Gender Priming Pitch Height and Instrument Timbre Perception: Differentiating the N400 ERP Component.

Aditya Chander, Sebastian James

Introduction

The N400 wave is thought to be involved in the processing of many different types of stimuli. written, spoken, and signed (pseudo) words, drawings, photos, and videos of faces, objects and actions, sounds, and mathematical symbols have all be shown to elicit an N400 response. What is curious about N400 is that despite its versatility as being summoned by many different types of stimuli, the latency of the negativity remains remarkably constant. N400 is an event related potential (ERP) component which can be used to analyze almost all forms of language processing.

N400 is known to be involved in semantic language processing. This ERP component shows the use of bottom-up and top-down attentive pathways to make sense of the world around us. The N400 component has been shown to reflect a deviant stimulus (based on the current stimulus as well as implicit knowledge about the world) in a semantic sentence. The N400 congruency effect has also been well documented. Wang et al. (2017) showed that despite gender priming, gender-biased words, and semantically congruent and incongruent words being semantically incongruent, congruency in gender elicited a smaller N400 wave then incongruence. They also showed that when words shared a social or stereotypical characteristic, they were processed faster and more intensively. Hallam, Rogers and Creech (2008) explored which instruments were preferred by children and adolescents of different genders. They surveyed the proportions of male and female children and adolescents (aged 5-19) learning each instrument offered by a music service in the UK. They found that Male-preferred instruments include electric guitar, drums, trombone, and tuba, while female-preferred instruments include: violin, harp, oboe, clarinet, and flute. Koelsch et al. (2004) found that the N400 component was elicited when target words were preceded by semantically unrelated musical primes

In our experiment we asked if gender priming using digital emoticons affected the congruence with either gender-biased instruments (as designated by Hallam et al. (2008)) or pitch height; it is commonly believed that higher pitches are associated with females and lower pitched are associated with males. Along with asking if the presentation of a gendered face prime the perception of a timbre as gendered, and/or the perception of pitch height as gendered, we asked if there's an interaction between timbre and pitch height in the perception of gender in instrumental tones. We hypothesized that If the gender of the face and the timbre match, the two will be perceived as congruent. The gender of the face should not prime the perception of pitch height – for example, male faces will be perceived to be neither congruent with low pitch height, or incongruent with high pitch height. Timbre should be the dominant factor on the elicitation of the N400 component.

Methods

Participant profile:
Participants (N = 7)

- Mean age = 23, SD = 3.16
- 3 female
- All right-handed
- 6 native English speaker
- 2 with absolute pitch
- Years of musical experience = 15, SD = 6.19

For the stimuli presented in this experiment we used synthesized instruments taken from Sibelius 7 Sounds. A 500ms note was presented followed by 500ms rest. Electric guitar and trombone were used for male gender instruments, clarinet and harp were used as female gender instruments, and piano and saxophone were used as gender neutral controls. The power of each stimuli across instruments used was equalized in MATLAB and set to 80 dB. The notes D3 and F3, and D5 and F5, were used. The faces presented in this experiment to prime the subject were emoticons with two male faces and two female faces as shown here:

- Male: 👱 🧔
- Female: 👱‍♀️ 👩

There was also a behavioural task used in this experiment for two reasons. Firstly, behavioural tasks are used to keep the subject attending the stimulus in order to record viable data. Secondly, the behavioural task data can later be analyzed for response time to draw further conclusions from out experiment. Subjects were given a remote to hold with both of their hands and when instructed on the screen they were to respond whether the face was male (button 1) or female (button 4).

