Title

Endogenous Auditory Event-Related Potentials of Feature Selective Attention and the Transition Bandwidths of Automatic Attention

Author

Bellato, Michael Warren

Publication Date

2017

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA,

IRVINE

Endogenous Auditory Event-Related Potentials of Feature Selective Attention and the Transition Bandwidths of Automatic Attention

DISSERTATION

submitted in partial satisfaction of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

in Psychology

By

Michael Bellato

Dissertation Committee:
Professor Bruce Berg, Chair
Professor Ramesh Srinivasan
Professor Kourosh Saberi

2017

DEDICATION

To my sisters for being by my side through all of the ups and downs.

“*Science is not only a discipline of reason but, also, one of romance and passion*”

Stephen Hawking

“*You miss 100% of the shots you don’t take*”

Wayne Gretzky

TABLE OF CONTENTS

LIST OF FIGURES

ACKNOWLEDGMENTS

To my advisor, Dr. Bruce Berg, for teaching me to have confidence in my studies during our time together when I was an undergraduate and for teaching me to have confidence in myself while I was a graduate student. Thank you for teaching me to take aim at any fundamental assumptions. These traits will serve me for the rest of my life, both as a researcher and as a person.

To my wife Hayley, for giving me your unconditional love and support. Words will never capture the amount of gratitude that I have for you. Thanks for keeping a grad student fed.

To my committee members, Dr. Kourosh Saberi and Dr Ramesh Srinivasan, for providing our department with an EEG facility so that graduate students have the opportunity to pursue experiments that they would otherwise not have access to.

To my lab mates, Dr. Borucki, Dr. Tan, Matt Richardson and Tyler Kaczmarek. Thank you for exposing me to new ideas and helping me develop my skills in the lab. I also deeply appreciate the conversations that had nothing to do with our work.

To Ani Orujyan, Anais Panossian, and Grant Martin for your help with capping our EEG subjects.

Curriculum Vitae

Michael Bellato

707 Leahy Street Apt 121
Redwood City, CA 94061
Mobile: (310) 699-5166
E-mail: mbellato@uci.edu

Education

M.A./Ph.D. University of California, Irvine, Psychology, 2017

B.A. University of California, Irvine, Psychology & Political Science, 2011

Lund University, Sweden, Education Abroad Program, 2011

Research Experience *University of California, Irvine – Dept. of Cognitive Sciences*

Graduate Student Researcher for Dr. Bruce Berg *September 2012 to August 2017*

Combining psychophysical measurements with electroencephalogram (EEG) to understand how humans attend to an auditory feature (e.g. pitch)

Investigating the automatic attention capabilities of the auditory system via Transition Bandwidths, reflected by changes in thresholds

Undergraduate Researcher Assistant for Dr. Bruce Berg *September 2011 to June 2012*

Researched temporal and spectral processing systems of the human auditory pathway through psychophysical measurements

Research Project for Experimental Psychology course (Dr. Michael Lee) *Spring 2010*

Designed a face-sorting experiment to determine that people can make abstract judgments and categorize faces

University of California, Irvine – Dept. of Informatix

Graduate Research Assistant for Dr. Alfred Kobsa *December 2015 to June 2017*

Applying Psychological research methods to user experience research to explain how environmental distracters inhibit CAPTCHA performance

Teaching Experience

Instructor, University of California, Irvine *August 2015 to September 2015*

Psychology Fundamentals (9A)

Teaching Assistant, University of California, Irvine *September 2012 to August 2017*

Teaching Assistant for 16 terms

Honors and Awards

Psi Chi Honors, University of California, Irvine *Spring 2015*

International Honors in Psychology awarded for commitment to research

Honors and Awards (continued)

Best Project, Racial Violence Archive – iConference 2015, Social Media Expo *March 2015*
Received the award for "Best Project". The award was for the most innovative application of government open data to social media.

