The effect of increasing acoustic and linguistic complexity on auditory processing: an EEG study

Fareeha S. Rana[1], Daniel Pape[1], & Elisabet Service[1]

[1]McMaster University, Hamilton, ON, Canada

ranafs@mcmaster.ca, paped@mcmaster.ca, eservic@mcmaster.ca

Abstract

This study explores the pre-attentive processing of auditory stimuli using EEG with a systematic (stepwise) increase in linguistic complexity. An oddball paradigm where rare deviants differed in pitch were compared to frequently-presented acoustic stimuli of two types: non-speech complex harmonic waves, and speech syllables. Syllables were generated using VocalTractLab [1] which allows for the synthesis of natural-sounding and highly controllable artificial speech. The complex waves were generated using Praat [2], and their intensity envelopes and harmonic structure were matched to those of the syllables. Thus, the two stimulus types were acoustically very comparable, differing only in the presence versus absence of modelled vocal tract influences and laryngeal control.

We hypothesized that both pitch deviants would evoke a Mismatch Negativity (MMN) brain response. We further explored how MMN size varied based on the nature of the stimulus trains. Results supported our hypothesis. An MMN was observed for both pitch deviants. The MMN was larger in amplitude for the complex waves than the syllables, suggesting differences between how linguistic syllable stimuli are processed compared to non-linguistic complex waves.

This study demonstrates how an increase in acoustic and linguistic complexity reflects in the MMN response and provides support for domain-specific theories of auditory processing.

Index Terms: auditory processing, neurolinguistics, EEG, speech perception, speech acoustics

1. Introduction

The experiment reported here explores differences in the auditory processing of acoustically highly-controlled complex waves and syllable stimuli. The aim of this experiment was to understand how the presence of linguistic content affects auditory processing when all other acoustic parameters have been controlled for.

The research in speech processing and perception makes use of speech tokens that are naturally recorded [e.g., 3], semi-synthesized [e.g., 4], or artificially produced with synthesizers [e.g., 5]. However, artificial speech synthesizers often produce distorted speech signals that are difficult to understand and a rather poor correlate of natural speech, leading to researchers increasingly opting for the natural and semi-synthesized options. There are pros and cons to both the more natural and the synthesized approaches. Notably, with the first two, speech is most natural, but various speech parameters are difficult to control for due to all the individual variation that inevitably exists within speakers. In speech synthesis, the advantage is being able to control all relevant acoustic/articulatory parameters of synthesis tightly and accurately. The disadvantage is a significant loss of perceptual speech quality. In the current experiment, we made use of a speech synthesis approach and program that generates speech using highly realistic vocal tract and laryngeal configuration modelling, namely articulatory synthesis using the software Vocaltractlab [1]. This synthesizer allows for fine and accurate control of all relevant speech signal parameters, while generating a syllable that has the same characteristics as a naturally spoken one. Considering that this syllable is completely artificially synthesized, although highly realistic (e.g., compared to a traditional Klatt approach), the question of how it would be processed arises. While natural to artificial speech comparisons have been made before [6], to our knowledge, little or no work has made use of a synthesizer that mimics vocal tract and laryngeal characteristics to the degree of detail studied here. VocalTractLab was used to generate syllables that sound highly natural despite being fully artificial and articulatorily parametrizable. Our research question was: would the artificial syllable be processed similarly to speech, or would processing be closer to that of other types of acoustic signals we are frequently exposed to (i.e., non-linguistic complex waves)?

Theoretical frameworks of auditory processing argue that speech is either processed the same way as all other auditory input (domain-general), or that it is processed via specialized processing pathways that are specific to speech only (domain-specific) [7]. While there are theories that incorporate elements of both [cf. 8], our aim in this experiment is to understand how similar the processing of linguistic stimuli is to acoustically similar non-linguistic stimuli.

