10th International Conference on Speech Prosody 2020
25-28 May 2020, Tokyo, Japan

Revisiting infant distributional learning using event-related potentials: Does unimodal always inhibit and bimodal always facilitate?

Liquan Liu[1234], *Varghese Peter*[1], *Jia Hoong Ong*[5], *Paola Escudero*[23]

[1]School of Social Sciences and Psychology, Western Sydney University, Australia
[2]MARCS Institute for Brain, Behaviour and Development, Western Sydney University, Australia
[3]Centre of Excellence in the Dynamics of Language, Australian Research Council, Australia
[4]Centre of Multilingualism across the Lifespan, University of Oslo, Norway
[5]School of Psychology and Clinical Language Sciences, University of Reading, United Kingdom
`liquanl@uio.no`, `v.peter@westernsydney.edu.au`, `jiahoong.ong@reading.ac.uk`,
`Paola.Escudero@westernsydney.edu.au`

Abstract

Infants can learn and generalize phonetic categories through speech sound frequency distributions. Nevertheless, previous research with varying participant ages and testing paradigms reported incongruent findings regarding the effect of distributional learning of phonetic contrasts.

The current study examines infants' distributional learning of non-native tones using electroencephalography. 5-6-month-old Australian infants were exposed to an 8-step continuum of a Mandarin Chinese high-level vs. high-falling tonal contrast. The bimodal condition had frequency peaks near the two ends of the continuum (steps 2, 7) whereas the peak was at the midpoint of the unimodal condition (steps 4, 5). Before and after listening to their corresponding distribution, both groups were tested on the same sounds (steps 3, 6) in a passive oddball paradigm.

The unimodal group (N = 8) showed strong sensitivity to the sound distinction at post- but not pre-distributional learning. The bimodal group (N = 8), no significant neural sensitivity or difference was observed in pre- or post-distributional learning. The finding that unimodal exposure enhances infant perception is novel and is explained by their acoustic sensitivity to peak location, highlighting the role of the magnitude of the acoustic distinction in the stimuli when prior training and exposure is insufficient to establish phonetic categories.

Index Terms: distributional learning, speech perception, lexical tone, perceptual salience, electroencephalography (EEG), event-related potentials (ERP), mismatch negativity (MMN), (positive) mismatch response (MMR)

1. Introduction

Statistical learning, the ability to extract and learn from the statistical regularities in the ambient environment, is a fundamental learning mechanism that drives infants' phonological acquisition since the beginning of life [1]. This mechanism has been argued to underlie infants' phonetic [2], semantic [3] and grammatical [4] acquisition. In this paper, we revisit previous mixed results by using a more sensitive method, namely infants' brain responses to a stimulus change before and after exposure to two types of distributional information.

To assess how input statistics can alter infants' rapid phonetic category discrimination and learning, Maye and colleagues [5,6] created a speech sound continuum of a voiced-voiceless unaspirated stop consonant contrast ([da]–[ta]) and stimuli were arranged with unimodal or bimodal frequency distributions. The two distributions differed in the number of Gaussian peaks (tokens with relatively high frequency) along the [t-d] continuum (Figure 1). The unimodal distribution was marked by one peak corresponding to single category learning, whereas a bimodal distribution was characterized by two peaks, resembling the learning of two categories. After familiarization to either distribution, infants' discrimination ability of the tokens presented with equal frequency in the two distributions (steps 3-6) was measured. Contrast perception was reduced after unimodal [5] and enhanced after bimodal [6] exposure.

After this seminal paper, studies have shown similar outcomes with respect to the influence of exposure to unimodal versus bimodal sound distributions from infancy [3,7-12] to adulthood [13-15]. Findings suggest that distributional learning is a viable learning mechanism used in sound perception and learning across the lifespan. Nevertheless, a recent study showed that only infants at 11 but not at 5 months showed enhanced sound perception after bimodal exposure and reduced perception after unimodal. Specifically, the 5-month-olds discriminated a non-native tone contrast after both unimodal and bimodal exposure [11]. This result somewhat conforms to a meta-analysis aiming at a comprehensive look on infant distributional learning: Cristia [16] compared 11 behavioural experiments from 6 studies and found that both age and testing paradigm influenced the outcome. In particular, this learning mechanism was more robust among older infants in the first twelve months after birth. Moreover, the learning effect was only robust in Habituation-Dishabituation but not in Familiarization-Alternating/Non-Alternating paradigms.

