults participated in the study (mean age = 34 years, age span = 23-60 years). Since stimuli were non-speech sounds, a specific native language was not a criterion for participation, so participants' first language(s) varied. Seventeen participants were right-handed, two were left-handed, and one had ambiguous handedness. All subjects received two movie vouchers as thanks for their participation. The study has been approved by the Ethical Review Board at Karolinska Institutet (2015/63-31).

2.2. Stimuli

Stimuli were rotated sounds created from the Swedish vowels /e/ and /i/. Two different CV-syllables, /te/ and /ti/, produced by a female native speaker of Swedish, were recorded in an echo-free chamber at the Phonetic Laboratory, Department of Linguistics, Stockholm University. The consonant in each syllable was removed using Wavesurfer 1.8.5 [33]. A continuum of six intermediate vowels between the /e/ and /i/ was created in Praat 6.0.21 [34], utilizing a freely available script [35]. The continuum was spectrally rotated using Mathematica (Wolfram Research Inc., Champaign, Illinois, USA), a procedure previously described in [22]. The resulting sounds will be referred to as R1 to R8, with R1 being the rotated /e/ and R8 the rotated /i/. In the MMN blocks of the experiment, R3 was the standard stimulus, whereas R6 was the deviant stimulus.

2.3. Experiment design

The experiment consisted of three blocks: a pre-exposure MMN block, an exposure block, and a post-exposure MMN block. The MMN blocks consisted of 1000 stimuli each, a total of 2000 stimuli, 20% of which were deviants (R6) and the remaining were standards (R3). Stimulus duration was 340 ms and stimulus onset asynchrony (onset-to-onset) was 1000 ms. The blocks were divided into five sections of two minutes each, with a 15 s pause in between to give participants the possibility to move around and shift position during the experiment without reducing data quality. Standards and deviants were presented in pseudo-random order, with at least one standard always preceding each deviant. The first ten stimuli in every section were always standards, and were not included in the analysis. The exposure block consisted of 320 presentations of sounds along the whole continuum, with the frequency distribution of the sounds bimodal for half of the participants (henceforth the bimodal group) and unimodal for the other half of the participants (henceforth the unimodal group), see Figure 1. E-Prime 2.0.10.352 (Psychology Software Tools, Sharpsburg, Pennsylvania, USA) was used for stimuli presentation, and the total duration of the experiment was approximately 45 minutes.

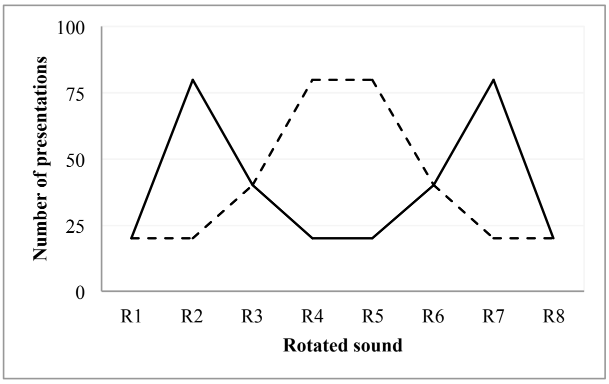

Figure 1: *The number of presentations of each rotated sound in the two groups during exposure. The solid line represents distribution of the sounds presented to the bimodal group and the dashed line represents the distribution of the sounds presented to the unimodal group. Proportions of the distributions taken from [6].*

2.4. Procedure

Participants were asked to fill out a consent form and a form asking for information about age, language background, and handedness. They were asked to select a movie to watch during the experiment, and the circumference of their heads was measured. Electrodes were placed above and below the left eye of the participants, as well as outside the lateral canthus of each eye and behind each ear on the mastoid bones. A head-cap was put on, and 16 electrodes were fastened (Fp1, Fp2, F3, F4, Fz, T7, T8, C3, C4, Cz, P3, P4, Pz, O1, O2 and Oz). The participants were instructed to remain as still as possible during stimuli presentation, and to move around, blink and stretch during the pauses. During the experiment, the participants watched their selected movie with no sound and Swedish subtitles. The movie window was centered on a screen in front of the participant, and its size was kept relatively small (approximately 15 by 20 cm), in order to minimize eye-movements. The purpose of letting the participants watch a movie, was to direct their attention away from the auditory stimuli while keeping them relatively alert throughout the experiment. Since the visual input from the movie was not time-locked to the auditory stimuli, it is assumed not to influence the ERPs systematically. The whole session, including preparations, lasted about an hour and a half.

