



# Towards a Neural Measure of Perceptual Distance—Classification of Electroencephalographic Responses to Synthetic Vowels

Manson C-M. Fong<sup>1,2</sup>, James W. Minett<sup>1,2</sup>, Thierry Blu<sup>1</sup>, William S-Y. Wang<sup>1,2</sup>

<sup>1</sup>Department of Electronic Engineering, The Chinese University of Hong Kong, Hong Kong, China

<sup>2</sup>Joint Research Centre for Language and Human Complexity,  
The Chinese University of Hong Kong, Hong Kong, China

cmfong@ee.cuhk.edu.hk, jminett@ee.cuhk.edu.hk, tblu@ee.cuhk.edu.hk, wsywang@ee.cuhk.edu.hk

## Abstract

How vowels are organized cortically has previously been studied using auditory evoked potentials (AEPs), one focus of which is to determine whether perceptual distance could be inferred using AEP components. The present study extends this line of research by adopting a machine-learning framework to classify evoked responses to four synthetic mid-vowels differing only in second formant frequency (F2 = 840, 1200, 1680, and 2280 Hz). 6 subjects attended 4 EEG sessions each on separate days. Classifiers were trained using time-domain data in successive time-windows of various sizes. Results were the most accurate when a window of about 80 ms was used. By integrating the scores from individual classifiers, the maximum mean binary classification rates improved to 70% (10 trials) and 77% (20 trials). To assess how well perceptual distances among the vowels were reflected in our results, discriminability indices ( $d'$ ) were computed using both the behavioral results in a screening test and the classification results. It was found that the two set of indices were significantly correlated. The pair that was the most (least) discriminable behaviorally was also the most (least) classifiable neurally. Our results support the use of classification methodology for developing a neural measure of perceptual distance.

**Index Terms:** vowel perception, electroencephalography, perceptual distance, classification, machine learning

## 1. Introduction

It is well-established that vowel perception hinges on the extraction of formant frequencies, especially the first two formants, F1 and F2 [1, 2, 3]. However, it is also recognized that the neural representation of speech in general and of vowels in particular is significantly altered by language experience [4, 5], thus it is unlikely that the perceptual vowel space can be described simply on acoustical ground using formant parameters. In order to shed light on the structure of the perceptual vowel space, behavioral research has made frequent use of the same-different judgment task (see [6] for an early review). Using the signal detection theory [7] to analyze the results of this task, the perceptual distance between a pair of vowels can be quantified using the discriminability index ( $d'$ ). Similarity judgment (e.g. [8, 9]), in which subjects judge the similarity of a pair of vowels on a scale (e.g., 1-10), has also been used to obtain a measure of perceptual distance directly. Perceptual distances can then be used to construct the perceptual vowel space by means of techniques such as multi-dimensional scaling (MDS) [10]. Such studies suggest that the perceptual vowel space can be thought of as being distorted or warped—acoustic differences that do not form a contrast in a language are shrunk while phonological

distinctions that are made in a particular language are stretched. Despite such success in mapping the perceptual space, our understanding remains incomplete with regard to the underlying cortical networks that give rise to such behaviors.

The aim of the present study is to use scalp electroencephalography (EEG) to examine one global organizational principle of vowels, namely, that vowels that are more perceptually distinct (as measured behaviorally) elicit neural responses that are more distinct. While this hypothesis is intuitive, it has not been directly tested using EEG. Instead, a similar hypothesis has only been tested using cortical local field potentials recorded using intracranial, high-density cortical surface arrays [11]. In that study, the responses to a continuum consisting of 14 synthesized speech syllables (perceived as either /ba/, /da/ or /ga/ according to the value of the onset frequency of the F2 transition) were recorded by surface electrodes placed over the superior temporal gyrus. By performing binary classification over all stimulus pairs, the authors were able to derive the perceptual distances for all stimulus pairs, from which a neurometric discrimination function was determined. This neurometric function reflected phonetic boundaries that were similar to those of the behavioral discrimination function.