As a widely-used neurophysiological measure, event-related potentials (ERP) provide a new way of investigating the discrepancies among previous studies. ERPs are derived from the electroencephalogram (EEG), and they show the brain responses to specific events. The ERP component mismatch negativity (MMN) is widely used as a measure of discrimination (in infants MMN is seen as a positivity, referred to as mismatch response, MMR [17]). MMN and MMR reflect the neural basis of acoustic-phonetic processing distinct from behavioural methods [18-21]. To date, only one study that we

10.21437/SpeechProsody.2020-67

know of has examined infant distributional learning using EEG. Two-to-three-month-old Dutch infants' distributional learning was examined via unimodal and bimodal distribution of an English /æ/-/ɛ/ contrast during their sleep [22]. Interestingly, an interaction was reported between distribution type and vowel: bimodally-trained infants discriminated better if the deviant was [æ] and unimodally trained infants discriminated better if the deviant was [ɛ]. Crucially, unimodal training did not inhibit but rather facilitated phonetic learning.

Recent adult studies may seem to further contest the facilitating role of bimodal training. German listeners' neural discrimination after training with a bimodal distribution of a Cantonese (high vs. mid) pitch height tonal contrast improved across all steps along the tonal continuum [23], with online neural improvement elicited during the learning phase [24]. A topography analysis also revealed that the MMN difference disappeared by the end of the experiment, indicating listeners' sensitivity to all tokens along the pitch height tonal continuum.

Different from the outcomes of most behavioural experiments, the two EEG studies using lexical tone as the non-native carrier jointly indicate that exposure to a bimodal distribution may not necessarily lead to the enhancement of specific steps or categories along the pitch continuum, but may increase listeners' overall tonal perception. Importantly, these studies did not include training with a unimodal distribution.

To increase our understanding of infant distributional learning, the current study tested 5-6-month-olds, the age at which distributional learning is arguably less robust [16], using EEG techniques. Specifically, we examined how 5-6-month-old Australian infants exhibited neural sensitivity to non-native tones after exposure to unimodal versus bimodal distributions. From previous research, we predicted increased neural sensitivity after bimodal but not unimodal exposure.

2. Method

2.1. Participants

Twenty typically developing infants participated in the study. All participants were naïve to tone or pitch-accent languages. The final sample consisted of 16 5-6-month-old infants being raised in monolingual Australian English households. Attention was captured by silent child-friendly videos during the EEG recording. Four participants were tested but were subsequently excluded from analysis due to excessive artefacts in their EEG data (see EEG Recording and Analysis below). The study protocol was approved by the Western Sydney University Human Research Ethics Committee (H11383). All families provided their written informed consent prior to the experiment.

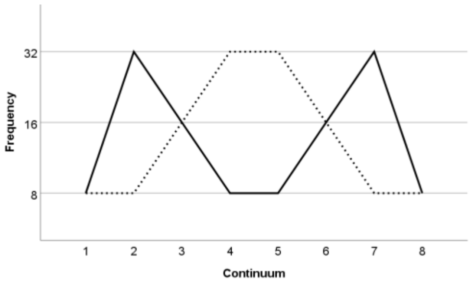

Figure 1: *Example of unimodal (dotted line) and bimodal (solid line) frequency distributions* [11].

2.2. Stimuli

A female Mandarin speaker produced natural tokens of the Mandarin high-level (T1) and high-falling (T4) tones with /ta/ as the tone bearing syllable in a soundproof booth at the phonetics lab of Utrecht Institute of Linguistics OTS, Utrecht University. Tokens were recorded by computer program Audacity. The tone pair was further divided into six equidistant in-between tokens which resulted in an eight-step continuum (Figure 2) via the computer program PRAAT [25]. Stimulus intensity was set to 65 dB and duration to 400ms. Five native adult Mandarin Chinese speakers listened to the stimuli and confirmed that they were acceptable tokens of Mandarin Chinese. The Mandarin T1–T4 contrast has been shown to be discriminated by adult speakers of tone [26,27] and non-tone [27,28] languages as well as non-tone-language learning infants [10].