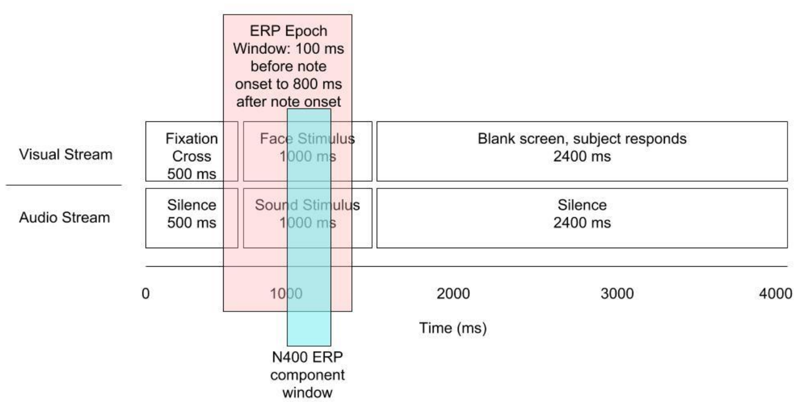


Figure 1: experimental design

The experimental design included two stimulus streams as can be seen in Figure 1. The visual stream consisted of a fixation cross for 500 ms to draw the attention of the subject to the middle of the screen followed by the presentation of the face stimulus for 1000 ms, followed by a blank screen when the subject was to respond whether the face was male or female. The audio stream consisted of silence for the first 500 ms when the fixation cross was on the screen. Then, when the was was flashed on the screen, the note was played for 1000 ms followed by 2400 ms of silence. Our epoch window (as seen in the red outline in Figure 1) was from 100 ms before the onset of the note to 800 ms after. The N400 ERP component was thought to lie from 250 to 450 ms after the onset of the note (Kutas, M., Federmeier, K.D., 2014). There were 96 trials per block and 4 blocks giving us 384 total trials. Each block took approximately 6 minutes and 15 seconds, making the total time for the recording portion of the experiment 25 minutes. EEG capping and preparation took another 30 minutes with time in between blocks for the subject. The total time to record each subject was an hour and a half.

Data analysis

Data was recorded in Curry 7 Neuroimaging suite and then transferred to Brainstorm, a plugin for MATLAB. The raw data underwent SSP eyeblink artifact rejection for blinks and horizontal eye movement. SSP peak-to-peak bad channel rejection at +/- 70 uV was also applied to the raw data. Each stim code was epoched from -100 ms to 800 ms and each individual trial was analyzed, removing further corrupted or bad data. A note should be made that the fourth block for subject NF contained corrupted data and so was disposed. Data was then broken down by timber and face as seen in Figure 2. The green cells represent

Instrument	Face (2 for each gender)	Pitch	Note	
Male-biased-1 *electric guitar*	M	Low	D (3 or 5)	F (3 or 5)
		High		
	F	Low		
		High		
Male-biased-2 *trombone*	M	Low		
		High		
	F	Low		
		High		
Female-biased-1 *clarinet*	M	Low		
		High		
	F	Low		
		High		
Female-biased-2 *harp*	M	Low		
		High		
	F	Low		
		High		
Neutral-1 *piano*	M	Low		
		High		
	F	Low		
		High		
Neutral-2 *saxophone*	M	Low		
		High		
	F	Low		
		High		

Figure 2: Data analysis breakdown for timber and face congruence. Green cells are congruent (C(T)) and orange cells are incongruent (IC(T)).

congruence and the orange cells represent incongruence. Data for pitch and face congruence was broken down similarly in Figure 3 with the same rules applied. A third condition was also

Instrument	Face (2 for each gender)	Pitch	Note	
Male-biased-1 *electric guitar*	M	Low	D (3 or 5)	F (3 or 5)
		High		
	F	Low		
		High		
Male-biased-2 *trombone*	M	Low		
		High		
	F	Low		
		High		
Female-biased-1 *clarinet*	M	Low		
		High		
	F	Low		
		High		
Female-biased-2 *harp*	M	Low		
		High		
	F	Low		
		High		
Neutral-1 *piano*	M	Low		
		High		
	F	Low		
		High		
Neutral-2 *saxophone*	M	Low		
		High		
	F	Low		
		High		

Figure 3: Data analysis breakdown for pitch and face congruence. Green cells are congruent (C(P)) and orange cells are incongruent (IC(P)).

analyzed for double congruence across timbre, pitch, and face gender. The analysis was

Instrument	Face (2 for each gender)	Pitch	Note	
Male-biased-1 *electric guitar*	M	Low	D (3 or 5)	F (3 or 5)
		High		
	F	Low		
		High		
Male-biased-2 *trombone*	M	Low		
		High		
	F	Low		
		High		
Female-biased-1 *clarinet*	M	Low		
		High		
	F	Low		
		High		
Female-biased-2 *harp*	M	Low		
		High		
	F	Low		
		High		
Neutral-1 *piano*	M	Low		
		High		
	F	Low		
		High		
Neutral-2 *saxophone*	M	Low		
		High		
	F	Low		
		High		

Figure 4: Data analysis breakdown for timbre, pitch, and face congruence. Green cells are doubly congruent (DC) and orange cells are doubly incongruent (DIC).

composed in a similar matter as the first two conditions. See figure 4. Difference waves were also analyzed for condition 1, D(T), condition 2 D(P), and condition 3 (DIC-DC).