2.5. EEG data

Electroencephalography (EEG) data were collected using the BioSemi ActiveTwo system, and the ActiView acquisition soft-

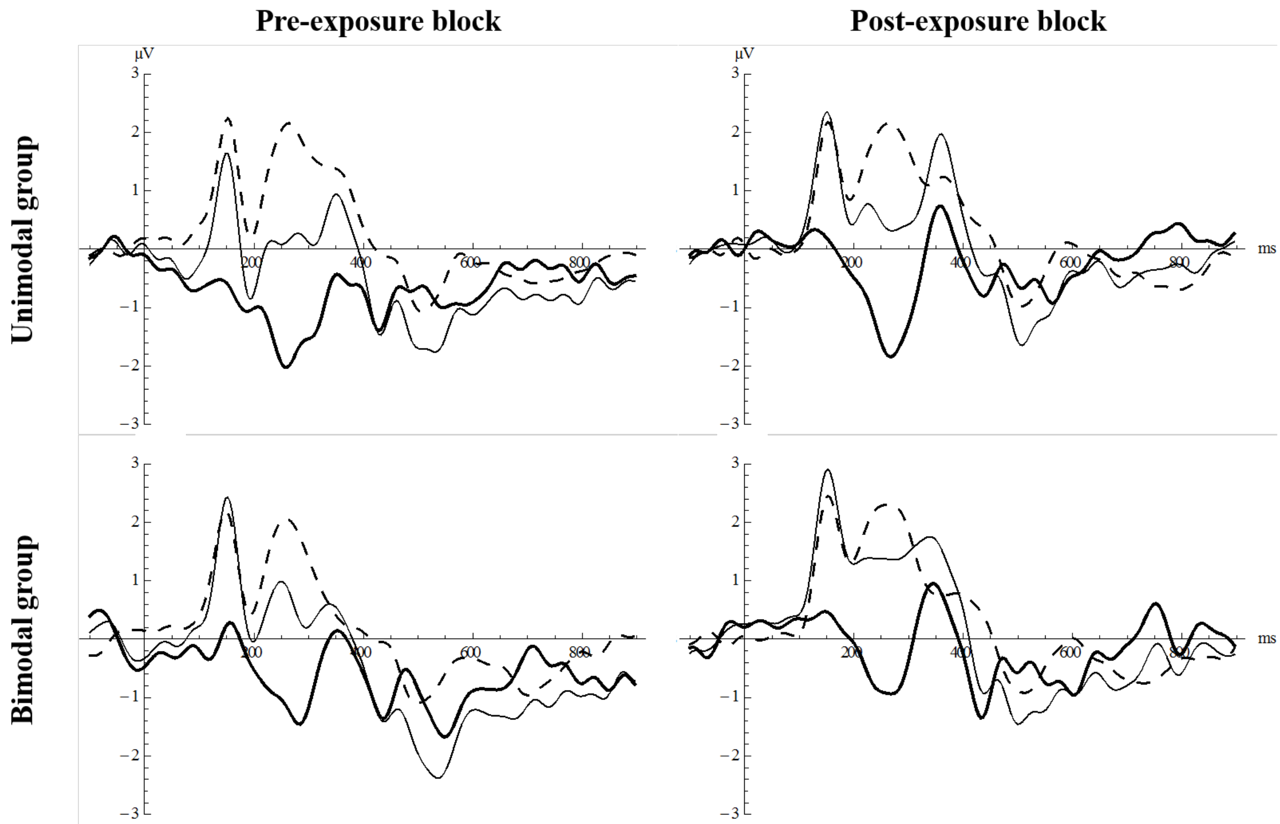

Figure 2: *The grand-average for standard (dashed line), deviant (solid thin line) and the difference waveform (solid thick line) obtained at the Fz electrode, for the unimodal group (top) and the bimodal group (bottom) separately, and for the pre-exposure block (left) and the post-exposure block (right) separately. A clear MMN can be seen for both groups in both blocks.*