Thus far, to our knowledge, only one study has classified scalp EEG responses to phonemes [12]. In this study, the focus was to classify EEG responses to phonemes using a hierarchical scheme according to distinctive features defined in traditional phonetics (e.g., [13]). Emphasis was put mainly on consonants, and it was found that EEG responses to consonants could be classified with an accuracy of 67% using a recognition model that fused three consonantal features (continuant, voicing, and place) hierarchically. This study also classified the EEG responses to the four English vowels, /i/, /æ/, /u/ and /a/, located at the four corners of the vowel quadrilateral. This choice was made to encourage classifiable EEG signals. The accuracy was up to 90.0% when hierarchical recognition was performed. These results suggested that phonemic information contained in EEG is sufficiently robust for accurate classification.

The present study draws inspiration from the two studies above in using neural classification performances to derive a metric of perceptual distance for vowels. In our case, instead of using classification results to find evidence that distinctive features are represented in the brain, our focus is to test whether neural discriminability of vowels correlates significantly with behavioral discriminability. In order to avoid ceiling effects (when all vowel pairs are equally discernable), we chose our stimuli to encourage confusion such that a varying degree of behavioral discriminability can be obtained. Specifically, we recorded EEG responses to four mid-vowels: [ɔ], [ɻ], [œ] and

[ɛ]. In Cantonese, the native language of the subjects in this study, only [ɔ], [œ] and [ɛ] are present but not [ɣ], which is present in Mandarin. Based on the behavioral results, it was determined that [ɣ] and [œ] exhibited considerable confusions. We took advantage of this to examine the correlation between behavioral dissimilarity and neuronal dissimilarity.

## 2. Materials and methods

### 2.1. Subjects

Six healthy native Cantonese participants (S1–S6, 3 male and 3 female), aged 19–21 (mean: 20.1, SD: 0.8), who were all undergraduate students, took part in the experiment. All subjects were right-handed [14]. S1–S4 were paid at a rate of about US \$7 per hour, while S5 and S6 were volunteers. Informed consent was obtained from each subject.

### 2.2. Stimuli

Synthetic stimuli with simple spectral characteristics were synthesized by summing two formants, each represented as a three-tone complex comprising three pure tones added in phase [15]. The middle tone was set to the value of the specified formant frequency, and with the frequency of the other two tones differing from the middle tone by  $\pm F_0$  ( $F_0 = 120$  Hz). Stimuli differ only in  $F_2$  ( $F_2 = 840, 1200, 1680, \text{ and } 2280$  Hz). Figure 1 shows the spectral representations of the four synthetic vowels. A single pure tone at a frequency of  $F_0$  was also added to improve naturalness. The duration of each vowel is 400 ms, and a cosine ramp of 20 ms was applied to the onset and offset portions of the stimuli.

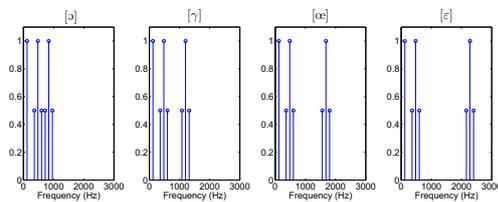


Figure 1: Spectral representations of the synthetic vowels.

### 2.3. Screening test

Prior to the main EEG experiment, all 6 subjects were given a screening test in order to ensure that they perceived the set of synthetic vowels as being phonologically similar to the intended vowels. Specifically, a same–different judgment task was adopted. In each trial, a pair of vowels—one synthetic and one naturally recorded from a native Cantonese–Mandarin bilingual male speaker—were played in succession (inter-stimulus interval about 2 s). Subjects were asked to judge whether they heard the same vowel or different vowels. 16 trials were administered for each of the 16 ( $4 \times 4$ ) vowel pairs. The criterion for inclusion was that the  $d'$  score computed for all four vowels should be significantly different from 0 ( $p < 0.01$ , Bonferroni corrected for multiple comparisons), which, according to signal detection theory, would indicate that the “same” (signal) trials were perceptually distinct from the “different” (noise) trials. Overall, 27 subjects sat the screening test, 22 of whom fulfilled the criterion above. S1–S4 were among the first few subjects who passed the screening test, while the two volunteers (S5–S6) were admitted into the study at a later stage. The remaining subjects were invited to sit a parallel study to the present one.