Results

Mean latency and amplitude for each condition where early negativity component and N400 ERP were found to be most negative.

Condition	C(T)	IC(T)	D(T)	C(P)	IC(P)	D(P)	DC	DIC	DIC-DC
Early negativity component amplitude (uV)	-2.8	-2.75	-.6	-2.99	-2.8	-.4	2.98	-2.8	-.6
Early negativity component latency (ms)	162	164	170	162	168	146	162	172	172
N400 amplitude (uV)	-3.2	-2.75	-.2	-3.25	-3.08	-.25	-3.46	-3.22	-.4
N400 latency (ms)	308	314	352	314	314	336	314	314	348

Apart from the difference waveforms, the latency and amplitude of the N400 component in each condition remain remarkably consistent was was previously stated. In each instance, the incongruent function was found to have a smaller amplitude than the congruent function.

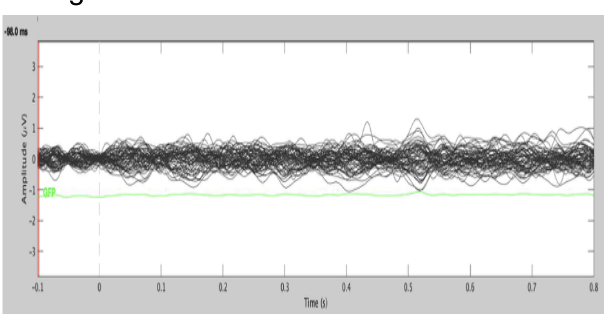


Figure 5: Butterfly plot of D(T)

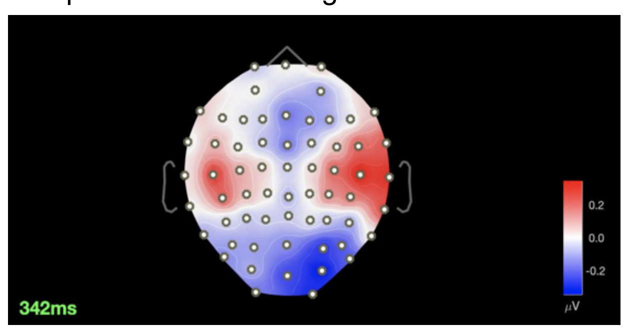
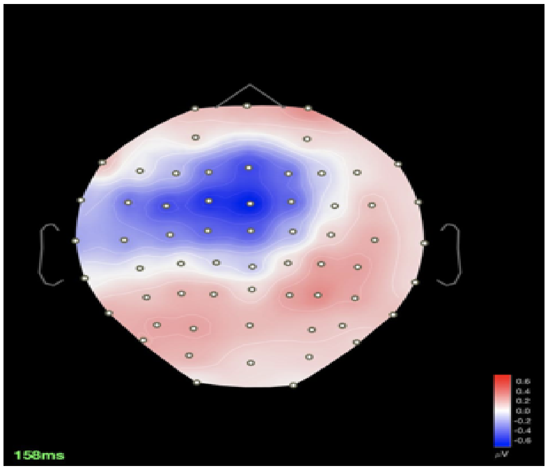


Figure 6: Topomap of D(T)

There was a slight negativity reported around 340-350 ms which had a maximum displacement of -0.2 uV which is barely significant to show any link between timbre and face gender incongruence. For this condition there was also centro-occipital negativity peaking at around 160 ms (see figures 7 and 8) which could possibly be the N170 ERP component which processes faces.