ware (BioSemi, Amsterdam, The Netherlands). The sampling rate was set to 2048 Hz, and a driven-leg reference (a CMS/DRL loop with voltage recorded relative to the CMS electrode) was used during recording. Preprocessing of data was done in EEGLAB [36]. Data were re-referenced to the mastoid electrodes, re-sampled to 256 Hz and band-pass filtered with cutoff frequencies of 0.5 to 20 Hz. The twelve pauses were manually rejected by visual inspection before an Independent Component Analysis (ICA) was performed to isolate eye-blink and eye-movement artifacts. The most prominent component for each of those artifacts was identified and removed. Data were epoched, time locked to -100 ms to 900 ms relative to stimulus onset, and consequently baseline corrected. Lastly, trials likely containing artifacts or otherwise bad data (fvoltage lower than -50 μV or exceeding 50 μV in any EEG channel) were rejected. Data from Fz was used in the analysis, as this is where the difference response is typically strongest [37]. The mean amplitude of each subject average difference wave within the time window of 150 to 300 ms after stimulus onset was extracted for analysis. This time window was selected based on visual inspection of the grand average difference waveform, pooling all participants and conditions, and is in line with the time window in which the MMN is typically found [37]. Subsequent statistical analyses were carried out in SPSS 21 (International Business Machines Corp., Armonk, New York, USA).

3. Results

To test whether an MMN response was evoked in the participants, one-tailed one-sample t-tests were performed on the MMN amplitude separately for the pre-exposure and the post-exposure block. The t-tests revealed that an MMN was elicited in both blocks ($t(19) = -5.672$, $p < 0.001$ and $t(19) = -3.201$, $p = 0.005$, respectively). Figure 2 shows the grand average waveforms for standard and deviant at Fz, as well as the difference waveform, for the two blocks.

In order to test the effect of exposure to the different distributions on the MMN amplitude in the second block, a 2x2 repeated-measures ANOVA was performed with Block (pre-exposure vs. post-exposure) as within-subject variable and Group (unimodal vs. bimodal) as between-subjects variable. A main effect was found for Block ($F(1,18) = 5.278$, $p = .034$). The interaction between Block and Group was not significant ($F(1,18) = .039$, $p = .846$). Figure 3 shows the MMN amplitude for each group in the two blocks. Both groups have a smaller MMN amplitude in the post-exposure block than in the pre-exposure block.

4. Discussion

Contrary to expectations, no difference in MMN amplitude after exposure was found between groups. Instead, both groups had smaller MMN amplitude in the post-exposure block compared to the pre-exposure block. There is thus no indication of

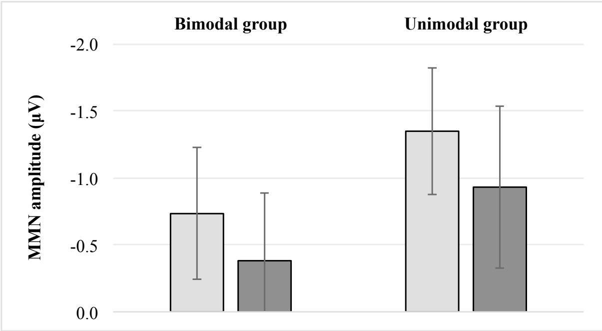

Figure 3: *Mean MMN amplitudes in μV for the bimodal group (left) and the unimodal group (right) in the pre-exposure block (light gray) and the post-exposure block (darker gray). Error bars show the 95% confidence interval. Note that negativity is plotted up.*

auditory perceptual category formation for the rotated sounds from passive exposure to different frequency distributions of the sound continuum.

The block effect is in line with a previous study, which has shown attenuated MMN amplitude within the last third of a block compared to the first third of a block [38].

The lack of group effect is however not in line with previous findings, which have demonstrated influence of passive exposure to different frequency distributions on the perception of speech sounds [23-29]. The present study differs from earlier studies on four counts, which may explain the discrepancy between the results of the present study and previous findings: the amount of exposure, the instructions to participants regarding the passive exposure, the use of a neurophysiological measure of categorization instead of behavioral measures, and the use of non-speech stimuli instead of speech stimuli.