### 2.4. EEG experimental paradigm and data acquisition

In each session, the 4 vowel stimuli were each presented to the participant 240 times. Subject completed 4 sessions, each comprising 24 experimental runs. Each run consisted of the presentation of 10 instances of each vowel, and 5 instances of a rare noise target. Subject was instructed to respond only to the noise targets within 1.5 s from the stimulus onset by pressing the space bar on a standard keyboard using the index finger of their preferred hand. The target noise sound was synthesized such that it had a flat spectrum, to minimize the chance that our results are influenced by phonetic distance between a given stimulus and the target (see e.g., [16]). The stimuli were played in pseudorandom order such that successive stimuli were always different. The interstimulus interval (ISI) was randomly jittered between 2.1–2.6 s. Stimuli were presented monaurally to the right ear. EEG data were acquired at a sampling rate of 2048 Hz using a 32-channel ActiveTwo EEG system (BioSemi B. V., Amsterdam, The Netherlands). Two flat-type electrodes were attached over the left and right mastoids for offline re-referencing. Stimulus presentation was controlled via a custom MATLAB program developed using the auditory presentation functionalities in Psychtoolbox [17].

### 2.5. Data analysis

#### 2.5.1. Pre-processing and EEG data classification

The raw EEG data were first downsampled from 2048 Hz to 64 Hz using EEGLAB [18], and segmented into runs. Each run was then filtered between 0.5–8 Hz using a sixth order Butterworth filter. The function `filtfilt` in MATLAB was used to ensure that the FIR filter created zero phase shift. For each run, epochs of duration 1.6 s were obtained for each stimulus presented. Each epoch started at 100 ms pre-stimulus, and lasted until 1500 ms post-stimulus. Within each epoch, the time-series associated with the two mastoid electrodes were averaged, and subtracted from that associated with the 32 main electrodes. Linear detrending was performed on each epoch. Epochs in which the potential recorded at any electrode exceeded  $100 \mu\text{V}$  between 0.1 s pre-stimulus and 0.4 s post-stimulus were rejected. Four-fold cross-validation was performed, using data from three sessions as the training data and the data from the remaining session as the testing data. For each trial, a feature vector  $\mathbf{x}$  corresponding to all time-samples in a specified analysis window in all 32 electrodes was constructed. Two parameters were varied—the starting time (0–250 ms) and duration of the analysis window (0–125 ms). For each pair of vowels (e.g., A and B), Fisher’s linear discriminant analysis (see e.g., [19]) was used to train a binary classifier, denoted as  $\mathbf{w}$ , using single-trials as the training data. The classifier was used to convert a feature vector  $\mathbf{x}$  in the testing data into a classifier score according to  $\text{score} = \mathbf{w} \cdot \mathbf{x}$  such that the score tends to be larger for vowel A. The classifier scores were then summed across  $N$  successive trials of the same type. Trials recorded within the same run were always analyzed together in that the vowel with the higher (lower) score was determined to be of vowel A (B).

#### 2.5.2. Analysis of behavioral and neural confusion matrices

To assess how well the perceptual distances among the vowels were reflected in our classification results, discriminability indices ( $d'$ ) were computed using both the behavioral results obtained during the screening test and the classification results. Spearman’s rank-order correlation coefficient was used to quantify the correlation between the two sets of  $d'$  scores.

### 3. Results

Figure 2 shows the mean classification rate as a function of the starting time and size of analysis window across all 6 vowel-pairs and 6 subjects. For sufficiently large window size (e.g.,  $> 40$  ms), we found that the classification rate quickly raised above chance-level as the starting time of the window increased from zero (i.e. stimulus onset), reaching a peak at about 150 ms post-stimulus. The optimal window of analysis was determined to be 78 ms (corresponding to 5 time-samples after the EEG data were down-sampled to 64 Hz).

