Binaural processing by temporal envelope and the effects of carrier frequencies

Nick Gang & Wisam Reid

1 Project Background

Distinguishing auditory events based on temporal features is an integral part of perceiving the world around us. Minute changes in timing, phase and intensity can inform us about where an object is located, if it is moving, its periodicity, and can allow us to differentiate sound sources from one another [Picton, 2013].

Specifically, interaural phase differences (IPD), interaural timing differences (ITD), and interaural intensity differenes (IID) are important in horizontal sound localization. In their 2007 study, Ross et al measured the detection of subtle shifts in IPD by binaurally presenting a diotic signal (same in both ears) and introducing a phase shift only in the right ear [Ross et al., 2007]. This created a dichotic signal with an IPD between the left and right ear of a quarter period. Response to this shift was examined using magnetoencephalography (MEG) and also behaviorally by asking participants to report if the tone moved. This study found that the carrier frequency significantly affected the perception of the IPD. As the length of a quarter period of the carrier frequency decreased below the distance between the ears (~ 1250 Hz), both the change response observed with MEG and the behavioral responses disappeared. Past this threshold, participants showed no observable difference in perception of the diotic and dichotic tones.

• *W. Reid and N. Gang are with the Department of Music, Stanford University, CA, 94305. Web: wisamreid.com, orangagang.com*

2 Introduction

Past research has also found evidence that complex features are used in sound localization, and an opponent-channel model was supported by the finding of spatially selective neurons in the inferior colliculus [Grothe et al., 2010]. The opponent-channel encoding of sound localization in bilateral auditory cortices has been further supported by MEG studies finding the opposite polarity of binaural beat responses at low frequencies [Ross et al., 2014]. Fluctuations in the temporal envelope of a sound have been shown to provide necessary information for speech processing [Ligeois-Chauvel et al., 2004]. However, little research exists that examines the effects of a shift in temporal envelope on perceived sound localization and more specifically, the effects of carrier frequency on this perception.

The current study aimed to determine the behavior of the change response to a non-varying shift in temporal envelope of AM tones with varying carrier frequencies. Rather than changing the phase of the carrier waves in one ear as [Ross et al., 2007] did, the stimuli in this study shifted the phase of the modulating envelope. This was meant to create a change in perceived sound localization that was neither strictly an IPD or an IID.

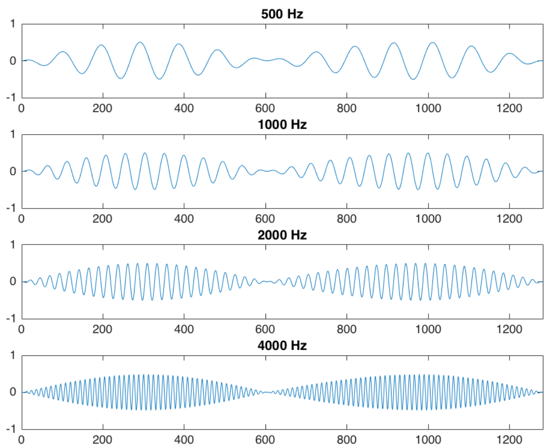

Fig. 1. Stimulus Centered on Temporal Envelope Shift

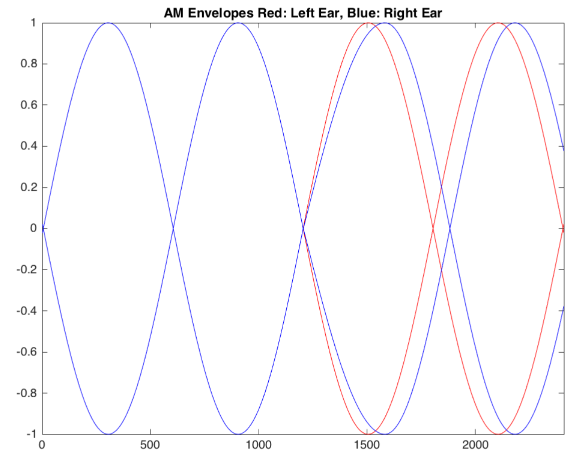

Fig. 2. Amplitude Modulation Envelopes

2.1 Hypothesis

If this change in temporal envelope is perceived more similarly to a IID, we should not expect to see the same rolloff in change response magnitude with a rise in carrier frequency. This is due to the fact that the intensity of the complex signals does not change with the carrier frequency. However, if the change in temporal envelope is perceived similarly to an IPD, we should expect to see the change response decrease in magnitude as the carrier wavelength decreases to below the distance between the human ears.