We evaluated differences in the auditory processing of the two types of stimuli through changes in a neural response known as the mismatch negativity (MMN). The MMN is a negative-polarity event-related potential, often evoked in the absence of attention, when the pattern of a presented stimulus stream is broken [9, 10]. The MMN relies on the formation of an expectancy based on the repeating standard stimulus; the violation of this expectancy elicits the MMN. The degree of violation—depending on many factors, including saliency, timing, and nature of the incoming stimulus—determines its amplitude (size) and latency (timing). Differences in MMN elicitation, therefore, will allow us to infer differences in stimulus processing. The results also have implications for future MMN studies, providing researchers with a benchmark of how this response behaves between tightly controlled sets of non-speech and speech stimuli.

Previous literature exploring MMN elicitation for complex waves and syllables has found mixed results. In one study, speech deviants within a stream of non-speech sounds elicited an MMN. However, in a reverse paradigm, non-speech deviants among frequent speech sounds did not, suggesting a pre-attentive saliency for speech [11]. Other studies have shown comparable results, with speech eliciting larger MMN responses compared to non-speech (e.g., complex waves, pseudowords) [12, 13]. On the other hand, in comparisons with simple tones, speech has been found to elicit smaller responses [14, 15], and complex waves have been found to elicit larger ones [16]. As the previous research confounds the acoustic details of stimuli with linguistic character, the findings do not allow us to conclude that increased linguistic or acoustic complexity necessarily elicits larger-amplitude MMNs.

2. Hypotheses

The present experiment was exploratory in nature. We did expect an MMN to the frequency (pitch) deviant for both the complex wave stimuli and the syllable stimuli. However, given the nature of the contrast between them—both sets of stimuli being entirely generated and artificially synthesized, but with vocal tract characteristics and non-harmonic structure applied only to one—it was difficult to predict how, if at all, their processing would differ.

There was a possibility that the elicited MMNs would be similar in amplitude and latency for both types of stimuli. Such an outcome would suggest that both complex waves and artificially synthesised syllables are processed similarly, lending credence to the domain-general model of processing which predicts no special auditory processing pathway for speech.

Alternatively, the MMN elicited may differ in size (or latency) between the two stimulus types. A larger MMN amplitude elicited by the complex waves might suggest better pattern-forming and expectancy-building for the non-linguistic, purely acoustic, stimuli compared to the artificial speech syllables. A bigger MMN for the syllables could, however, suggest that an intrinsic importance is assigned to the processing of linguistic (speech) stimuli over that for other auditory stimuli, leading to a stronger violation response. In some cases, this might be accompanied by a positive response known as the P300, signalling that the stimulus triggered attention during its processing [17].

3. Methods

1.1. Participants

Twenty Canadian university students (mean age = 21.8; s.d. = 3.1; 15 females) with no reported visual/auditory problems were recruited to participate in this experiment. Data from two participants was rejected from further analyses due to excessive artifacts and noise, and partial data was used from one.

1.2. Stimuli

The oddball paradigm consists of a series of standard stimuli interspersed with a randomly presented deviant that differs from them on one or more acoustic features, eliciting the MMN response. In this experiment, the paradigm was conducted with two types of stimuli: complex waves, and syllables. Complex wave and syllable stimuli were kept as acoustically similar as possible, with differences only in the presence of linguistic content within the syllables.

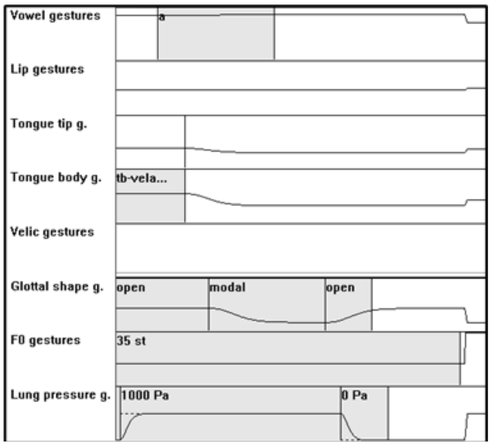

Figure 1: *Example of the gestural score for the synthesized syllable /ka/ from VocalTractLab [1].*