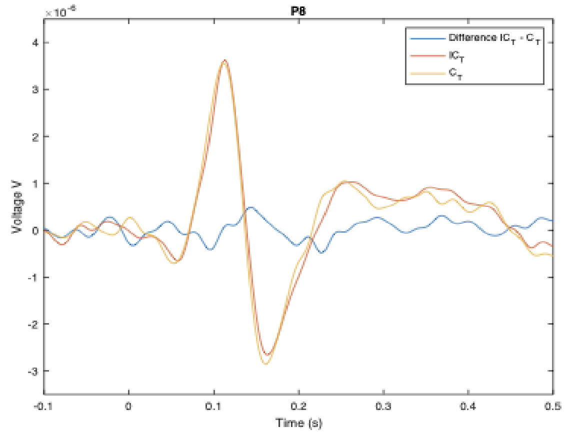
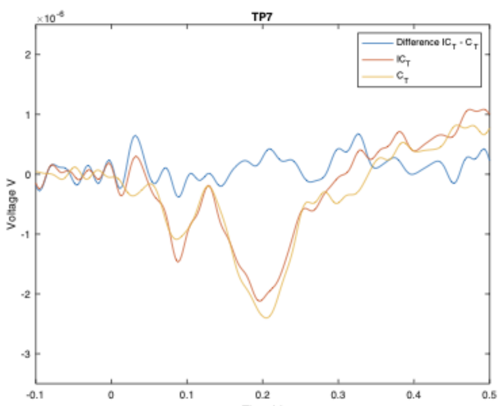


Figure 7: Topomap of frontal negativity D(T) *Figure 8: Condition 1 P8 electrode overlay*

The N170 component was centered around P8 (figure 9) where there is a sharp positive deflection at 100 ms – this is most likely the P100 component which is responsible for preattentive processing. At 160 ms there is a sharp negative deflection which could be the N170 component which displays the processing of the faces presented in the experiment. This wave had a decent negativity (-2.75 uV) and was consistent in both trials. Looking forward to the N400 component in figures 10 and 11 there is a negative deflection at 200 ms for electrode TP7 which could still be the N170 component. In figure 11 there is a decent negativity peaking at 350 ms. This is the right latency and amplitude for an N400 component but there is not change in the change in the difference wave which tells us that timbre and gendered face congruence vs. incongruence had no effect.



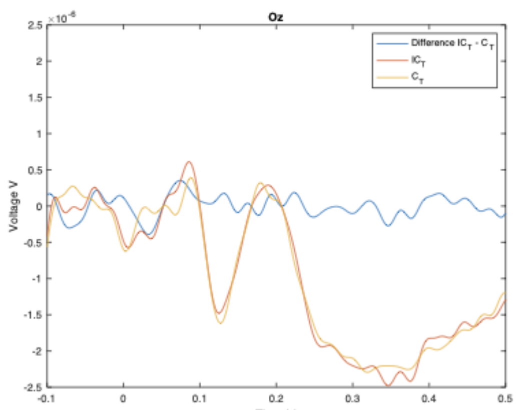


Figure 9: Condition 1 TP7 electrode overlay *Figure 10: Condition 1 Oz electrode overlay.*

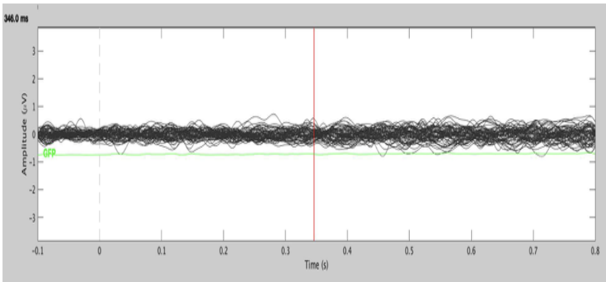


Figure 11: Condition 2 butterfly plot D(P)

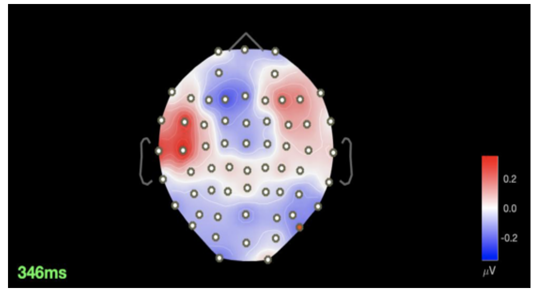


Figure 12: Condition 2 N400 topomap D(P)

In condition 2 we found similar results. The most likely candidate for the N400 component in pitch by gendered face congruence vs. incongruence was at 340-350 ms. In figure 12 the slight negativity concentrated in the occipital lobe and the negative deflection on the Oz electrode in figure 14 which peaks at -2.4 uV at 350 ms has the characteristics of the N400 component. This is evidence for neural processing of the semantic structure of the stimuli presented. However, the lack of any difference in the congruence vs. incongruence functions is evidence for an absence of neural processing for semantic incongruencies between gender priming with a face and the pitch presented (high pitches are characteristic of female and low pitches are characteristic of males).