It was hypothesized that the temporal shift would be perceived as a complex phase shift and thus, the change response should decrease as a function of the carrier frequency. In order to compare the processing of temporal envelope differences with IPDs in this study, carrier frequencies both above and below the IPD detection threshold were chosen. The AM frequency of 40 Hz was chosen to create the most prominent steady state response. This is the same frequency chosen by Ross et al in their aforementioned study.

3 Methods

Subjects were presented with 4 second long amplitude modulated sine tones with a change in temporal envelope after 2 seconds. Between each tone was 1.5 seconds of silence for a total stimulus onset asynchrony (SOA) of 5.5 seconds. The carrier frequency of these tones were either 500 Hz, 1000 Hz, 2000 Hz, or 4000 Hz. The amplitude modulation frequency stayed constant at 40 Hz **Fig 1**.

3.1 The Stimuli

For the first 2 seconds of each tone, both ears received the same diotic signal. At 2 seconds, the phase of the temporal envelope in the right ear was shifted by 1.5 ms, equivalent to a quarter cycle of the 40 Hz modulation. To smooth this transition, a 1.5 ms section of a 31.8471 Hz AM signal was fit to this delay

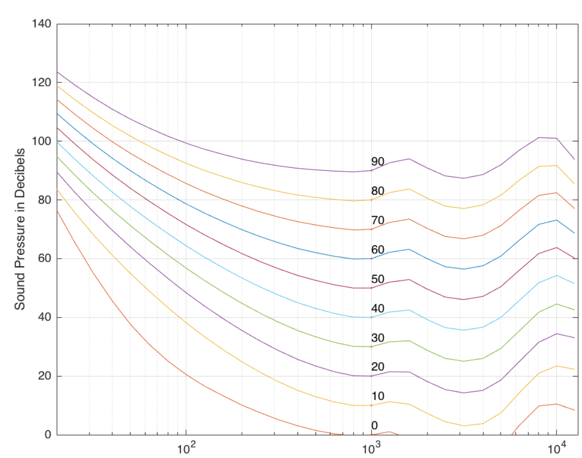

Fig. 3. Equal loudness contours

Fig 2. resulting in a temporal envelope difference between the left and right ear inputs. Note that the carrier frequency phase was kept identical between the ears throughout the tone duration. The effect of this was a delay in temporal envelope that was not perceivable when listening only to the signal in the right ear. This created a dichotic signal when heard binaurally, shifting the apparent location of the sound. While the carrier frequency varied by condition, the phase of the carrier signal was always kept identical between the left and right ears. To prevent a discontinuity at the end of the delayed signal, the last 6.25 ms of both channels were ramped to zero with a cosine curve.

3.1.1 *Equal Loudness Compensation*

In order to control for the apparent loudness of the 4 different carrier frequencies, an equal loudness curve (ELC) was applied to the stimuli. Although based on perceived loudnesses, the dB[EQL] sound pressure weighting is not a specific loudness measurement, does not consider critical band formation, and does not yield masking or psychoacoustic loudness data [Bray, 2010]. The ELC served as a spectral 'weighting' intended to match the sensitivity of hearing as a function of frequency within "use-level" ranges, specifically accounting for cavum conchae resonance of the ear. Sound intensity was set as reference at 80 Phon for the 1000 Hz carrier frequency prior to amplitude modulation. Relative amplitudes for all other carrier frequencies were tuned according to their respective ISO 226:2003 standard ELC characteristics **Fig 3**. However, since these stimuli were complex tones, additional calibration was needed. A pilot hearing threshold test was conducted on 2 subjects. Both subjects heard stimuli from all 4 conditions starting at silence until they reported hearing the tone. The results from both subjects were averaged to create an additional volume offset and the averaged hearing threshold using STIM2 [Compumedics Nueroscan, 2015] stimulus presentation software. The final volume offsets were based on a sound intensity of 80 dB-SPL as