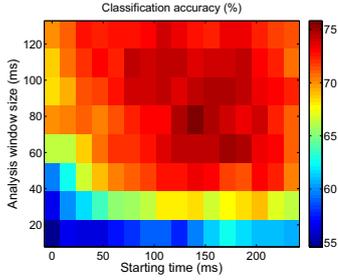


Figure 2: Classification accuracy is plotted as a function of starting time (ms) and size of the analysis window (ms).

Figure 2 suggests that information about vowel identity was contained throughout the duration of the vowel. This in turn suggests that an improvement in accuracy might be obtained by integrating the classifier scores in successive analysis windows. Figure 3 shows the classification rate as a function of the starting time and size of integration window, for the average rate across all 6 vowel-pairs and for every vowel-pair. The best classification rate (77%) was attained when the starting time corresponded to the stimulus onset and the integration window was such that the scores from the first 7 analysis windows were summed.

Table 1(a) shows the mean classification accuracy for the six vowel-pairs obtained using the best parameters described above. Table 1(b) shows the corresponding matrix of  $d'$  scores. Table 1(c) shows the  $d'$  matrix computed based on behavioral data. It is clear that the pair that was the most (least) discriminable behaviorally was also the most (least) classifiable neurally. Figure 4 shows that the two sets of discriminability indices were significantly correlated, according to both the Spearman's rank-order correlation coefficient ( $\rho(34) = 0.54, p < 10^{-3}$ ) and Pearson's correlation ( $r(34) = 0.49, p = 0.003$ ). We further examined the correlation as a function of  $N$  (figure 5). It was observed that as the classification accuracy increases, correlation between the two set of  $d'$  scores also tends to increase.

(a)

	ɔ	ɣ	œ	ɛ
ɔ	—	72	77	86
ɣ	—	—	66	83
œ	—	—	—	76
ɛ	—	—	—	—

(b)

	ɔ	ɣ	œ	ɛ
ɔ	—	1.16	1.54	2.33
ɣ	—	—	0.87	2.04
œ	—	—	—	1.49
ɛ	—	—	—	—

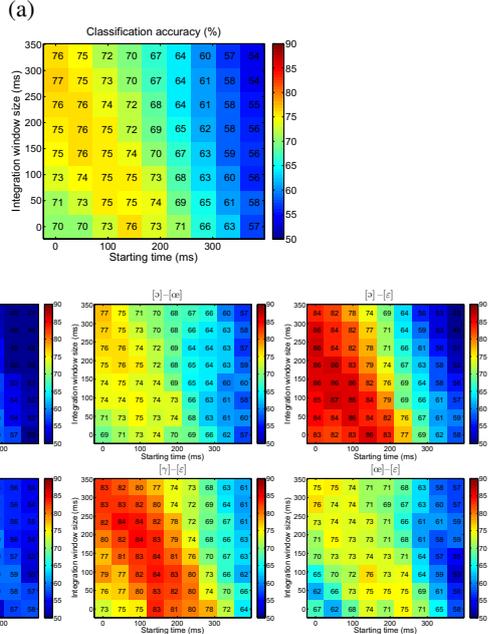


Figure 3: (a) Mean classification rate over all subjects and all 6 vowel-pairs. (b) Mean classification rate for each vowel-pair.

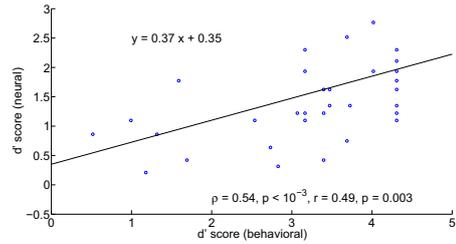


Figure 4: Behavioral performance is significantly correlated with classification performance.

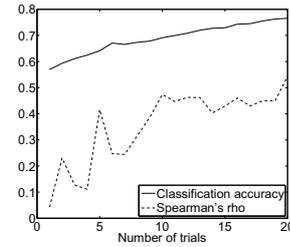


Figure 5: Spearman's rho for the correlation between behavioral and neural  $d'$  scores tends to increase with increasing classification accuracy (scaled to 0-1).