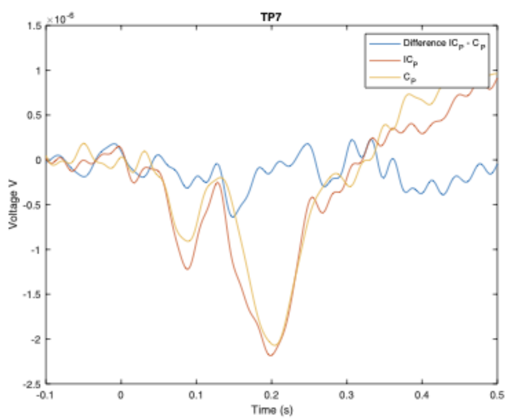


Figure 13: Condition 2 TP7 electrode overlay

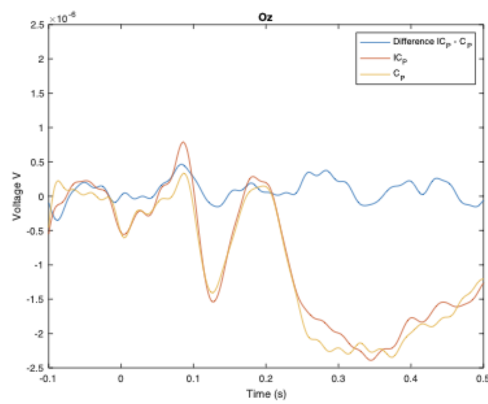


Figure 14: Condition 2 Oz electrode overlay

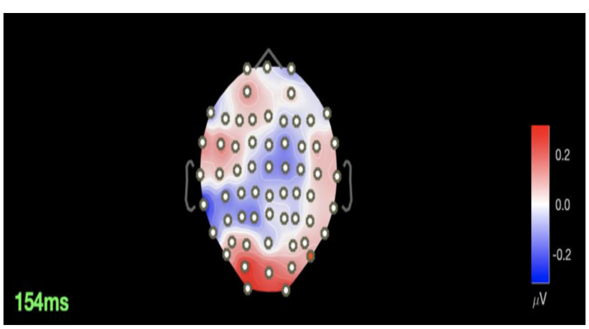


Figure 15: Condition 2 N170 topomap D(P)

Again, like in condition 1 there is a fronto-occipital negativity at 150-160 ms as seen in the topography centered around Fz in figure 15. The sharpt negative deflection for both the congruent and incongruent fucntions in condition 2 (see figure 13) shows negativity in the temporal region reaching a maximum at 200 ms. This is what we might expect from the N170 component. Most likely, this is evidence for the processing of the faces presented in the visual stream in the experiment.

Condition 3 looked to analyze functions which were doubly congruent and incongruent for timbre, pitch, and face gender. The difference waveform butterfly plot and topography for DIC-DC (doubly incongruent - doubly congruent) shows a small concentration of negativity (-0.4 uV) in the occipital region at 340-360 ms. This latency matches what we would expect in the N400 component congruent factor and there is some amplitude. It could be debated whether it is statistically significant, but it could also be a weak causal relationship between timbre, pitch, and gender priming congruence vs. incongruence.

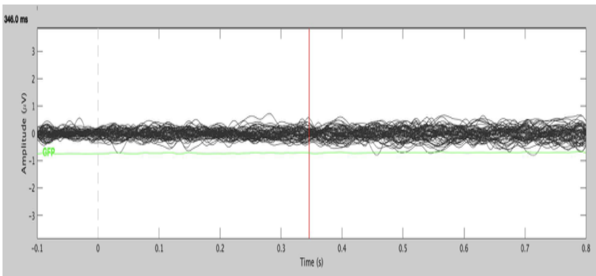


Figure 16: Condition 3 butterfly plot DIC-DC

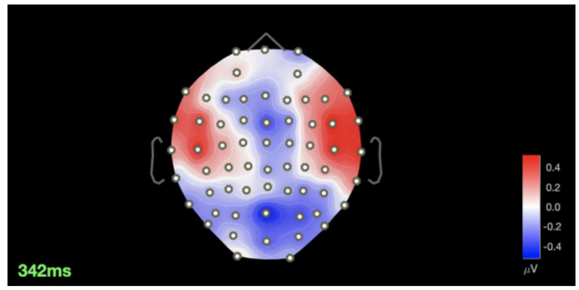


Figure 17: Condition 3 topomap DIC-DC

The CB2 and Oz electrode overlay plot has consistent similarities with conditions 1 and 2 with a slow negative peak starting at 200 ms and reaching an absolute maximum (negative) at 300 ms. There is, however, no difference between the congruent and incongruent factors which suggests that perhaps N400 wasn't elicited by the congruent factor of timbre vs. pitch vs. gender, but rather by the semantic processing of music.