Social Sciences Student Scholar, University of California, Irvine *June 2010*
Awarded to students who demonstrate a high level of scholarly achievement in the school of Social Sciences while contributing to academically based on-campus organizations

Publications

Conference

Lee, Hosub; **Bellato, Michael**; Jain, Sowmya; Spanghero, Fernando; Singer-heinze, Roeland; Lin, Ya-Wen; Gupta, Sunakshi; Ward, Geoff; Kobsa, Alfred. *Racial Violence Archive: Public Information System on Incidents of Violence during the Civil Rights Period*
iConference *Newport Beach, March, 2015*

Berg, Bruce; **Bellato, Michael**; Baltzell, Luke. *Increased Reaction Time Adherence to Response Criterion in a Sample Discrimination Task*
American Acoustical Society *Boston, June 2017*

Presentations

Differentiating Endogenous Auditory Event Related Potentials: Top-Down Feature Selectivity
Luce Graduate Conference *Irvine, May, 2015*
UC Irvine Institute for Mathematical and Behavioral Sciences

Racial Violence Archive
iConference, Social Media Expo *Newport Beach, March, 2015*
Selection was based on contributing government open data in a social media platform.

Poster Presentations

Berg, B., **Bellato, M.**, & Baltzell, L. *Peak Reaction Time Adheres to Criterion*
Acoustical Society of America *Boston, June, 2017*

Bellato, M., Berg, B., & Richardson, M. *Differentiating Endogenous Auditory Event Related Potentials: Top-Down Feature Selectivity*
UC Irvine, Data Science Initiative *Irvine, May, 2015*

Bellato, M., Andrews, L., & Lee, M. *Categorical Perception in Making Judgments on Facial Trustworthiness*
UC Berkeley, Psychology Undergraduate Research Conference *Berkeley, June, 2011*

Funding

Microsoft Research – Racial Violence Archive *December, 2014*
Granted $3,000 to develop an application of government open data provided in social media. Presented at iConference, 2015.

UC Irvine, Cognitive Sciences Summer Research Stipend *Summer, 2013 & 2014*

ABSTRACT OF THE DISSERTATION

Endogenous Auditory Event-Related Potentials of Feature Selective Attention and the Transition Bandwidths of Automatic Attention

By

Michael Bellato

Doctor of Philosophy in Psychology

University of California, Irvine, 2017

Professor Bruce Berg, Chair

The ability to engage in auditory selective attention relies on listeners discriminating between acoustic features that are simultaneously available within a sound and can differ between auditory objects. Quantifying the qualitative state of these acoustic features has been advanced via COSS analysis. This methodology employs identical stimuli across trials, while directing the listener to attend to one of several acoustic features (e.g. loudness, timbre, or pitch). Weight profiles for loudness and timbre judgments were replicated from previous experiments, while under EEG recordings. While listeners were engaged in making loudness judgments, a characteristic ERP response of the stimulus fundamental frequency (F0) and second harmonic (2F0) was observed. Timbre ERPs displayed a similar response to F0, but a decrease in activity related to 2F0. The dissertation then moves to studying automatic attention. Transition bandwidths are calculated by estimating thresholds at various bandwidths as the number of components in a complex sound is increased at a fixed frequency difference. Thresholds increase with the number of components as the signal to noise ratio is decreased. Eventually a breakpoint is achieved where thresholds decrease as listeners perform profile analysis. This experiment

manipulates previous experiments by fixing the phase of each component, instead of randomly sampling each phase independently. Results show a dramatic decrease in threshold estimates associated with profile listening from randomized phase to fixed phase conditions. The physiological basis of this change is investigated on the cortical level. EEG recordings show that F0 power increases with stimulus bandwidth.

Chapter 1

Introduction

1.1 Listening Strategies

Multiple mechanisms process auditory information, allowing for listeners to perform task oriented behaviors: a process that we call attention. Although common people and researchers alike use the term "attention" frequently, the definition and study of this proves challenging. In human listeners, the summation of attention related mechanisms provides listeners with their sense of listening. This dissertation generalizes auditory attention to have both automatic and selective attributes. The first form of attention discussed is selective attention, where the listener selectively attends to either a particular feature of a stimulus or an auditory object, while other features/objects are simultaneously present. Auditory features, such as loudness, timbre (roughness), pitch and location are derived from stimulus statistics including the amplitude, duration, phase, envelope and frequency of a sound. All of the above features, except for localization are discussed herein. Second to be discussed is the automatic auditory attention, where attention is manipulated by the characteristics of the stimulus.