2.3. EEG Paradigm

The experiment consisted of three phases: (1) EEG recording: pre-training, (2) distributional learning, and (3) EEG recording: post-training. At pre- and at post-training phases, participants were presented with 2 oddball blocks. To counterbalance, stimulus step 3 was presented as the standard (presented 80% of the time) and step 6 was presented as the deviant (20% of the time) in one block, and the standards and deviants were reversed in the other block. There were 500 stimuli in each oddball block (400 standards, 100 deviants).

In the distributional learning phase, participants were randomly assigned to one of two conditions: unimodal or bimodal distribution. Each of the 8 tokens was played with the same distributions as in Figure 1 but repeated twice (128 tokens x 2 times). The two conditions had different distributional peaks (a single central category vs. two separate categories) but were equal in the total amount of distributional learning tonal input (256 tokens) and duration (6 minutes). Stimulus steps 3 and 6 were presented an equal number of times in both unimodal and bimodal conditions. The stimulus presentation was controlled using Presentation 18.1 (Neurobehavioral Systems Inc.).

2.4. EEG Recording

EEG was recorded using a 128 channel Hydrocel Geodesic Sensor Net (HCGSN), Netamps 300 amplifier and NetStation 5.1 recording software (EGI Inc.) Infants sat on the caregiver's lap and watched either an age-appropriate silent video or a puppet show while the sounds the . The continuous EEG was recorded at a sampling rate of 1000 Hz with the ground reference electrode placed at the Cz scalp location.

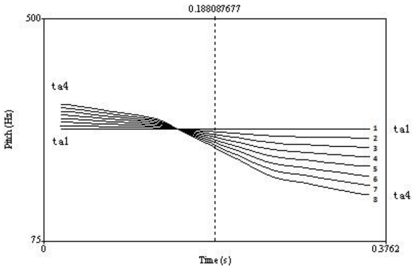

Figure 2: *Pitch contours along a [taT1-taT4] continuum* [11].

2.5. EEG Analysis

The EEG was analysed using fieldtrip toolbox [29] in MATLAB 2019a. The continuous EEG was first bandpass filtered between 0.3-20 Hz. The EEG was then divided into epochs between -100 to 400 ms, relative to sound onset. Epochs were baseline corrected between -100 and 0 ms. Noisy channels and trial rejection were determined as follows: If a trial has amplitudes exceeding ±100μV at any time point in more than 20 channels, the trial was rejected. If the number of bad channels was less than 20, the trial was kept, and the channels with amplitudes exceeding ±100μV were interpolated. If a channel was noisy for more than 50% of the trials, that channel was interpolated for all the trials. Participants with less than 50 good trials in each condition were excluded. EEG was then re-referenced to the average of the mastoids. Trials were averaged separately for deviants and standards to get the ERP waves. Difference waves were computed by subtracting the ERP for the standard from the deviant. To improve the signal to noise ratio, the ERPs for the two pre-training blocks were averaged together. Similar averaging was done for the post-training block as well. Individual ERP waves were averaged to create grand-averaged ERPs.

2.6. Statistical Analysis

The significance of the MMN/MMR was measured by calculating the average amplitude in the standard and deviant waveforms between 100-400 ms in 100 ms steps at the fronto-central electrode (FCz). The time window of 100-400 ms was selected as it encompasses both the MMN and MMR ranges. This amplitude values were subjected to a paired sample t-test between standard and deviant.

2.7. Results

The grand averaged standard and deviant waveforms showed a broad positive response between 100-200 ms (P1) as expected from infants of this age (Figure 3). The grand averaged deviant-standard waveforms are shown in Fugure 4. Results of the amplitude comparison between standard and deviants across 100-400 time windows are listed in Table 1. As the number of participants per group was small ($N = 8$), the two groups were analyzed separately. Results of *t-tests* were summarized in Table 1 below. In the unimodal group, there was no significant MMN or MMR in the pre-training phase whereas a significant MMR was seen between 300-400 ms post-training in the unimodal group. In the bimodal group, although marginal significances were elicited in both pre- and post-training phases ($ps < .079$), the influence of training, evaluated by comparing the MMR amplitudes pre- and post-training, did not show any significant effects $t(6) = 0.20$, $p = .84$.