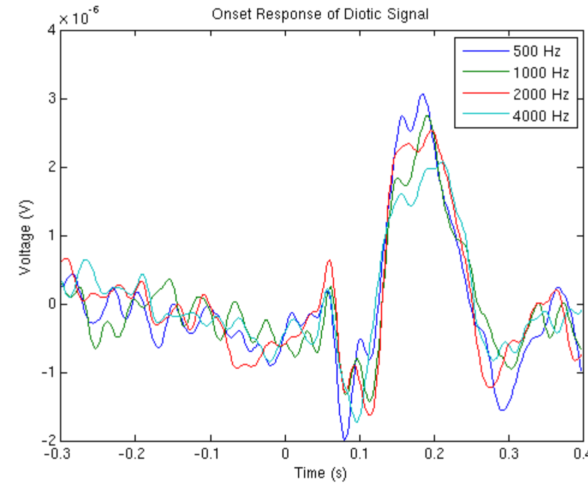

Fig. 5. Grand Average: Onset Response

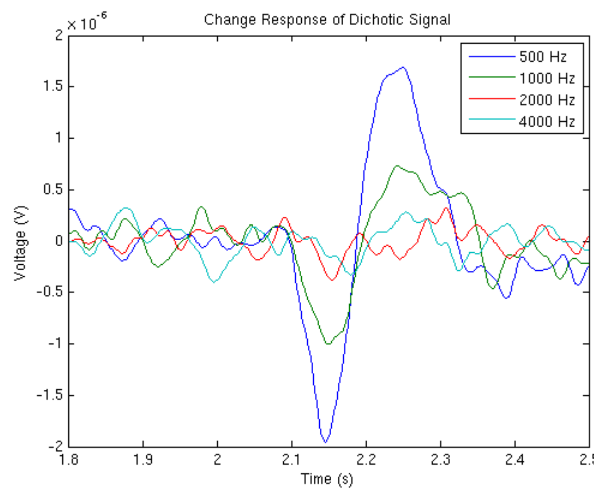

Fig. 6. Grand Average: Dichotic 500 Hz

Grand Average Time Series [FCz]								
	Onset Response				Change Response			
ERP Peak	N1		P2		N1		P2	
Carrier [Hz]	Latency [ms]	Peak [μV]	Latency [ms]	Peak [μV]	Latency [ms]	Peak [μV]	Latency [ms]	Peak [μV]
500	104	2.90	192	2.85	148	2.10	246	1.65
1000	98	2.90	196	2.80	148	1.05	246	0.75
2000	100	2.75	204	2.25	148	0.35	246	0.30
4000	96	2.30	194	2.30	148	0.30	246	0.75

Fig. 4. N1 and P2 Peaks and Latencies

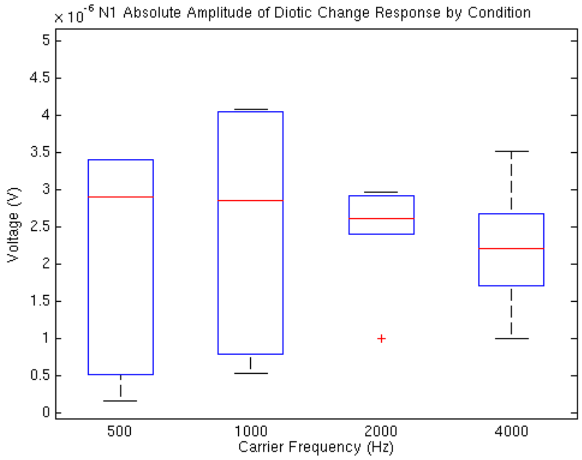

Fig. 7. N1 Onset Response

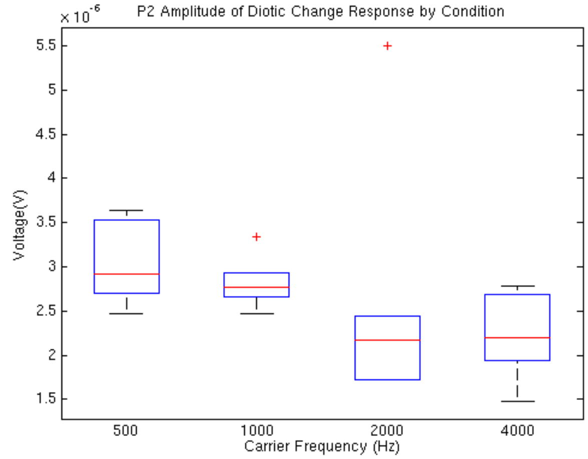

Fig. 8. P2 Onset Response

reference for the 1000 Hz stimulus, resulting in normailzed gains of 87, 75.5, 76 dB-SPL for the 500, 2000 and 4000 Hz stimulus respectively.