Listeners have the ability to isolate an acoustic feature in the presence of other features, at the cost of acquiring information from those other features (Hafter, 2007). Hafter also notes how the theories on attention have a wide range of focus, ranging from feature binding (Treisman and Gelade, 1980) to feature segregation. The latter of the two attention theories is of interest to the following research, as the research is concerned with auditory features.

Evidence of the existence of two distinct forms of auditory attention is supported by dichotic auditory streaming tasks (Cherry 1953). Initially, listeners performed a speech shadowing task by verbally repeating the target speech from two mixed speeches. Highly probable phrases in dual streams were played simultaneously, subjecting the listener to equally probable selection of stream. Dichotic speech tasks allowed the subject to effortlessly shadow the target stream of speech while rejecting the non-target speech. However, little could be recalled by the listener in the dichotic tasks. Finally, alterations were made to the non-target stream of speech, in either the details of the speech or the structure of the sound (e.g. the language of the speaker vs. the sex of the speaker). Listeners were unaware of a language change in the unattended ear, but were able to identify the change of speaker sex and the presence of a 400 Hz tone. This dichotic alteration suggests a hierarchical mechanism of attention where listeners can process the content of a target sound source, but are limited in their ability to process non-target information. Such mechanisms are of particular relevance for the studies included in this dissertation.

Automatic attention suggests that a listener's attention can be restricted and therefore justifies the study of the auditory periphery where sound is processed in both spectral and temporal domains. These two mechanisms are not consciously employed by the observer and are therefore of concern when studying auditory attention.

The study of auditory attention has gained insights via electrophysiology studies that reveal the neural processes that are associated with multiple attention mechanisms. Such studies begin with the assumption that the physiological response to a stimulus is a mixing of the biological processes performed succeeding transduction and the active cognitive mechanisms of the observer. The former of these two assumptions includes the exogenous auditory event related potential. This response can be recorded from the scalp when a listener receives auditory input while either not attentive to, or suppressing that information. Exogenous ERPs are the sum of neural mechanisms that process an input which does not provide information related to the task the listener is engaged in. Endogenous brain potentials are the sum of neural components that are directly related to the task engagement of the listener, as they exploit task-relevant information.

The endogenous ERP can be measured by the difference between ERPs recorded during passive listening from those recorded during active listening. Hansen and Hillayd demonstrated this by measuring the N1 wave deflection during a binaural tone pip task (Hansen and Hillyard, 1979). The frequency separation between high and low tone pips was varied across three conditions. Exogenous recordings were collected while the stimulus was played for the subjects as they were engaged in a reading task, and no feedback was required from the subjects. The endogenous recordings were collected as subjects identified the signal from a set of standards, which were a longer duration or a shorter duration, respectively. The difference in evoked potentials across conditions was a function of the frequency difference between tone pips. Furthermore, d’ calculations for all subjects found a significant relationship between the magnitude of the difference between EEG recordings and the efficiency of the listener.

Despite psychophysical, technical and computational advances, cognitive electrophysiology is not without limitations. The ERP provides details related to the physiological response to the stimulus and the endogenous state of the listener. However, the ERP only represents the sum of all associated components, the decomposition of which can be ambiguous. For example, the amplitude of the N1 component, associated with attention, can be influenced by stimulus characteristics such as intensity (Schwent et al., 1976). The source of this effect could either be an increase in the amplitude of that negative component, or the addition of a new negative component (Hillyard et al., 1978).

Controlling measurements of attention by contrasting against measures of passive exposure has provided a macroscopic approach to understanding the mechanisms responsible task relevant behavior. However, this contrast does not provide microscopic details that distinguish between separate distinct listening strategies. The data collection methods involved in this research include auditory psychophysics and electroencephalography (EEG). The fundamental stimulus across all experiments consists of a harmonic complex structured from multiple pure tones. The motivation for EEG as a viable neuroimaging technique for these studies is that the temporal resolution of this method serves the researchers hypotheses about the dynamics of attention. Specifically, the interaction that the stimuli have with cortical modulation (minimum, 15 Hz or 40 Hz) is rapid enough to warrant a temporally precise technique. These recording will be analyzed via frequency decomposition with the fundamental frequency and secondary harmonics given particular attention condition. Differentiating between evoked responses for each condition will first be subject to a qualitative inspection to explain the directionality of an effect (e.g. which condition has greater power). Bootstrapping will be applied to the frequency representations of ERPs between conditions to conclude the significance of the effect. A within subjects experimental design was used in all experiments, where conditions were presented in blocks of trials to all subjects. For EEG recordings, these blocks were interchanged to counterbalance any time dependent variations in arousal or recording quality.