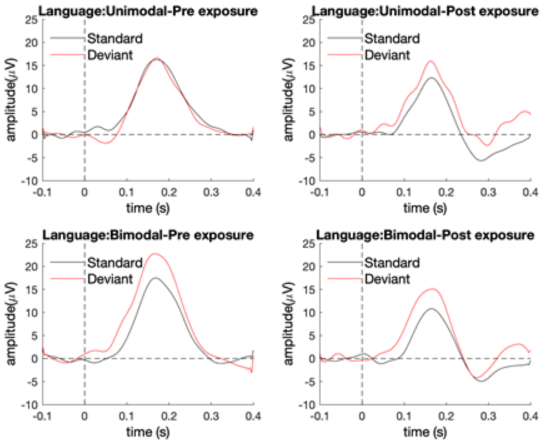

Figure 3. *Deviant and standard waveforms for before (left panel) and after (right panel) training for unimodal (top panel) and bimodal (bottom panel) distributions. As the number of participants per group is small (N = 8), the pre-training difference may dissipate with more infants per group.*

Table 1. The significance of grand averaged difference between standard and deviant across 100-400 time windows.

	Unimodal		Bimodal	
ms	Pre	Post	Pre	Post
100-200	$p>.05$	$p>.05$	$p>.05$	$p>.05$
200-300	$p>.05$	$p>.05$	$p>.05$	$p>.05$
300-400	$p>.05$	$p=.002$	$p>.05$	$p>.05$

3. Discussion

To further our understanding of distributional learning, we examined Australian-English learning infants' neural sensitivity to non-native lexical tones using a passive oddball listening paradigm in pre- and post-training phases with a distributional learning phase in between. The same methodology has been used in behavioural and EEG studies to show phonetic category formation. Improved MMR amplitudes were observed only after unimodal but not bimodal exposure, contrary to prediction. Given the relatively small sample size, results need to be interpreted with caution and trends in the two distributions were discussed separately.

Infants exhibited near-identical MMRs to the contrast before and after bimodal distribution. The non-significant difference between pre- and post-training phases makes it difficult to disambiguate whether bimodal exposure enhanced the particular contrast as in previous behavioural studies [5,6], or led to a general enhancement across steps which benefited overall tone perception as in previous neural studies [23,24], or else, did not alter perception. In the first scenario, under relatively high neurocognitive plasticity [30], 5-6-month-old infants may face little constraint by top-down influences, maintaining their less rigid sensitivity to acoustic differences after bimodal distribution. However, this would predict a change from pre- to post-training which was not observed in our study. With respect to the second scenario, the distributional information in the bimodal condition may be insufficient for infants to establish proto-tonal categories and alter infants' discrimination.

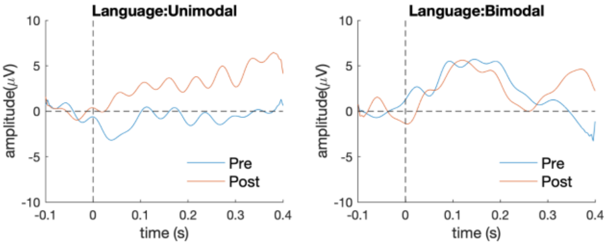

Figure 4. *Mismatch responses before (Pre) and after (Post) unimodal (left) and bimodal (right) distributions. Our preliminary results may show that hearing the two very similar tokens may enhance sensitivity to their small differences.*

On the other hand, facilitation was observed in the unimodal condition, contrary to outcomes from behavioural studies. Unlike those in the bimodal condition, infants in the unimodal condition did not exhibit evidence of target contrast discrimination before training. In other words, the two distribution groups may not have started at the same level prior to the distributional learning phase. This is likely due to a combination of factors such as individual differences in pitch perception, sampling error, and the number of participants. With that in mind, any group differences should be interpreted with caution. Nonetheless, according to the traditional view of distributional learning, unimodal participants should *not* show any improvement; yet, it appeared they do, at least here and in the only other EEG study on infant distributional learning [22].