1.2.1. Syllables

In our experiment, we used three different syllables that were entirely synthesized (as opposed to natural speech recordings, or semi-synthesized) and presented them in an auditory stream intended to emulate natural speech. Four CV syllables were generated for this experiment using plosives /t k/ and vowels /i u ə a/ to form the syllables /ti/, /tu/, /tə/ and /ka/ (**Figure 1**). The voiceless plosives /t k/ were chosen for the deviant onset contrast because they are amongst the most frequently occurring speech sounds in the world [18], in principle allowing us to test with a participant pool that was not limited to native Canadian English speakers. Syllables were synthesized using VocalTractLab (version 2.1) [1], a software that allows highly controllable manipulation of vocal tract and laryngeal parameters and configurations to generate natural-sounding artificial speech using a state-of-the-art articulatory synthesis approach. The following parameters were used to generate these syllables:

- Tongue tip (alveolar stop) or tongue body (velar stop) configuration for 0 - 150 ms, Time Constant (tc): 15.2 ms.
- Vowel gesture (u, i, @, a) from 100 – 350 ms, tc: 15.0 ms.
- Glottal shape open (0 – 152 ms, tc = 15.1 ms), modal (152 – 402 ms, tc = 15.1 ms) and open (402 – 100 ms, tc = 15.0 ms).

These parameters were set to ensure a plosive length of at least 50 ms, measured from the end of the stop burst to the beginning of the following vowel. The synthesized syllables were exported into Praat and truncated to be 150 ms in length (by adjusting the overall vowel length). The average consonant length was 49.5 ms, and the average vowel length was 95.5 ms. A 5 ms rise-time was introduced at the beginning, after which the stop burst occurs. Following time normalization, the stimuli were loudness-normalized in Sound Forge Pro (v.3.0) and linear fade-in/fade-out effects applied for the first and last 5 ms. The sound pressure level was adjusted for all participants to 70 dB SPL_C (except for one, where it was further lowered to comfort). The first three syllables /ti, tu, tə/ had an f0 of 100 Hz and were presented as standards; the pitch deviant /ka/ had a f0 of 120 Hz.

1.2.2. Complex waves

A base complex acoustic signal was constructed by taking an initial (f0) frequency component of 100 Hz, and then adding additional sine waves (i.e., harmonics) in frequency steps of 100 Hz up until 8000 Hz with a spectral slope of -6 dB/octave. Thus, we effectively generated a complex waveform with a f0 of 100 Hz, and a large number of harmonics decreasing in intensity in a way comparable to the harmonics in a standard speech sound. This signal was generated using Praat [2] and the complex waves with an f0 of 100 Hz were used as the standard stimuli.

The deviant waveform differed in the base (f0) frequency. An initial frequency of 120 Hz was used instead of 100 Hz. Adjustment of the base frequency resulted also in differing harmonic frequencies; all other features were identical between standards and deviants.

The generated complex waves were processed to acoustically match the syllables (150 ms duration, 5 ms fade-in and fade-out), and loudness-normalized using identical procedures. The sound pressure level was adjusted for all participants to 70 dB SPL_C (except for two, where it was adjusted lower for comfort).

Intensity envelopes of three different speech syllables used as the standards, /ti tu tə/, and for the one deviant, /ka/, were used to transfer the amplitude-over-time information onto the generated complex waveform. This resulted in four different complex waveforms with differing amplitude envelopes, each with a specific syllabic speech amplitude distribution but no linguistic content. The main difference between the complex waves and syllables, therefore, was only in the presence of modelled vocal tract influences and laryngeal control for the syllables (e.g., voiceless VOT phase with transient and noise elements driven by articulator interactions).