This dissertation will merge psychophysical and neuroimaging techniques to advance the study of cognitive electrophysiology by strengthening the control between conditional manipulations. Here, EEG is aiding in the study of attention, and the results produced should confirm several theories regarding feature selectivity. Additionally, psychophysical advancements for studying selective and automatic attention are presented. The remainder of the introduction presents the challenges, goals and methods for studying the multiple mechanisms of auditory attention.

1.2 Selective Attention - EEG

The ability to engage in auditory selective attention relies on listeners discriminating between acoustic features that are simultaneously available within a sound and can differ between auditory objects. Quantifying the qualitative state of these acoustic features has been advanced via COSS analysis. This methodology employs identical stimuli across trials, while directing the listener to attend to one of several acoustic features (e.g. loudness, timbre/roughness, or pitch). Serving as an absolute control between conditions, this paradigm is well suited for neuroimaging because condition differences will yield purely endogenous responses. Weight profiles for loudness and timbre judgments were replicated from previous experiments during EEG recordings. While listeners were engaged in making loudness judgments, a characteristic ERP response of the stimulus F0 and secondary harmonic was observed. Timbre ERPs displayed a similar response to F0, but a decrease in activity related to the secondary harmonic.

1.3 Automatic Attention – Psychophysics

A phenomenon referred to as a transition bandwidth was first observed by Berg (2007) in a band-widening, discrimination experiment designed to estimate the bandwidth of sensitivity to a change in the temporal envelope. The signal is an increment in the central tone of a stimulus complex composed of *n* equally intensity tones, evenly spaced with respect to frequency by Δf. Pitch and loudness cues were degraded by randomization techniques, which presumably limited the source of valid information to the envelope of the waveform. As the number of tones was increased, thresholds increased up to a certain critical bandwidth. It was predicted that thresholds would asymptote beyond the critical bandwidth because it exceeded the bandwidth of envelope sensitivity. As the stimulus bandwidth was increased beyond this *breakpoint*, however, many listeners displayed decreases in thresholds by more than 10 dB. The theory is that listeners engaged in a spectral process of profile listening (Green, 1993) beyond the breakpoint, when the bandwidth of the stimulus became broad enough to support across channel comparisons (i.e. the output energy of a channel containing the signal component is greater than the energy from comparison channels.) The breakpoint presumably represents a change in underlying process from one that relies on information from the temporal envelope to one the relies on across-channel comparisons of spectral energy (i.e. detecting a peak in the spectrum).

The current interest in estimated transition bandwidths (ETB) is that the phenomena offers an approach to investigating automatic attention in hearing. Evidence suggest that the switch in the underlying processes is triggered by the bandwidth of the stimulus and is not under the volitional control of the listener or affected by instructions from the experimenter. One might say that ETBs manifest a previously unrecognized form of attention explicitly controlled by characteristics of the external sound. ETBs are extremely robust, are unaffected by changes in Δf ranging from 10 Hz to 400 Hz, and are stable following extended practice (Turner dissertation, 2010). As discussed in the next section, the transition bandwidths offer a potentially valuable paradigm for neuroimaging techniques. The intent here is to refine the psychophysical measurements of ETBs.

An interesting aspect of the ETB paradigm is the range of individual differences. With a stimulus centered at 1000 Hz, ETBs across listeners range from 100 Hz to over 1000 Hz (Berg, 2007).   Rather than a sharp decrease in thresholds beyond the ETB, some listeners display an asymptotic segment beyond the breakpoint, making it difficult to ascertain a clear breakpoint. One contribution of this dissertation is to introduce an enhanced technique for locating the ETB. Previous experiments used a new sample of starting phases for every presentation of the stimulus. It was discovered that by fixing the starting phase of each component to 0, all listeners show an immediate decrease in signal thresholds at the ETB of at least 10 dB, with some showing improvements by as much as 20 dB.