We argue that any type of distributional exposure to sounds (and non-native ones in particular) may show a practice effect, allowing listeners to hear a difference regardless of the frequency distributions when they were not sensitive to the contrast in the first place. Alternatively, similar to the attunement to the acoustic difference explanation for the bimodal infants, infants may have a tendency to process stimuli efficiently and may therefore focus on the most frequently presented (i.e., most salient/prominent [31,32]) steps (peaks of distribution) in the distribution which would impact subsequent perception. Note that infants have shown different performances regarding whether a contrast is easy to perceive [33]. Training on a unimodal distribution in which the peak contrast (steps 4-5) is extremely difficult to discriminate may ease the processing of a contrast with a larger acoustic difference (i.e., steps 3-6). This efficiency may be grounded in infant cognitive processing. Infants rely on existing heuristics to maximise efficiency and conservation of energy. When such strategies on input statistics break down, non-native listeners' neural sensitivity may be insufficient to discern the fine-grained tone continuum steps. This explanation assumes interactions between acoustic and statistical cues. Listeners across ages are able to abstract and retain the memory of non-native pitch directional cues as well as shift their acoustic/phonetic cue weighting and learning strategies [34,35]. When various types of cues (e.g., acoustic feature, frequency distribution) are presented, listeners' weighting of these cues may be dynamic and change in real-time during experimental training [36,37]. One study found that when both prosodic and statistical cues are presented, listeners rely on prosodic cues in speech segmentation [38]. Crucially, a unimodal distribution does not always inhibit perception. Perceptual constraints such as the acoustic properties and distance between tokens may exist for distributional learning,

The overall EEG findings appear to mostly align with the heuristic explanation. From a theoretical perspective, certain perceptual constraints like acoustic distance may play a role when the listeners do not use the cues phonologically. Since the infants did not have experience using lexical tones for their phonological categories, acoustic differences may influence the results differently from other contrasts. However, prosody is used in non-tone languages in other forms such as intonation, stress, etc., rendering it difficult to predict whether and how listeners might treat pitch information.

From a methodological perspective, it is worth mentioning that all distributional learning studies using similar EEG techniques across ages [22-24], including the current study, have reported neural discrimination patterns distinct from traditional bimodal-facilitation and/or unimodal-inhibition effects. Distribution-induced processing changes appear to be a more complex measure than the previously reported simple, binary distinction. Studies with tones show that tone perception seems to emerge earlier than vowel and consonant native-like perception [39]. We believe that factors such as statistical information, stimuli acoustics (properties, distance, etc.), accumulated linguistic experience, and processing capacity may all play a role.

Our findings suggest that EEG may be more sensitive than behavioural measures in phonetic discrimination at least when reporting outcomes with large individual variations, which have been found in the discrimination of tone contrasts. A larger number of subjects may be required to reach a better picture of the overall perceptual pattern. We leave this investigation for future studies.

4. Conclusions

This study examined Australian-English learning infants' neural responses before and after exposure to unimodal or bimodal distributions of a non-native lexical tone contrast. In the bimodal condition, no difference was observed pre- and post-distributional exposure. In the unimodal condition, no sensitivity was observed before training, but the neural response was enhanced post-training. These results indicate a previously ignored effect of distributional learning where unimodal exposure may also lead to facilitation in the sensitivity to difficult novel phonetic categories along the continuum of tonal distributions. Factors including but not limited to statistical information, stimuli acoustics, accumulated linguistic experience, and processing capacity may all play a role in infants' distributional learning outcomes. Overall, results reflect listeners' sensitivity to ambient acoustic and statistical information, paving the way for perception and learning of contrasts in foreign language speech.