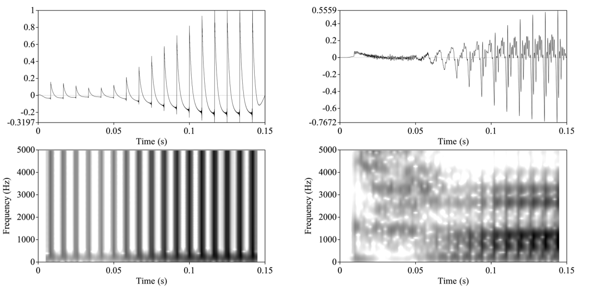

Figure 2: *Oscillogram and spectrogram for the generated complex wave (left) containing amplitude-over-time information from the syllable /ka/ (right).*

1.3. Experimental and testing procedure

An auditory oddball paradigm was used with three standards and one deviant in each block. Complex waves were presented to participants in a separate block from syllables. A total of 900 standards (90%) and 100 deviants (10%) were presented in each block with a standard Stimulus Onset Asynchrony (SOA) of 650 ms. To mimic the variation present in naturally spoken speech, three different tokens were presented equiprobably for the standards (300 x 3 types of tokens for a total of 900 standards). Only one deviant was presented (either the syllable /ka/, or its complex wave equivalent).

Participants provided informed consent at the beginning of the testing session, and answered some demographic questions (e.g., age, handedness, languages spoken, any concussion history etc.). The EEG experiment followed, with participants watching a silent film while ignoring the sounds presented to them. As is standard in an oddball paradigm, no responses were required of the participants. The stimuli were presented to them via ER-1 Insert earphones [19] using the software Presentation [20]. Participants completed two blocks, one of each stimulus type. Each block was approximately 10 minutes long.

1.4. EEG recording and data processing

EEG data were pre-processed and cleaned using BrainVision Analyzer (v2.1.2.327). Data were re-filtered to 0.1-30Hz (24 dB/oct), and up to four noisy channels were statistically interpolated and replaced where needed [21]. Sections of data with artifacts greater than 100 μV were removed, as were responses to trials with vertical and horizontal eye movements using an Ocular Independent Component Analysis (ICA).

The data were then segmented into chunks of -100 ms to 600 ms. Difference waveforms for MMN and P300 strength were produced by subtracting the averaged waveform in the standard condition from that in the deviant condition. Automated peak detection [22] was conducted to find the maximum negative amplitude pertaining to the MMN within a window of 100 ms to 300 ms, and the positive maximum for the P300 within a window of 250 ms to 500 ms [17]. As the MMN is a fronto-central component, electrodes grouped into relevant regions of interest (ROI) were chosen for analyses [9, 10]. One-tailed *t*-tests and paired *t*-tests were conducted to evaluate MMN and P300 mean peak amplitude and latency values within and between the two stimulus types using the R Studio [23] environment for R [24].

Table 1: *Means, standard deviations and significance (t test) results for the mean peak amplitude (AMP) and mean peak latency (LAT) for MMN and P300.*

			n	*Mean peak value*	*t (df)*
MMN	AMP	Complex	18	-3.27 ± 1.67	-9.30 (17) *
		Syllables	17	-2.14 ± 1.71	-5.34 (16) *
	LAT	Complex	18	211 ± 39.9	1.26 (17)
		Syllables	17	213 ± 52.6	1.25 (16)
P300	AMP	Complex	18	2.18 ± 2.04	4.78 (17) *
		Syllables	17	2.20 ± 1.40	7.09 (16) *
	LAT	Complex	18	382 ± 79.6	0.483 (17)
		Syllables	17	373 ± 76.7	-0.153 (16)

* Indicates significance $p < 0.05$ — Average ± s.d

4. Results

An MMN subtraction wave of varying amplitude was elicited for deviants of both stimulus types (**Figure 3**). Mean peak amplitude and mean peak latency values were extracted for each stimulus type within the time window 100 ms – 300 ms. One tailed *t*-tests were used to evaluate the difference of the peak MMN amplitude from 0 μV, and of the peak MMN latency from the midpoint of the time window (200 ms) for both stimulus types (**Table 1**). Both complex waves and syllables elicited an MMN difference wave between deviant and standard responses that was significantly negative in amplitude compared to baseline ($p<0.05$). There were no latency differences for either the complex waves or syllables ($p>0.05$).