1.4 Automatic Attention – EEG

Few auditory paradigms have the ability to manipulate the characteristic of the stimulus which the listener attends to. The ETB paradigm provides a unique opportunity to study the cortical dynamics of attention associated with the listeners transitioning between envelope processing and profile listening. One challenge that this paradigm does not control for are the individual differences observed when estimating a breakpoint. Because the attentional switch is controlled by external stimulus, identifying the traces in the ERP record should be traceable by the stimulus differences in each condition. Such a change in stimulus characteristics will improve the SNR of the neral activity of interest.

1.5 Broad Hypotheses and Direction

The experiments included in this dissertation demonstrate the dynamics of attention both between and within selective and automatic listening tasks. Chapter 2 applies and develops psychophysical procedures aimed towards quantifying the qualitative states of attending to various acoustic features. EEG supplements psychophysical modeling in ascertaining the mechanism used by the listener and the stimulus characteristics which the decision statistics are derived. Stimuli are then abbreviated to demonstrate the efficacy of psychophysical models. Chapter 3 develops techniques designed to capture the dynamic properties of automatic attention. Varying stimulus characteristics by increasing the bandwidth of the stimulus, changes the reliability of temporal/spectral information, and listener performance supports a respective change in decision strategies.

Chapter 2

Selective Auditory Attention and Endogenous ERPs

2.1 Overview

An innovative approach for investigating feature selective attention was introduced by Southworth and Berg (1995). Their discrimination experiment used a narrowband stimulus in which the standard consisted of three equal-intensity tones and the signal was an increment in the intensity of the central component. An interesting property of the stimuli that is exploited here is that the signal can be detected by either a change in pitch, loudness, or roughness. Southworth and Berg demonstrated that the preferred acoustic feature used by a listener can be determined by estimating a set of decision weights that summarize the relative contribution of each tone to a listener's trial-by-trial decision (Berg, 1989, 1990). Presumably, attending to different features produces different patterns of decision weights. Figure 2.1 shows decision weights obtained from simulations of various models that calculate a decision variable based on either pitch, loudness, or roughness (Feth 1974, Green, et al 1992; , Berg, Nguyen and Green, 1989). As shown by Southworth and Berg, listeners produce similar patterns of weights when instructed to attend exclusively to each of the corresponding features.

Of critical importance is the ability to measure the degree of compliance to instructions to attend exclusively to a particular feature. This is done by comparing weight estimates for a listener to the weights obtained from simulations of the corresponding model. The paradigm offers a unique control technique for neuroimaging applications because the stimulus remains the same while the underlying discrimination process presumably changes with instructions. Control of the stimulus across condition is essentially perfect. Any differences in the EEG signature for pitch discriminations and loudness discriminations, for instance, are attributable entirely to endogenous sources because the exogenous sources arising from transduction of the stimulus are presumed to be constant.

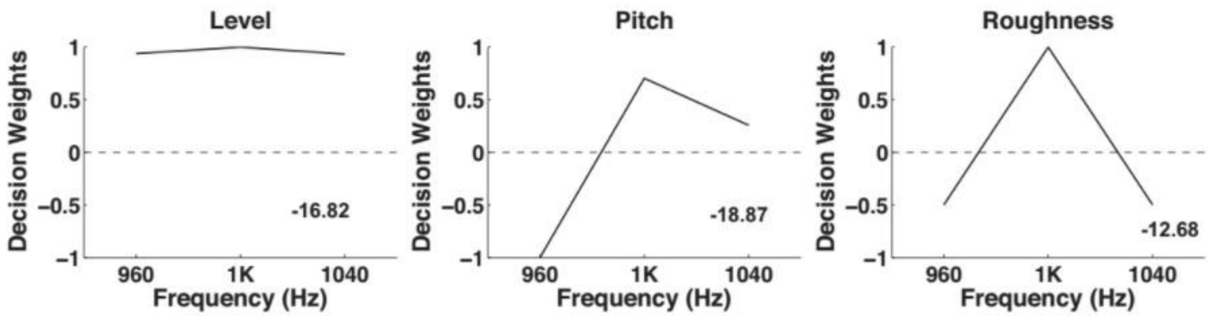

Figure 2.1: Acoustic cue, model predictions. Simulations adopted from Tan, 2015.