Table 1: (a) Binary classification rate for each of the six vowel-pairs arranged in the form of an upper triangular matrix; (b) Matrix of discriminability indices  $d'$  computed based on classification results; (c) Matrix of discriminability indices  $d'$  computed based on behavioral results.

## 4. Discussion

In this study, we used a machine-learning framework to extract information about vowel identity from EEG data. The best classification rate was achieved when the size of analysis window for extracting features was set to about 80 ms. Using the behavioral data in the screening test to estimate the discriminability indices of different vowel-pairs, we found that discriminability indices calculated based on behavioral data and classification results were significantly correlated. In particular, the pair that was neurally the most discriminable, [ɔ]–[ɛ] (mean accuracy: 86%), and the pair that was the least, [ɣ]–[α] (mean accuracy: 66%), were also the most and the least discriminable behaviorally, respectively.

### 4.1. Implications to the study of vowel perception

First, our results indicate that at the global level, the cortical representation of vowels follows the organizational principle that vowels that are perceptually more dissimilar are represented more dissimilarly at the cortical level. While psychoacoustic theory can be applied to estimate the perceptual similarity between any pair of sounds by modelling the representation of each sound as an excitation pattern at the level of basilar membrane, it is also clear that such a model is inadequate to accurately reflect perceptual distance, which is significantly influenced by language experience. The present study offers a direct method for probing the perceptual distances between vowels that are indicative of cortical processes.

Second, above-chance-level classification performances were achieved throughout the stimulus interval (figure 2), when the window size was sufficiently large (e.g., > 40 ms). This suggests that vowel-dependent information is distributed across the whole stimulus-presentation interval. Also, the amount of information reaches a peak at about 140 ms post-stimulus. In the local field potential study by Chang et al. [11], the time-interval between 110–150 ms was found to be the most informative for discriminating the three stop consonants. Taken together, these results suggest that the N1–P2 complex in the auditory evoked responses likely contain information that can be used to classify phonemes.

Third, there have previously been only two widely-used neurophysiological methods for measuring perceptual distance—source localization using magnetoencephalography (MEG), and the use of mismatch negativity (MMN [20]), using both EEG and MEG. Regarding source localization, the focus was placed on one event-related magnetic field (ERF) component—the N1m component. While previous studies [21, 22, 23, 24, 25] have demonstrated that the acoustic distances between vowel-pairs and the relative distances in the N1m source location between the vowel-pairs follow the same ordering, suggesting that vowels are organized in an orderly fashion cortically, it is unclear whether the relative distance between source reflects acoustic differences or phonological differences. Similar problems may also exist regarding MMN studies. Many MMN studies using vowels of a single language generally support the notion that MMN magnitude primarily reflects the differences in formant structures in both active [26, 16] and passive [26, 27, 16, 28] listening conditions—the larger the differences in formant structures, the larger the MMN amplitude. Thus, according to these studies, their results suggest that MMN could serve as a metric of perceptual distance that is based on formant parameters. On the other hand, other studies find that MMN magnitude is influenced by whether the deviant is a native or non-native vowel (e.g., [29]).

Specifically, a non-native vowel deviant elicits smaller MMN magnitude. However, since the vowel deviant is not confused with the standard in this study, the perceptual distance, as defined behaviorally, should be similar to other native stimuli instead of smaller. This inconsistency suggests that MMN magnitude does not always reflect perceptual distance.

Fourth, in the study by Wang et al. [12], EEG responses to the four English vowels, /i/, /æ/, /u/ and /ɑ/ were classified using a multi-class, hierarchical model with accuracy up to 90.0% using 25 trials. Several other factors might have contributed to the higher accuracy in this study compared to ours. (a) Wang et al. recorded 1,792 trials per vowel compared to 960 in our study, thus the amount of training data available was about doubled in each analysis. (b) Only 1 subject participated in their vowel classification experiment, thus as far as accuracy is concerned, there is no basis to determine the generalizability of their results. (c) The four English vowels varied in both F1 and F2, and were more separated in the vowel space compared to our set of vowels, which differ only in F2. Also, natural vowels produced by a male speaker were used in their case, while synthetic vowels were used in ours. Thus it is likely that their stimuli contain greater amount acoustic cues for vowel discrimination. This might also have contributed to their better classification performance. Given these differences, the present study is not directly comparable with Wang et al.'s study. Further experiments are required to determine the factors influencing vowel classification performance.