The amount of exposure is critical. In a previous study, the ability to discriminate a non-native fricative contrast was influenced by 200 exemplars of exposure to a bimodal frequency distribution of a fricative continuum, but not by 100 exemplars of exposure [29]. However, the amount of exposure in the present study was 320 exemplars in total. This is substantially more exemplars than the previous studies on adults, which ranged from 128 exemplars [28] to 200 exemplars [29], in the cases where influence on speech perception was achieved. Other differences between the present study and earlier studies are therefore considered more likely to be the cause of the unexpected results.

Concerning participant instructions, in most previous similar studies [23-29] these were that they should listen carefully to the sounds during exposure, whereas in the present study participants were instructed to pay no particular attention to auditory stimuli. The discrepancy in attention to the sounds might influence the impact they have on perceptual categorization. In infant studies, where an effect has been found [6, 8, 9, 10], and where explicit instructions are in principle impossible to give, it is nevertheless reasonable to assume that the attention to exposure stimuli varies considerably between infants. It would be interesting to test the effect of attention to the exposure sounds explicitly in the future; however, this is not currently considered the most likely explanation for the results of the present study.

In the present study, the measure of categorization was the MMN amplitude, whereas in previous similar studies on adults, various behavioral measures have been used [23-29]. Intuitively, neurophysiological responses ought to be at least as sensitive as related behavioral responses, since the latter presumably are dependent upon the former. The MMN amplitude has been well documented as a neurological correlate to discrimination [30], and has been shown to be influenced by explicit discrimination training [39]. It is unlikely that the lack of a measurable effect in the present study is simply due to the present measure being less sensitive than those used earlier. The MMN response is however less studied as an explicit correlate to perceptual categorization. The rationale for using the MMN as an indicator of when two sounds are perceived as belonging to different categories, that is, categorization, is sound, but has not been explicitly investigated. It is conceivable that the enhanced MMN amplitude found when standard and deviant belong to different speech sound categories, compared to when they belong to the same speech sound category [31, 32], is not actually indicative of categorization, but caused by something else. This is an important methodological issue, warranting further study.

Lastly, the present study was designed to test whether the findings from previous studies [23-29] extended to non-speech stimuli, and it appears that they do not. This is unexpected, since previous research has shown that perceptual category formation is found for non-speech stimuli [17, 19, 18], and that the mechanisms involved are likely the same as in speech sound categorization to the extent of engaging brain areas typically considered specifically devoted to speech perception [40]. However, those previous studies all involved either explicit [17] or implicit training [18, 19, 40]. It is thus possible that although the mechanisms involved in perceptual category formation are domain-general, passive exposure to different frequency distributions does not have an impact on perception if a system of perceptual categories is not already in place (or alternatively, that they do, but more exposure is needed). If this is the case, it explains the results of the present study, and why they differ from previous similar studies [23-29].

In conclusion, the present study provides no evidence that passive exposure to different frequency distributions of a non-speech continuum induces perceptual categorization of the sounds, or otherwise influences the perception of them. There are multiple possible explanations for this, but the two considered most likely are that the MMN amplitude is not a suitable measure for perceptual categorization, or that passive exposure does not influence the perceptual categorization of sounds not already recognized as part of a categorization system. To disentangle which of the potential explanations of the present results in relation to previous studies is the most likely, it would be necessary to use a behavioral paradigm with the stimuli of the present study, and the paradigm of the present study with speech stimuli.

5. Acknowledgements

This study was funded by the Department of Linguistics, Stockholm University. We would like to thank all of our subjects for their participation, Petter Kallioinen and Hatice Zora for help with the EEG-processing, and Johannes Bjerva, Iris-Corinna Schwarz and David Pagmar for helpful comments on earlier versions of the manuscript.

6. References

[1] J. F. Werker and S. Curtin, “PRIMIR: A developmental framework of infant speech processing,” *Language learning and development*, vol. 1, no. 2, pp. 197–234, 2005.

[2] D. Maurer and J. F. Werker, “Perceptual narrowing during in-

fancy: A comparison of language and faces," *Developmental Psychobiology*, vol. 56, no. 2, pp. 154–178, 2014.

[3] P. K. Kuhl, E. Stevens, A. Hayashi, T. Deguchi, S. Kiritani, and P. Iverson, "Infants show a facilitation effect for native language phonetic perception between 6 and 12 months," *Developmental science*, vol. 9, no. 2, pp. F13–F21, 2006.