The estimation of relative weights has allowed for the attainment of spectral weights in a narrow-band task that produces distinct weight profiles for separate acoustic features (Southworth 1995). Three auditory features were observed, and all of which can be adequately explained by their respective models. This application advances psychophysics to study higher order cognitive processes that are employed when multiple sources of information are available. For example, instructing a listener to perform a specific strategy when engaged in a task is subject to the bias of the interpretation by both the subject and the experimenter. Deriving weights on the multiple observations available, from the listener's responses ascertains the internal representations of the listener in a quantifiable fashion.

This study applies the weighting procedure to a three tone experiment where subjects are instructed to attend to distinct acoustic features of loudness and timbre, identical to Southworth and Berg (1995). The opportunity provided by the spectral weighting procedure allows for a constant stimulus across conditions, where the conditional manipulation is the instruction given to the subjects. Such a control allows isolating scalp activity specific to qualitative states of attention because the stimulus variance is limited to with conditions, instead of between conditions. The motivation to apply a neuroimaging technique is largely exploratory, but also offers the unique opportunity to test the hypothesis of weighted modulation channels.

2.2 Introduction

Auditory stimuli are composed of several features that are derived from the various physical characteristics of a stimulus. These features are fundamental to a listener's ability to attend to a sound when the features convey task relevant information or when that sound is in competition with other auditory objects. Communicating the quality of these features can be highly subjective, and the study of which requires objectification of the relevant characteristics. In turn, this provides an understanding of how sound is processed. Auditory models have been developed to quantify subjective classes of auditory attention (e.g. loudness, timbre, and pitch) and provide an understanding of how they are influenced by the stimulus characteristics (Feth 1974, Green 1983, Berg 1989).

The models that quantify subjective listening strategies provide an opportunity to electrophysiology to further the study of the neural correlates of consciousness. Recent developments in psychophysical techniques offer a unique experimental control that provides both a constant stimulus, and a method to parse distinct processes. Combining these psychophysical procedures with neuroimaging allows researchers to ascertain the source of condition dependent differences in scalp activity. This experiment employs a method developed by Berg 1989, that quantifies the effect that stimulus perturbations have on responses. These methods are supported by previously developed psychophysical models that aid the understanding of auditory features, such as loudness, timbre and pitch (Feth, 1974; Green et al., 1992).

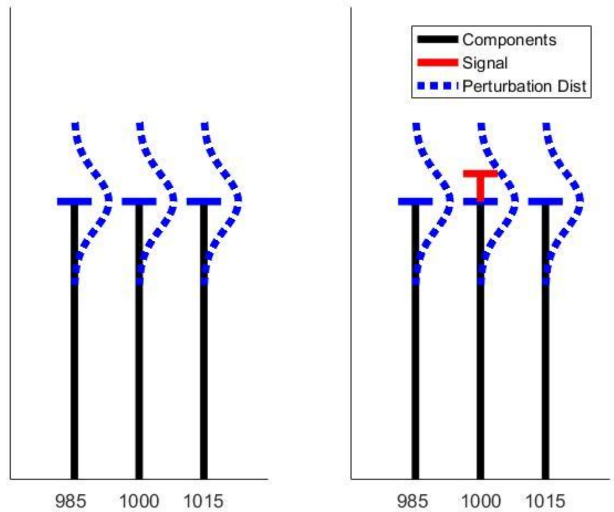

Figure 2.2: Two Interval, Three-tone Stimulus. Sinusoids are separated by 15 Hz, and centered around 1000 Hz. A signal is randomly assigned to interval 2 above (right panel). Perturbations are drawn from a normal distribution, with $\mu = 0$ dB and $\sigma = 1$ dB.

Recently, the efficacies of models that explain auditory features have been verified with psychophysical methods via hypothesis testing. Tan and Berg (2017) demonstrated that normalizing stimulus intensity and employing a frequency rove, in a forced alternative task, deprives listeners of loudness and pitch information, thus producing psychophysical weights characteristic of roughness models. In a similar task, Bellato and Berg (Appendix A) abbreviated stimulus duration to 30 milliseconds. The psychophysical weights produced by listeners suggested a strong dependence on stimulus intensity, supporting predictions provided by loudness models.