3.2 Participants

There were 6 participants in this study (3 male, 3 female) between the ages of 23-31. 5 of the 6 participants had some form of musical training. One participant was left-handed.

3.3 EEG recording procedure

Each block consisted of 100 trials in a single condition. Participants each experienced 8 blocks, 2 for each condition. This brought the time for each block to 9 minutes and 10 seconds, and the total time for each subject to 73 minutes and 20 seconds. Prep time and capping each participant added roughly 30 minutes to each session.

During the recording time, participants were seated in a comfortable reclining chair in an electrically shielded room. Participants were presented the movie or show of their choice with subtitles and were instructed to fixate their eyes on the screen and move a little as possible.

3.4 Data analysis

Before processing the data, a signal space projector (SSP) was created to detect eye blinks using data from electrodes located above and below the left eye and on each temple. This SSP was generated and applied with the Brainstorm3 Matlab toolbox [Neuroimage, 2011] using ocular artifact reduction (OAR) based on vertical electrooculographic (EEG-VEOG) and horizontal electrooculographic (EEG-HEOG) covariance analysis, derived from EEG measurements taken at the right temple, lower brow, upper check bone, and temple surrounding the left eye of each subject. This SSP was applied to each epoch to remove eye blink artifacts. There was no peak to peak artifact detection performed and all 200 trials from each condition were used to compute single subject and grand averages. The electrode FCz showed very strong onset and change responses, and was chosen to visualize the ERP components in the overlay and box plots.