[4] P. K. Kuhl, F.-M. Tsao, and H.-M. Liu, "Foreign-language experience in infancy: Effects of short-term exposure and social interaction on phonetic learning," *Proceedings of the National Academy of Sciences*, vol. 100, no. 15, pp. 9096–9101, 2003.

[5] J. Lany and J. R. Saffran, "Statistical learning mechanisms in infancy," *Comprehensive developmental neuroscience: Neural circuit development and function in the brain*, vol. 3, pp. 231–248, 2013.

[6] J. Maye, J. F. Werker, and L. Gerken, "Infant sensitivity to distributional information can affect phonetic discrimination," *Cognition*, vol. 82, no. 3, pp. B101–B111, 2002.

[7] R. N. Aslin, "Experiential influences and sensitive periods in perceptual development: A unified model," in *Development of Perception: Psychobiological Perspectives*, R. N. Aslin, J. R. Alberts, and M. R. Petersen, Eds. Academic Press, 1981, pp. 45 – 93.

[8] J. Maye, D. J. Weiss, and R. N. Aslin, "Statistical phonetic learning in infants: Facilitation and feature generalization," *Developmental science*, vol. 11, no. 1, pp. 122–134, 2008.

[9] K. A. Yoshida, F. Pons, J. Maye, and J. F. Werker, "Distributional phonetic learning at 10 months of age," *Infancy*, vol. 15, no. 4, pp. 420–433, 2010.

[10] A. Cristià, G. L. McGuire, A. Seidl, and A. L. Francis, "Effects of the distribution of acoustic cues on infants' perception of sibilants," *Journal of phonetics*, vol. 39, no. 3, pp. 388–402, 2011.

[11] J. F. Werker, F. Pons, C. Dietrich, S. Kajikawa, L. Fais, and S. Amano, "Infant-directed speech supports phonetic category learning in English and Japanese," *Cognition*, vol. 103, no. 1, pp. 147–162, 2007.

[12] A. J. Lotto, M. Sato, and R. L. Diehl, "Mapping the task for the second language learner: The case of Japanese acquisition of /r/ and /l/," *From sound to sense*, pp. C381–C386, 2004.

[13] L. Lisker and A. S. Abramson, "A cross-language study of voicing in initial stops: Acoustical measurements," *Word*, vol. 20, no. 3, pp. 384–422, 1964.

[14] J. R. Saffran, E. K. Johnson, R. N. Aslin, and E. L. Newport, "Statistical learning of tone sequences by human infants and adults," *Cognition*, vol. 70, no. 1, pp. 27–52, 1999.

[15] N. Z. Kirkham, J. A. Slemmer, and S. P. Johnson, "Visual statistical learning in infancy: Evidence for a domain general learning mechanism," *Cognition*, vol. 83, no. 2, pp. B35–B42, 2002.

[16] J. R. Saffran, E. L. Newport, and R. N. Aslin, "Word segmentation: The role of distributional cues," *Journal of memory and language*, vol. 35, no. 4, pp. 606–621, 1996.

[17] L. L. Holt and A. J. Lotto, "Cue weighting in auditory categorization: Implications for first and second language acquisition," *The Journal of the Acoustical Society of America*, vol. 119, no. 5 Pt 1, pp. 3059–3071, 2006.

[18] T. Wade and L. L. Holt, "Effects of later-occurring nonlinguistic sounds on speech categorization," *The Journal of the Acoustical Society of America*, vol. 118, no. 3, pp. 1701–1710, 2005.

[19] Y. Gabay and L. L. Holt, "Incidental learning of sound categories is impaired in developmental dyslexia," *Cortex*, vol. 73, pp. 131–143, 2015.

[20] B. Blesser, "Speech perception under conditions of spectral transformation. i. phonetic characteristics," *Journal of Speech, Language, and Hearing Research*, vol. 15, no. 1, pp. 5–41, 1972.

[21] C. Narain, S. K. Scott, R. J. Wise, S. Rosen, A. Leff, S. D. Iversen, and P. M. Matthews, "Defining a left-lateralized response specific to intelligible speech using fMRI," *Cerebral Cortex*, vol. 13, no. 12, pp. 1362–1368, 2003.