### 4.2. Future work

Although our results demonstrated a significant brain–behavior correspondence, in that behavioral  $d'$  scores are correlated with neural  $d'$  scores computed based on classification accuracy, there is likely plenty of room to improve the robustness of this result. First, we observed a trend for the brain–behavior correlation to improve as classification rate increases (figure 5). Thus, we anticipate that the use of a more sophisticated classification method might strengthen the brain–behavior correspondence. Thus, our future work includes testing different classification algorithms, such as support vector machine (SVM), which is also adopted in [12], to improve upon the current accuracy achieved.

Second, our behavioral results were collected prior to the main EEG experiment. Because of the nature of the noise-detection task adopted, it is unknown what the subject actually perceived, which could vary from trial to trial. One possibility is to use a delayed oral report task during the EEG experiment, in which the subject would have to orally repeat the stimulus in every trial. By recording the behavioral and EEG data together, it is possible that the correlations found in the present study could be further strengthened.

In summary, we used classification analysis to recognize the EEG responses to vowels. By comparing behavioral discriminability indices with those derived from classification, we found results that support the intuitive idea that vowels that are behaviorally more distinct evoke brain activities that are more distinct. The work presented here demonstrate the viability of using EEG data classifications to derive a neural measure of perceptual distance.

## 5. Acknowledgements

This work was supported in part by a research grant awarded to William S-Y. Wang by the Patent Committee of The Chinese University of Hong Kong (Grant No.: TBF/12/ENG/019).