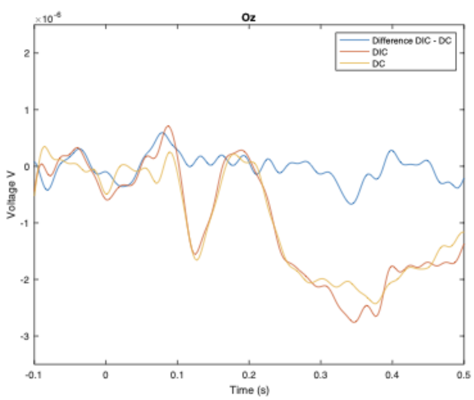


Figure 18: Condition 3 Oz electrode overlay

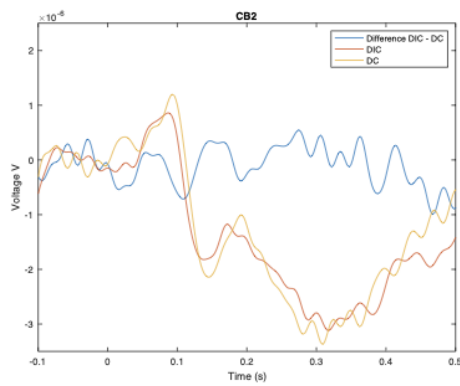
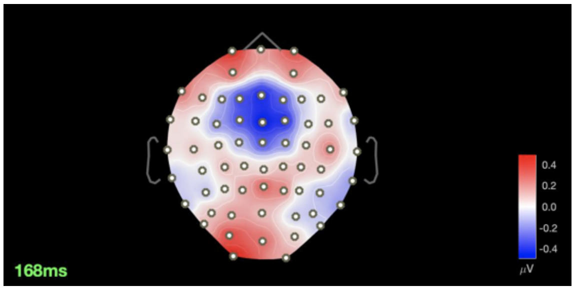


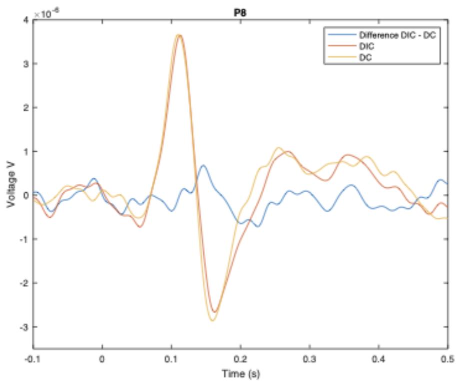
Figure 19: Condition 3 CB2 electrode overlay

Just as in conditions 1 and 2 there is the fronto-occipital negativity peaking at 168 ms in the individual functions. The lack of a substantial negativity in the difference topomap in figure 20 tells us that it is not typical of MMN. Rather, this could again by the N170 component which is characteristic of facial processing. N170 is most concentration in the temporal regions and in figure 22 a clear negative peak can be seen reaching its maximum at 200 ms.



Just as in conditions 1 and 2 there is the earlier negative deflection seen in the line plot overlay of the P8 electrode. This could possibly be N170.

Figure 20: Condition 3 N170 topomap DIC-DC



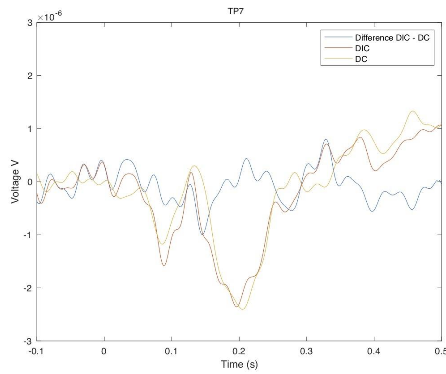


Figure 21: Condition 3 P8 electrode overlay *Figure 22: Condition 3 TP7 electrode overlay*

Discussion

ERP component/condition	timbre x face	pitch height x face	timbre x pitch height x face
MMN-like negativity	-0.6µV, 158ms, FCz	-0.3µV, 154ms, TP7 (but this seems like an artifact)	-0.5µV, 168ms, FCz
N400	-0.3µV, 342ms, CB2	-0.1µV, 346ms, Oz	-0.6µV, 342ms, Oz

Figure 23: The overall mean amplitudes of each ERP component of interest in the most prevalent electrode.