A positivity, the P300 component, was also observed, and one tailed *t*-tests for the deviant minus standard difference waves were used to evaluate mean peak amplitude (from 0 μV) and mean peak latency (375 ms; time window: 250 ms – 500 ms) for both stimulus types. A significantly positive mean peak amplitude was observed for both complex waves and syllables ($p<0.05$). There were no significant latency differences for either stimulus type for the P300 ($p>0.05$).

A *z*-normalization to control for individual differences in skull and scalp conductivity was performed on the raw data points for each participant, and paired t-tests were conducted to evaluate the difference in mean peak amplitude and mean peak latency comparing the two types of stimuli. Results showed a significantly larger MMN for the complex wave deviant than for the syllable deviant ($t(16)=2.46$, $p<0.05$). There was no difference between the P300 amplitudes between stimulus types ($t(16)=0.545$, $p>0.05$). Furthermore, no difference in peak latency was observed between the two stimulus types for either the MMN ($t(16)=0.426$, p>0.05) or the P300 ($t(16)=-0.124$, $p>0.05$).

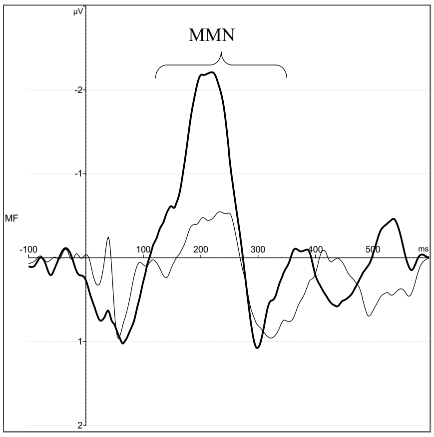

Figure 3: *Difference wave showing the response to the pitch deviant for complex waves (bold) and syllables (thin). (Middle-Frontal ROI, negative amplitude plotted upwards. Image is filtered at 1-25 Hz)*

5. Discussion and Conclusion

This experiment was mostly exploratory. Although we did expect an MMN to be elicited for both types of stimuli, whether the MMN would differ between stimulus types was our main question. In our experiment, a MMN was observed for both the complex waves frequency deviant and the syllables frequency deviant. A larger size MMN was elicited by the complex waves deviant compared to the syllable deviant. Both stimulus types elicited a P300, as well, suggesting an engagement of attention during stimulus presentation for both complex waves and syllables.

The larger response to the complex wave pitch deviant suggests that it was more salient during presentation compared to the syllable deviant. Considering both sets of stimuli were controlled so that the only acoustic difference between them was of that of linguistic content, we can assume that the difference in response was a direct result of the introduction of linguistic complexity or, at least, introduction of articulatory information and articulator interaction. Therefore, we find that the presence of this information elicits a shallower response, suggesting that it is processed differently.

This result is different from those observed previously in the literature, where speech stimuli has elicited larger responses compared to non-speech stimuli. While the MMN has been extensively used to study phoneme learning and discrimination [25-27], CV-syllables have been found to elicit little to no MMN in some studies [e.g., 28]. The authors suggested several reasons for this, including specialized processing of speech over non-speech. In our experiment, both the complex waves and syllables elicit a clear MMN; the difference lies only in the size of the response, which was larger for complex waves than syllables. This result seems puzzling at first. The deviant for the syllable differed along three features (consonant, vowel, and f0), while the complex wave deviant differed only on pitch. This should have led to an increased perceptual saliency eliciting a larger response, or at the very least, an equal response; this was not the case.

One possible reason for this may be the experimental design, which aimed to replicate speech variation by presenting three different tokens for each stimulus type. If syllables were perceived as part of a potential speech stream, it may have led to a wider 'acceptability' window for the pattern violations caused by the deviant, as the already existing phonemic differences among the syllables would confound pattern formation. This would have led to a slower build-up of onset expectancy and ergo a higher threshold for violation detection, meaning it would take longer for an MMN to emerge. The ratio of unique standards to deviants (300:100) may also have contributed to a weaker expectancy formation [29]. However, this multi-token design was presented for complex waves as well, which showed no MMN size attenuation as would be expected had the design been a contributing factor.