5. Acknowledgements

The experiment was funded by Australian Research Council grants CE140100041 (L.L. & P.E.) and FT160100514 (P.E.), and by Western Sydney University grants: Early Career Researcher [20311.87608], Research Fund [20311.82935] and Start-up [20311.86565]. For writing, the first author has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 798658 and is hosted by Center for Multilingualism across the Lifespan at the University of Oslo, financed by Research Council of Norway through its Centers of Excellence funding scheme grant agreement No. 223265.

6. References

[1] J. R. Saffran, R. N. Aslin, and E. L. Newport, "Statistical learning by 8-month-old infants." *Science*, vol. 274, no. 5294, pp. 1926-1928, 1996.

[2] P. Escudero, and D. Williams. "Distributional learning has immediate and long-lasting effects." *Cognition*, vol. 133, no. 2, pp.408-413, 2014.

[3] S. M. M. ter Schure, C. M. M. Junge, and P. P. G. Boersma. "Semantics guide infants' vowel learning: computational and experimental evidence." *Infant Behavior and Development*, vol. 43, pp.44-57, 2016.

[4] S. P. Thompson, and E. L. Newport. "Statistical learning of syntax: The role of transitional probability." *Language learning and development*, vol. 3, no. 1, pp.1-42, 2007.

[5] 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.

[6] 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.

[7] D. J. H. Capel, E. H. De Bree, M. A. De Klerk, A. O. Kerkhoff, and F. N. K. Wijnen. "Distributional cues affect phonetic discrimination in Dutch infants." *Sound and sounds. Studies presented to MEH (Bert) Schouten on the occasion of his 65th birthday. Utrecht: UiL-OTS*, pp.33-43, 2011.

[8] L. Liu, and R. Kager, “How Do Statistical Learning and Perceptual Reorganization Alter Dutch Infant's Perception to Lexical Tones?” in *ICPhS*, vol. 17, pp. 1270-1273, 2011.

[9] L. Liu, and R. Kager, "Is perceptual reorganization affected by statistical learning? Dutch infants’ sensitivity to lexical tones." in *Proceedings of the 35th annual Boston University conference on language development, Cascadilla Press, Somerville*, pp. 404-413, 2011.

[10] L. Liu, and R. Kager, "Perception of tones by infants learning a non-tone language." *Cognition*, vol. 133, no. 2 pp.385-394, 2014.

[11] L. Liu, and R. Kager, "Statistical learning of speech sounds is most robust during the period of perceptual attunement." *Journal of experimental child psychology*, vol. 164, pp.192-208, 2017.

[12] 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.

[13] P. Escudero, T. Benders, and K. Wanrooij, K. "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.

[14] J. H. Ong, D. Burnham, C. J. Stevens and P. Escudero. "Naïve learners show cross-domain transfer after distributional learning: the case of lexical and musical pitch." *Frontiers in psychology* 7, pp. 1189, 2016.

[15] J. H. Ong, D. Burnham, and C. J. Stevens. "Learning novel musical pitch via distributional learning." *Journal of Experimental Psychology: Learning, Memory, and Cognition*, vol. 43, no. 1, pp.150, 2017.

[16] A. Cristia, "Can infants learn phonology in the lab? A meta-analytic answer." *Cognition*, vol. 170, pp.312-327, 2018.

[17] C. He, L. Hotson, and L. J. Trainor, "Mismatch responses to pitch changes in early infancy." *Journal of Cognition Neuroscience*. vol. 19, pp. 878–892, 2007.

[18] N. Kraus, T. McGee, T. D. Carrell, and A. Sharma. "Neurophysiologic bases of speech discrimination." *Ear and hearing*, vol. 16, no. 1, pp.19-37, 1995.

[19] S. C. Lipski, P. Escudero, and T. Benders. "Language experience modulates weighting of acoustic cues for vowel perception: An event‐related potential study." *Psychophysiology*, vol. 49, no. 5, pp.638-650, 2012.

[20] R. Näätänen, and I. Winkler. "The concept of auditory stimulus representation in cognitive neuroscience." *Psychological bulletin*, vol. 125, no. 6, pp.826, 1999.

[21] M. Sams, K. Alho, and R. Näätänen. "Short‐term habituation and dishabituation of the mismatch negativity of the ERP." *Psychophysiology*, vol. 21, no. 4, pp.434-441, 1984.