[22] E. Marklund, F. Lacerda, and I.-C. Schwarz, "Using rotated speech to approximate the acoustic mismatch negativity response to speech," *Brain and Language*, Under review.

[23] J. Maye and L. Gerken, "Learning phonemes without minimal pairs," in *Proceedings of the 24th annual Boston University Conference on Language Development*, vol. 2. Citeseer, 2000, pp. 522–533.

[24] ——, "Learning phonemes: How far can the input take us?" in *Proceedings of the 25th annual Boston University Conference on Language Development*, vol. 1. Citeseer, 2001, pp. 480–490.

[25] R. Hayes-Harb, "Lexical and statistical evidence in the acquisition of second language phonemes," *Second Language Research*, vol. 23, no. 1, pp. 65–94, 2007.

[26] M. Gulian, P. Escudero, and P. Boersma, "Supervision hampers distributional learning of vowel contrasts," in *Proceedings of the 16th International Congress of Phonetic Sciences.*, vol. 2. Saarbrucken: University of Saarbrucken., 2007, pp. 1893–1896.

[27] P. Escudero, T. Benders, and K. Wanrooij, "Enhanced bimodal distributions facilitate the learning of second language vowels," *The Journal of the Acoustical Society of America*, vol. 130, no. 4, pp. EL206–EL212, 2011.

[28] K. Wanrooij, P. Escudero, and M. E. Raijmakers, "What do listeners learn from exposure to a vowel distribution? An analysis of listening strategies in distributional learning," *Journal of Phonetics*, vol. 41, no. 5, pp. 307–319, 2013.

[29] J. Gilkerson, "Categorical perception of natural and unnatural categories: Evidence for innate category boundaries," *UCLA working papers in linguistics*, vol. 13, pp. 34–58, 2005.

[30] N. Novitski, M. Tervaniemi, M. Huotilainen, and R. Näätänen, "Frequency discrimination at different frequency levels as indexed by electrophysiological and behavioral measures," *Cognitive Brain Research*, vol. 20, no. 1, pp. 26–36, 2004.

[31] I. Winkler, A. Lehtokoski, P. Alku, M. Vainio, I. Czigler, V. Csépe, O. Aaltonen, I. Raimo, K. Alho, H. Lang *et al.*, "Pre-attentive detection of vowel contrasts utilizes both phonetic and auditory memory representations," *Cognitive Brain Research*, vol. 7, no. 3, pp. 357–369, 1999.

[32] A. Sharma and M. F. Dorman, "Cortical auditory evoked potential correlates of categorical perception of voice-onset time," *The Journal of the Acoustical Society of America*, vol. 106, no. 2, pp. 1078–1083, 1999.

[33] K. Sjölander and J. Beskow, "Wavesurfer - an open source speech tool." in *Interspeech*, 2000, pp. 464–467.

[34] P. P. G. Boersma, "Praat, a system for doing phonetics by computer," *Glot international*, vol. 5, no. 9/10, 2002.

[35] M. B. Winn, "Formant continuum v. 30, april 2014," Retrieved from http://www.mattwinn.com/praat.html, November 16, 2016.

[36] A. Delorme and S. Makeig, "EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis," *Journal of Neuroscience Methods*, vol. 134, no. 1, pp. 9–21, 2004.

[37] R. Näätänen, P. Paavilainen, T. Rinne, and K. Alho, "The mismatch negativity (MMN) in basic research of central auditory processing: A review," *Clinical Neurophysiology*, vol. 118, no. 12, pp. 2544–2590, 2007.

[38] E. Sussman and I. Winkler, "Dynamic sensory updating in the auditory system," *Cognitive Brain Research*, vol. 12, no. 3, pp. 431–439, 2001.

[39] M. Atienza and J. L. Cantero, "Complex sound processing during human REM sleep by recovering information from long-term memory as revealed by the mismatch negativity (MMN)," *Brain research*, vol. 901, no. 1-2, pp. 151–160, 2001.

[40] R. Leech, L. L. Holt, J. T. Devlin, and F. Dick, "Expertise with artificial nonspeech sounds recruits speech-sensitive cortical regions," *Journal of Neuroscience*, vol. 29, no. 16, pp. 5234–5239, 2009.