This experiment employs a narrowband complex tone, which allows for the acquisition of spectral weights. Coupling these psychophysical techniques with electroencephalography allows the opportunity to ascertain the feature of sound that the listener was attending to. Subjects will be instructed to attend to attend to a difference in loudness or a difference in roughness. Provided that pre-frontal cortex activity is associated with top-down attention, it can be hypothesized that this activity will onset earlier in the loudness condition than the roughness condition. The justification for this is that a roughness judgment requires the acquisition of sustained envelop information. Meanwhile, loudness judgments can be made within the first 50 milliseconds of stimulus onset. Because stimulus modulation is a critical characteristic of roughness, it is expected that the auditory steady state response (ASSR) of the cortex will be stronger and more sustained when listeners are making decisions related to stimulus roughness.

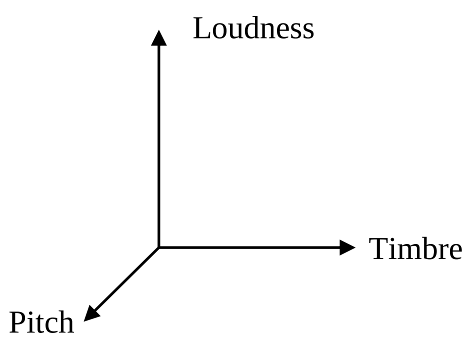

Figure 2.3: Multi-dimensional space of perceptual acoustic features. Acoustical features are conceptualized as the internal representations of the listener.

The profile of weights that are obtained from instructing listeners to attend to a single cue throughout each condition can be predicted and/or explained by different auditory models. Zwicker proposed a model for loudness that sums the energy across frequencies (Zwicker, 1957). The level model that predicts the weights obtained in the loudness condition sums the energy from each component of the stimulus, and predicts that the listener will respond signal to interval one when:

$$\sum_{i=1}^{3}(x_{i1}+L_1)^2 > \sum_{i=1}^{3}(x_{i2}+L_2)^2 \qquad \text{Equation 2.1}$$

where x_i is the respective set of perturbations for each interval and the *L* variables represent the overall level of the tone. The model that best explains the weight profile attained from instructing the listener to attend to roughness computes the Fourier coefficients from the Hilbert envelope of the stimulus. This model is developed by Green *et al.* (1992) and follows as:

$$\sum_{i=2}^{m}(\eta_i - \tau_i)z_{i1} > \sum_{i=2}^{m}(\eta_i - \tau_i)z_{i2} \qquad \text{Equation 2.2}$$

where the observation vectors of ηi and τi represent the Fourier coefficients computed from the Hilbert envelope. The *z* coeffieient is a weighting variable for the normalization of the envelope.

2.3 Methods

2.3.1 Participants

Participants were students and researchers at the University of California, Irvine and ranged in age 19 to 29. Prior to beginning data collection, all listeners were screened to be normal listeners, as defined by thresholds equal to or less than 20 dB SPL for the frequency regions to be presented.

2.3.2 Materials

Stimuli were constructed and presented via Matlab. Complex tones were constructed in the spectral domain, so that amplitude perturbations could be added to the stimulus with more computational efficiency. Sounds were presented at a sampling rate of 44100 Hz at 70 dB SPL for a duration of 800 milliseconds (ms). The frequency of each complex was stationary, with a center frequency of 1 kHz and two side bands at 15 Hz separation (ΔF). Non-signal amplitude perturbations were added to every component, and were randomly sampled for each trial from normal distribution with a mean of 0 dB and a standard deviation of 1 dB (N(0,1), Figure 2.1). The perturbations were added in phase to each component, with all phases fixed at Θ = 0.

Stimuli were presented in blocks of 70 trials where each trial consisted of a two interval forced choice task. The signal was an amplitude increment to the central component at 1000 Hz. Listeners were directed to identify the signal by responding 1 or 2 on a keyboard, corresponding to the interval which they deemed the signal present.

The signal for each trail was randomly ($p = 0.5$) added to either the first or second interval. A two-down one-up procedure was used to estimate thresholds, where amplitude of the signal was averaged across the last 3 reversals (Levitt, 1971).