[22] K. Wanrooij, P. Boersma, and T. Van Zuijen. "Fast phonetic learning occurs already in 2-to-3-month old infants: an ERP study." *Frontiers in psychology*, vol. 5, pp.77, 2014.

[23] J. S. Nixon, N. Boll-Avetisyan, T. O. Lentz, S. van Ommen, B. Keij, C. Cöltekin, L. Liu, and J. van Rij. "Short-term exposure enhances perception of both between-and within-category acoustic information." In *Proceedings of the 9th International Conference on Speech Prosody, 13-16 June 2018, Poznan, Poland*, pp. 114-118.

[24] N. Boll-Avetisyan, J. S. Nixon, T. O. Lentz, L. Liu, S. van Ommen, Ç. Çöltekin, and J. van Rij. "Neural Response Development During Distributional Learning." In *Interspeech*, 2018, pp. 1432-1436.

[25] P. Boersma, and D. Weenink. "Praat: doing phonetics by computer [Computer program]. Version 5.3. 51." *Online: http://www. praat. org/retrieved, last viewed on* 12, 2013.

[26] T. Huang, and K. Johnson. "Language specificity in speech perception: Perception of Mandarin tones by native and nonnative listeners." *Phonetica*, vol. 67, no. 4, pp.243-267, 2010.

[27] A. Chen, L. Liu, and R. Kager, "Cross-linguistic perception of Mandarin tone sandhi." *Language Sciences*, vol. 48, pp.62-69, 2015.

[28] L. Liu, A. Chen, and R. Kager, "Perception of tones in Mandarin and Dutch adult listeners." *Language and Linguistics*, vol. 18, no. 4, pp.622-646, 2017.

[29] R., Oostenveld, P. Fries, E. Maris, & J. M. Schoffelen, J. M., "FieldTrip: open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data." *Computational intelligence and neuroscience,* vol. 1, 2011.

[30] J. F. Werker, and R. C. Tees. "Cross-language speech perception: Evidence for perceptual reorganization during the first year of life." *Infant behavior and development* vol. 7, no. 1, pp. 49-63, 1984.

[31] B. Chandrasekaran, J. T. Gandour, and A. Krishnan. "Neuroplasticity in the processing of pitch dimensions: A multidimensional scaling analysis of the mismatch negativity." *Restorative neurology and neuroscience* vol. 25, no. 3-4, pp. 195-210, 2007.

[32] B. Chandrasekaran, J. Hornickel, E. Skoe, T. Nicol, and N Kraus. "Context-dependent encoding in the human auditory brainstem relates to hearing speech in noise: implications for developmental dyslexia." *Neuron* vol. 64, no. 3, pp. 311-319, 2009.

[33] P. Escudero, and M. Kalashnikova. "Infants use phonetic detail in speech perception and word learning when detail is easy to perceive." *Journal of Experimental Child Psychology*, vol. 190, pp. 104714, 2020.

[34] P. Escudero, T. Benders, and S. C. Lipski. "Native, non-native and L2 perceptual cue weighting for Dutch vowels: The case of Dutch, German, and Spanish listeners." *Journal of Phonetics* vol. 37, no. 4, pp. 452-465, 2009.

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

[36] E. D. Thiessen, and J. R. Saffran. "When cues collide: use of stress and statistical cues to word boundaries by 7-to 9-month-old infants." *Developmental psychology* vol. 39, no. 4, pp.706, 2003.

[37] A. Origlia, F. Cutugno, and V. Galatà. "Continuous emotion recognition with phonetic syllables." *Speech Communication* vol. 57, pp. 155-169, 2014.

[38] M. Marimon, amd B. Höhle. "Prosody Outweighs Statistics: Evidence from German." Paper presented at the *Biennial International Conference on Infant Studies (ICIS)*, Philadelphia, 2018.

[39] Yeung, H. H., Chen, K. H., & Werker, J. F. (2013). When does native language input affect phonetic perception? The precocious case of lexical tone. *Journal of Memory and Language*, *68*(2), 123-139.