Thus, it is likely, that the difference in the elicited MMN response lies purely in the linguistic nature of the stimulus. The complex waves, having been processed via a generalized auditory processing route that has been classically shown for many MMN experiments, elicited a larger MMN. The syllables, on the other hand, may have engaged a more language-specific pathway, sensitive to the linguistic familiarity of the stimuli. This result provides evidence for the domain-specific theory of auditory processing, suggesting that language stimuli intrinsically are processed in a manner that is different from purely auditory/acoustic stimuli, when all else remains the same (both generated or synthesized artificially, loudness-normalized, with normalized intensity envelopes, durations and fade times).

The results of this experiment, therefore, suggest that there is a difference in the way complex waves and syllables are processed, with a language-specific pathway used to process the linguistic syllable stimuli as opposed to the purely acoustic, non-linguistic complex waves. This has implications for future MMN studies that may choose to use synthesized speech instead of natural speech, as the results show how MMNs compare between complex waves and artificially synthesized syllables.

6. Acknowledgements

This work was supported by the ARiEAL Research Supplement Fund (first author) and NSERC grant RGPIN-2018-06518 (second author).

7. References

[1] P. Birkholz, “Perceptual Optimization of an Enhanced Geometric Vocal Fold Model for Articulatory Speech Synthesis,” in *Proceedings INTERSPEECH 2019,* pp. 3765-3769, 2019.

[2] P. Boersma, and Dr. Weenink, “Praat: Doing Phonetics by Computer (6.0.18) [Computer Software].” 2019.

[3] L. Plug, and R. Smith. “Segments, syllables, and speech tempo perception.” *Proceedings of the 9th International Conference on Speech Prosody*, pp. 279-283, June, 2018.

[4] A. Sorokin, P. Alku, and T. Kujala, “Change and novelty detection in speech and non-speech sound streams.” *Brain research,* vol. 1327, pp. 77-90, 2010.

[5] L. Sanders, E. L. Newport, and H. J. Neville, “Segmenting nonsense: an event-related potential index of perceived onsets in continuous speech.” *Nature neuroscience*, vol.5, no.7, pp. 700-703, 2002.

[6] A. Brown, “Neural coding of natural and synthetic speech,” M.S. thesis, Department of Ot. Head and Neck Sur. And Comm. Disorders, Univ. of Louisville, Louisville, Kentucky, 2017. Accessed on: Mar. 13, 2022. [Online]. Available: https://ir.library.louisville.edu/cgi/viewcontent.cgi?article=3833&context=etd

[7] R. Zatorre, and J. T. Gandour, “Neural specializations for speech and pitch: moving beyond the dichotomies,” *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 363, no. 1493, pp. 1087-1104, 2008.

[8] D. Poeppel, “The analysis of speech in different temporal integration windows: cerebral lateralization as ‘asymmetric sampling in time’”, *Speech communication*, vol. 41, no.1, pp. 245-255, 2003.

[9] R. Näätanen, W. Teder, K. Alho, J. Lavikainen, J., “Auditory attention and selective input modulation: a topographical ERP study”, Cognitive Neuroscience and Neuropsychology, vol. 3, no. 6, pp. 493–496, 1992.

[10] R. Näätanen, T. Kujala, and G. Light, Mismatch Negativity: A Window to the *Brain*. Oxford University Press, 2019.

[11] C. M. Pettigrew, B. M. Murdoch, J. Kei, H. J. Chenery, R. Sockalingam, C. W. Ponton, S. Finnigan, and P. Alku, “Processing of English words with fine acoustic contrasts and simple tones: A mismatch negativity study,” *Journal of the American Academy of Audiology*, vol. 15, no. 1, pp. 47–66. 2004

[12] P. Korpilahti, C. M. Krause, I. Holopainen, and A. H. Lang, “Early and Late Mismatch Negativity Elicited by Words and Speech-Like Stimuli in Children. Brain and Language”, *Brain and Language,* vol. 76, no. 3, pp. 332–339, 2001.