## 6. References

- [1] Delattre, P., Liberman, A. M., Cooper, F. S. and Gerstman, L. J., "An experimental study of the acoustic determinants of vowel color; observations on one- and two-formant vowels synthesized from spectrographic patterns", *Word*, 8:195-210, 1952.
- [2] Peterson, G. E. and Barney, H. L., "Control methods used in a study of the vowels", *The Journal of the Acoustical Society of America*, 24(2):175-184, 1952.
- [3] Rosner, B. S. and Pickering, J. B., "Vowel perception and production", Oxford University Press, 1994.
- [4] Kuhl, P. K., "Human adults and human infants show a 'perceptual magnet effect' for the prototypes of speech categories, monkeys do not", *Perception & psychophysics*, 50(2):93-107, 1991.
- [5] Kuhl, P. K., Williams, K. A., Lacerda, F., Stevens, K. N. and Lindblom, B., "Linguistic experience alters phonetic perception in infants by 6 months of age", *Science*, 255(5044):606-608, 1992.
- [6] Macmillan, N. A., Kaplan, H. L. and Creelman, C. D., "The psychophysics of categorical perception", *Psychological Review*, 84(5):452, 1977.
- [7] Macmillan, N. A. and Creelman, C. D., "Detection theory: A user's guide", Psychology press, 2004.
- [8] Mohr, B. and Wang, W. S-Y., "Perceptual distance and the specification of phonological features", *Phonetica*, 18(1):31-45, 1968.
- [9] Kewley-Port, D. and Atal, B. S., "Perceptual differences between vowels located in a limited phonetic space", *The Journal of the Acoustical Society of America*, 85(4):1726-1740, 1989.
- [10] Iverson, P. and Kuhl, P. K., "Mapping the perceptual magnet effect for speech using signal detection theory and multidimensional scaling", *Journal of the Acoustical Society of America*, 97(1):553-562, 1995.
- [11] Chang, E. F., Rieger, J. W., Johnson, K., Berger, M. S., Barbaro, N. M. and Knight, R. T., "Categorical speech representation in human superior temporal gyrus", *Nature neuroscience*, 13(11):1428-1432, 2010.
- [12] Wang, R., Perreau-Guimaraes, M., Carvalhaes, C. and Suppes, P., "Using phase to recognize English phonemes and their distinctive features in the brain", *Proceedings of the National Academy of Sciences*, 109(50):20685-20690, 1983.
- [13] Jakobson, R., Fant, G. and Halle, M., "Preliminaries to speech analysis. The distinctive features and their correlates", M.I.T. Press, 1967.
- [14] Oldfield, R. C., "The assessment and analysis of handedness: the Edinburgh inventory", *Neuropsychologia*, 9(1):97-113, 1971.
- [15] Hose, B., Langner, G. and Scheich, H., "Linear phoneme boundaries for German synthetic two-formant vowels", *Hearing research*, 9(1):13-25, 1983.
- [16] Deguchi, C., Chobert, J., Brunelli re, A., Nguyen, N., Colombo, L. and Besson, M., "Pre-attentive and attentive processing of French vowels", *Brain research*, 1366:149-161, 2010.
- [17] Borgo, M., Soranzo, A. and Grassi, M., "Psychtoolbox: Sound, Keyboard and Mouse", In *MATLAB for Psychologists*, 249-273, Springer New York., 2012.
- [18] Delorme, A. and Makeig, S., "EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis", *Journal of neuroscience methods*, 134(1):9-21, 2004.
- [19] Duda, R. O., Hart, P. E. and Stork, D. G., "Pattern classification", John Wiley & Sons, 2012.
- [20] N at nen, R., Paavilainen, P., Rinne, T. and Alho, K., "The mismatch negativity (MMN) in basic research of central auditory processing: a review", *Clinical Neurophysiology*, 118(12):2544-2590, 2007.
- [21] Diesch, E., Eulitz, C., Hampson, S. and Ross, B., "The neurotopography of vowels as mirrored by evoked magnetic field measurements", *Brain and language*, 53(2):143-168, 1996.
- [22] M kel , A. M., Alku, P. and Tiitinen, H., "The auditory N1m reveals the left-hemispheric representation of vowel identity in humans", *Neuroscience letters*, 353(2):111-114, 2003.
- [23] Obleser, J., Elbert, T., Lahiri, A. and Eulitz, C., "Cortical representation of vowels reflects acoustic dissimilarity determined by formant frequencies", *Cognitive Brain Research*, 15(3):207-213, 2003.
- [24] Eulitz, C., Obleser, J. and Lahiri, A., "Intra-subject replication of brain magnetic activity during the processing of speech sounds", *Cognitive brain research*, 19(1):82-91, 2004.
- [25] Shestakova, A., Brattico, E., Soloviev, A., Klucharev, V. and Huotilainen, M., "Orderly cortical representation of vowel categories presented by multiple exemplars", *Cognitive Brain Research*, 21(3):342-350, 2004.
- [26] Aaltonen, O., Niemi, P., Nyrke, T. and Tuhkanen, M., "Event-related brain potentials and the perception of a phonetic continuum", *Biological Psychology*, 24(3):197-207, 1987.
- [27] Savela, J., Kujala, T., Tuomainen, J., Ek, M., Aaltonen, O. and N at nen, R., "The mismatch negativity and reaction time as indices of the perceptual distance between the corresponding vowels of two related languages", *Cognitive brain research*, 16(2):250-256, 2003.
- [28] Tuomainen, J., Savela, J., Obleser, J. and Aaltonen, O., "Attention modulates the use of spectral attributes in vowel discrimination: Behavioral and event-related potential evidence", *Brain research*, 1490:170-183, 2013.
- [29] N at neiv, R., Lehtokoski, A., Lennest, M., Luuki, A., Alliki, J., Sinkkonen, J. and Alho, K., "Language-specific phoneme representations revealed by electric and magnetic brain responses", *Nature*, 385:432-434, 1997.
- [30] Winkler, I., Kujala, T., Tiitinen, H., Sivonen, P., Alku, P., Lehtokoski, A., Czigler, I., Cspe, V., Ilmoniemi, R. J. and N at nen, R., "Brain responses reveal the learning of foreign language phonemes", *Psychophysiology*, 36(5):638-642, 1999.