A very miniscule negativity as found consistently between 300 - 400 ms which is the time window if interest for N400. There is no difference in negativity however, like we would expect. Koelsch et al. (2004) found differences in the negativity for non congruent functions. Our experiment did not answer our research questions and failed to back up our hypothesis. From data gathered in this pilot study there is no evidence that gender priming affects the semantic processing of gender biased instruments and pitch corresponding to high vs. low voices. This study drew from the paradigms of the Koelsch et al. (2004), Wang et al. (2017), and Hallam et al. (2008).

While we cannot prove that gender priming affects the semantic processing of gender biased instruments, we also can't determine with confidence if this is because semantic language paradigms such as those in Koelsch et al. (2004) can't be elicited by music, or if the brain can't determine gender biases from instruments. Wang et al. (2008) suggested that gender stereotypes in semantic word pairs elicited a larger N400 component than lexical semantic word pairs. Therefore, that might suggest that gender is something which is easily processed and know throughout speech processing. We did not find such a result. Hallam et al. (2008) found that populations which study music tend to fall into specific bins segregated by gender. For example, in any one population of musicians, there are more likely to be more females who play the harp as males just as there are more likely to be more males who play the electric guitar that females. When looking at Condition 1 and 3 which examined the effects of musical timbre on gender congruence we found little effect on eliciting a stronger or weaker N400. Both responses for timbre vs. face and pitch vs. face really weak, but timbre x face had a slightly stronger negativity than pitch height x face. There appeared to be some sort of interaction between the two. This was shown in condition 3 where we found more negativity in the DIC-DC condition

On piece of our experiment which could have hindered the development of the N400 wave could have been the instrument tones used to form the auditory stimuli. After further review it was evident that the synthesized instrument timbre used from Sibelius Sounds 7 lacked the authentic timbre one might hear in person. This could explain why the N400 component we found had a weak negativity. Another variable which could be changed for further research into this topic was the time at which the faces were presented before the sound. There was a 60 ms delay between the onset of the visual stimuli and the onset of the audio. This could not have been enough time to fully prime the visual and auditory cortices to make top-down processing decisions on the sound which followed. Further research into the hypothesis postulated in this experiment might require adjusting the timing of the stimuli, or controlling for the pitch so as to only isolate the instrument timbre and face priming. Another potential avenue for further reading could be into the use of an audio primer instead of a face. The elimination of the cross-modality integration between the visual and auditory cortices could isolate N400 to create a stronger response.

Out data was very consistent throughout each condition which provides evidence that there was little artifact or other experimental noise clouding our data. Though the negativity where N400 should be found wasn't as strong as shown in Koelsch et al. (2004), we had distinct P1, N1, N2, and P2 wave components. The N400 which we did examine was, on average, 60 ms before the latency at which N400 was found in Koelsch et al. (2004). While each of the paradigms drawn upon in this experiment showed strong evidence for semantic processing of gender stereotypes and separate semantic processing of music, our results show little ability for the brain to use both indices to gender-prime a musical instrument timbre or pitch.

References

Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: finding meaning in the N400 component of the event-related brain potential (ERP). Annual review of psychology, 62, 621-47.

Hallam, S., Rogers, L., & Creech, A. (2008). Gender differences in musical instrument choice. International Journal of Music Education, 26(1), 7–19. https://doi.org/10.1177/0255761407085646

Pei Wang, Ya-Ping Yang, Chen-Hao Tan, Qin-Wei Chen & Thomas Cantfort (2017) Gender Stereotype Activation versus Lexical Semantic Activation: An ERP Study, The Journal of General Psychology, 144:4, 283-308, DOI: 10.1080/00221309.2017.1310714

Koelsch, S., Kasper, E., Sammler, D., Schulze, K., Gunter, T., & Friederici, A. D. (2004). Music, language and meaning: Brain signatures of semantic processing. Nature Neuroscience, 7, 302. Retrieved from https://doi.org/10.1038/nn1197