[13] C. M. Pettigrew, B. M. Murdoch, J. Kei, C. W. Ponton, P. Alku, and H. J. Chenery, “The mismatch negativity (MMN) response to complex tones’ and spoken words in individuals with aphasia,” *Aphasiology*, vol. 19, no. 2, pp. 131–163, 2005.

[14] C. Sebastian, and I. Yasin, “Speech versus tone processing in compensated dyslexia: Discrimination and lateralization with a dichotic mismatch negativity (MMN) paradigm,” *International Journal of Psychophysiology*, vol. 70, pp. 115–126, 2008.

[15] T. Ilvonen, T. Kujala, H. Kozou, A. Kiesiläinen, O. Salonen, P. Alku, P., and R. Näätänen, “The processing of speech and non-speech sounds in aphasic patients as reflected by the mismatch negativity (MMN)”, *Neuroscience Letters*, vol. 366, pp. 235–240, 2004.

[16] M. Tervaniemi, T. Ilvonen, J. Sinkkonen, A. Kujala, K. Alho, M. Huotilainen, and R. Näätänen, “Harmonic partials facilitate pitch discrimination in humans: electrophysiological and behavioural evidence,” *Neuroscience Letters*, v. 279, pp. 29-32, 2000.

[17] J. Polich, “Updating P300: An integrative theory of P3a and P3b,” *Clinical Neurophysiology*, vol. 118, no. 10, pp. 2128–2148, 2007.

[18] I. Maddieson, and S. F. Disner, *Patterns of Sounds*. Cambridge University Press, 1984.

[19] Etymotic Research, Inc., Elk Grove Village, IL, www.etymotic.com)

[20] Presentation software (v. 18.0), Neurobehavioral Systems, Inc., Berkeley, CA. URL https://www.neurobs.com

[21] V. Duda-Milloy, P. Tavakoli, K. Campbell, D. L. Benoit, and A. Koravand, “A time-efficient multi-deviant paradigm to determine the effects of gap duration on the mismatch negativity,” *Hearing Research*, vol. 377, pp. 34–43, 2019.

[22] R. E. Barr, J. J. Ackmann, and J. Sonnenfeld, J, “Peak-detection algorithm for EEG analysis,” *International Journal of Bio-Medical Computing*, vol. 9, no. 6, pp. 465-476, 1978.

[23] “R Studio: Integrated Development Environment for R [v. 1.4.1106],” RStudio, PBC, Boston, MA, 2021. URL http://www.rstudio.com/.

[24] “R: A language and environment for statistical computing [v. 4.1.0],” R Foundation for Statistical Computing, Vienna, Austria. 2021. URL https://www.R-project.org/

[25] R. Näätänen, “The perception of speech sounds by the human brain as reflected by the mismatch negativity (MMN) and its magnetic equivalent (MMNm),” Psychophysiology, vol. 38, no. 1, pp.1-21, 2001.

[26] I. Winkler, T. Kujala, H. Tiitinen, P. Sivonen, P. Alku, A. Lehtokoski, I. Czigler, V. Csépe, R. J. Ilmoniemi, and R. Näätänen, "Brain responses reveal the learning of foreign language phonemes," Psychophysiology, vol. 36, no. 5, pp. 638-642, 1999.

[27] O. Aaltonen, P. Niemi, T. Nyrke, and M. Tuhkanen, “Event-related brain potentials and the perception of a phonetic continuum,” *Biological Psychology*, vol. 24, no. 3, pp. 197–207, 1987.

[28] J. L. Wunderlich, and B. K. Cone-Wesson, “Effects of stimulus frequency and complexity on the mismatch negativity and other components of the cortical auditory-evoked potential,” *The Journal of the Acoustical Society of America*, vol. 109, no. 4, pp. 1526–1537, 2001.

[29] 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, pp. 2544-2590, 2007.