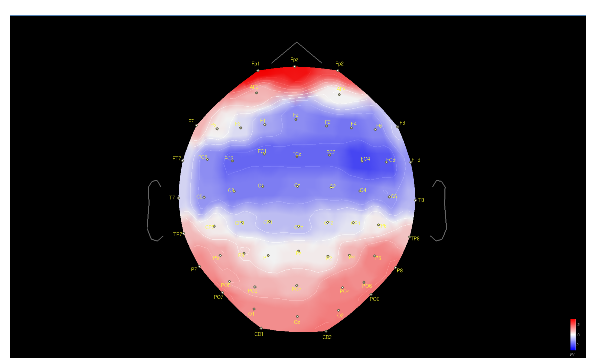

Fig. 9. Grand Average: Change N1 500 Hz

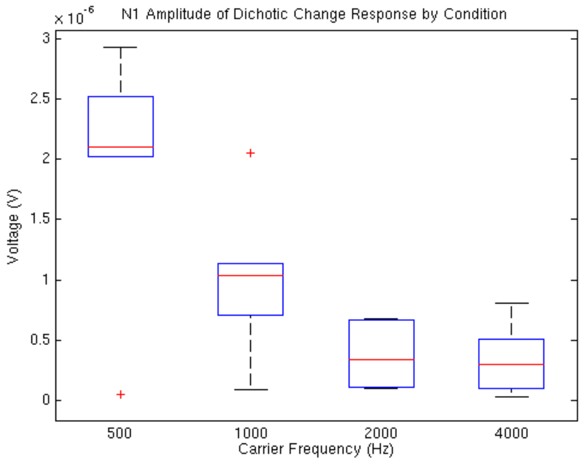

Fig. 10. N1 Change Response

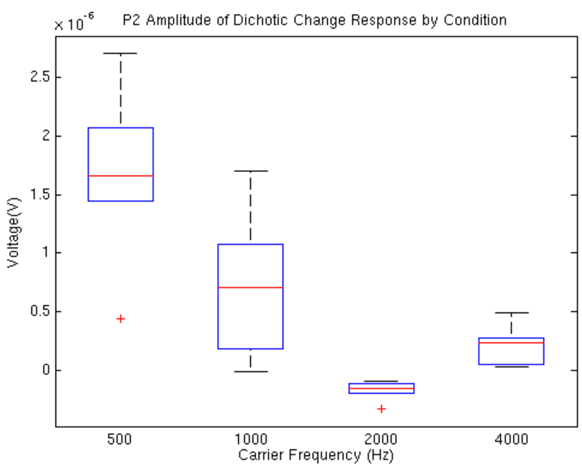

Fig. 11. P2 Change Response

4 Results

The diotic conditions showed no significant changes in peak and latency characteristics across conditions, confirming successful equal loudness calibration across carrier frequencies **Fig 5**. In contrast, a large difference was clearly seen between the 500 Hz stimulus and the 1000 Hz stimulus in the grand average time series waveforms **Fig 6** for the dichotic conditions. N1 peaks in the grand averged change response were observed at 148 ms latencies and were observed to be 2.10 and 1.05 μV for the 500 and 1000 Hz stimulus respectively. The P2 peaks were observed at 246 ms latencies and were seen to be 1.65 and 0.75 μV for the 500 and 1000 Hz stimulus respectively **Fig 4, 6**. Without further data collection and statistical analysis not much can be said about the difference in the 2000 Hz and 4000 Hz carrier frequencies ERPs. The responses at these these frequencies were significantly weaker, largely due to greater statistical variance between subjects **Fig 4, 6**. These results were confirmed by the topographic maps, which showed theses results with both the onset and change response being similarly front central localized **Fig 9**. Another view of this data and the underlying distribution across subjects can be seen in the box plots for the N1 and P2 peaks in both the onset and change responses **Fig 7, 8, 10, 11**. All time series plots for each individual participant, as well as topographic maps for the grand average can be seen in the attached **appendix**.

5 Discussion

While there is not enough data to determine statistical significance, these initial findings show that the response to a dichotic change in temporal envelope does decrease as the carrier frequency increases. However, above 2000 Hz this relationship was not consistent. The difference between the 500 Hz condition and the 1000 Hz condition and between the 1000 Hz condition and the 2000 Hz condition are quite obvious. A small change response may be present in both of these conditions, which could be an important finding. A difference in the detection threshold frequency between IPD shifts and temporal envelope shifts would suggest that the two features are being processed differently in the brain.

It seems logical that the threshold of temporal envelope detection would not be exactly related to the size of a quarter cycle of the carrier frequency, since the carrier frequency phase does not change over the course of the stimuli. However, the reasons for this behavior as a function of carrier frequency is unclear. There are several aspects of this study that could be improved in future revisions. First, the carrier frequencies chosen did not include those closest to the threshold found by Ross et al [Ross et al., 2007]. Adding frequencies near 1250 Hz could help to discern if the threshold for detection of temporal envelope difference differs from the threshold for IPD in carrier frequency. The drawbacks to adding more conditions like this would be the risk of fatiguing participants to the point where their data is no longer reliable.

Another future consideration will be to add both behavioral and psychoacoustic components to the study. It is ambiguous whether a change response is present in the 2 highest pitched conditions, but a difference may still have been perceived by the participants. Forming questions similar to Ross et al, future revisions could ask participants if they heard the sound move or if there was any apparent change in timbre. With more research, more participants, and improved control, this study has the potential to further the understanding the effect of complex audio features on sound localization.

REFERENCES

[Bray, 2010] Bray, W. (2010). The db[eql]: An alternative sound pressure weighting according to the equal loudness contours of the international standard iso 2262003. *The Journal of the Acoustical Society of America*, 127(3).

[Compumedics Nueroscan, 2015] Compumedics Nueroscan (2015). Stim.

[Grothe et al., 2010] Grothe, B., Pecka, M., and McAlpine, D. (2010). Mechanisms of sound localization in mammals. *Physiological Reviews*, 90(3):983–1012.

[Ligeois-Chauvel et al., 2004] Ligeois-Chauvel, C., Lorenzi, C., Trbuchon, A., Rgis, J., and Chauvel, P. (2004). Temporal envelope processing in the human left and right auditory cortices. *Cerebral Cortex*, 14(7):731–740.

[Neuroimage, 2011] Neuroimage (2011). Brainstorm.

[Picton, 2013] Picton, T. (2013). Hearing in Time: Evoked Potential Studies of Temporal Processing. *Ear and Hearing KW -*, 34(4).

[Ross et al., 2014] Ross, B., Miyazaki, T., Thompson, J., Jamali, S., and Fujioka, T. (2014). Human cortical responses to slow and fast binaural beats reveal multiple mechanisms of binaural hearing. *Journal of Neurophysiology*, 112(8):1871–1884.

[Ross et al., 2007] Ross, B., Tremblay, K. L., and Picton, T. W. (2007). Physiological detection of interaural phase differences. *The Journal of the Acoustical Society of America*, 121(2).