2.3.3 Conditions

The primary conditional manipulation was the instruction given to the participants preceding each block and/or set of blocks. Each condition allows for the listener to identify the signal interval along the dimension of the respective acoustic feature. Loudness judgments are made by identifying the interval that is louder (equation 2.1). This coincides with the signal across trials because provided that the perturbations equal zero, on average, the signal interval will have greater intensity. The roughness condition requires that the listener identify the signal as the interval which is less rough. The increment in amplitude of the signal component creates the sensation of a smoother sound. The decision variable that predicts this strategy is expressed in equation 2.2. Identical weighting procedures are applied to data from each condition. Differences between conditions indicate the change in listening strategy for each listener.

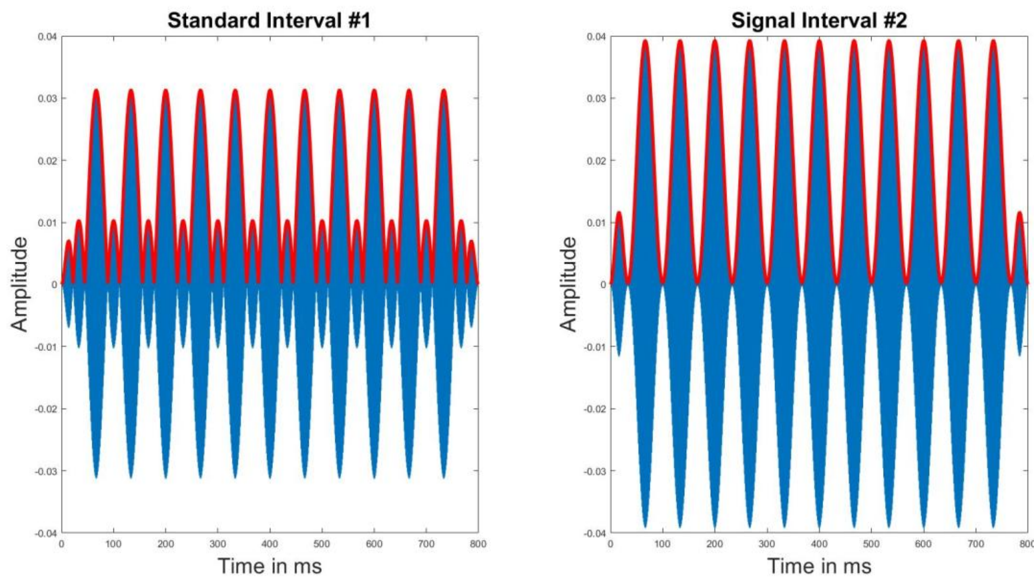

Figure 2.4: The time domain representation of the first and second interval. The signal is in the second interval. The Hilbert envelope is plotted in red. Time is in milliseconds.

2.3.4 EEG Methods

EEG signals were recorded from the Compumedics Neuroscan 128 channel Quik Cap, sampled at 1000 Hz with a 100 Hz low-pass filter. Bad channels were adjusted during intermittent breaks for subjects to rest. Rejected channels were not used during the analysis of individual channels and their values over time were interpolated (Perrin et al., 1989) as a weighted sum when computing scalp topographies.

2.4 Results

2.4.1 Psychophysical Results

Psychophysical weight profiles were derived from COSS analysis by calculating the probability of the identification of a signal as a function of the amplitude perturbation for each component (Berg 1989). Figure 2.5 demonstrates the relationship between the perturbation level, the probability of the listener responding “signal” and the summary profile of weights. The probabilities are attained by calculating 20 equal area bins across the normally distributed perturbations for each component, and each interval. The ratio of “signal” responses for all perturbations within the range of each bin serves as each probability point. Probabilities across perturbations are fit to the minimum mean squared error (MSE) cumulative Gaussian for the respective variance and mean (lower 3 panels for each subject/condition in Figure 2.5). This procedure is not dissimilar to estimating the best fit psychometric function. The weights (upper panel) for each component are derived from taking the ratio of estimated variances.

Consistent with findings from Southworth and Berg (1995) the profile of weights is similar to model predictions for each condition. However, individual differences between subjects are observed. The indication that each individual subject changed their decision strategy is the relative difference of the psychometric function of the side components, between conditions. Specifically this can be determined by the sign change in the psychometric function which the weights for each component are derived from. The central component does not exhibit a sign change, but the side components change from a positive sign to a negative sign between the loudness and roughness conditions respectively. The change in sign indicates a change in the cognitive process employed for processing the perturbation of that component.

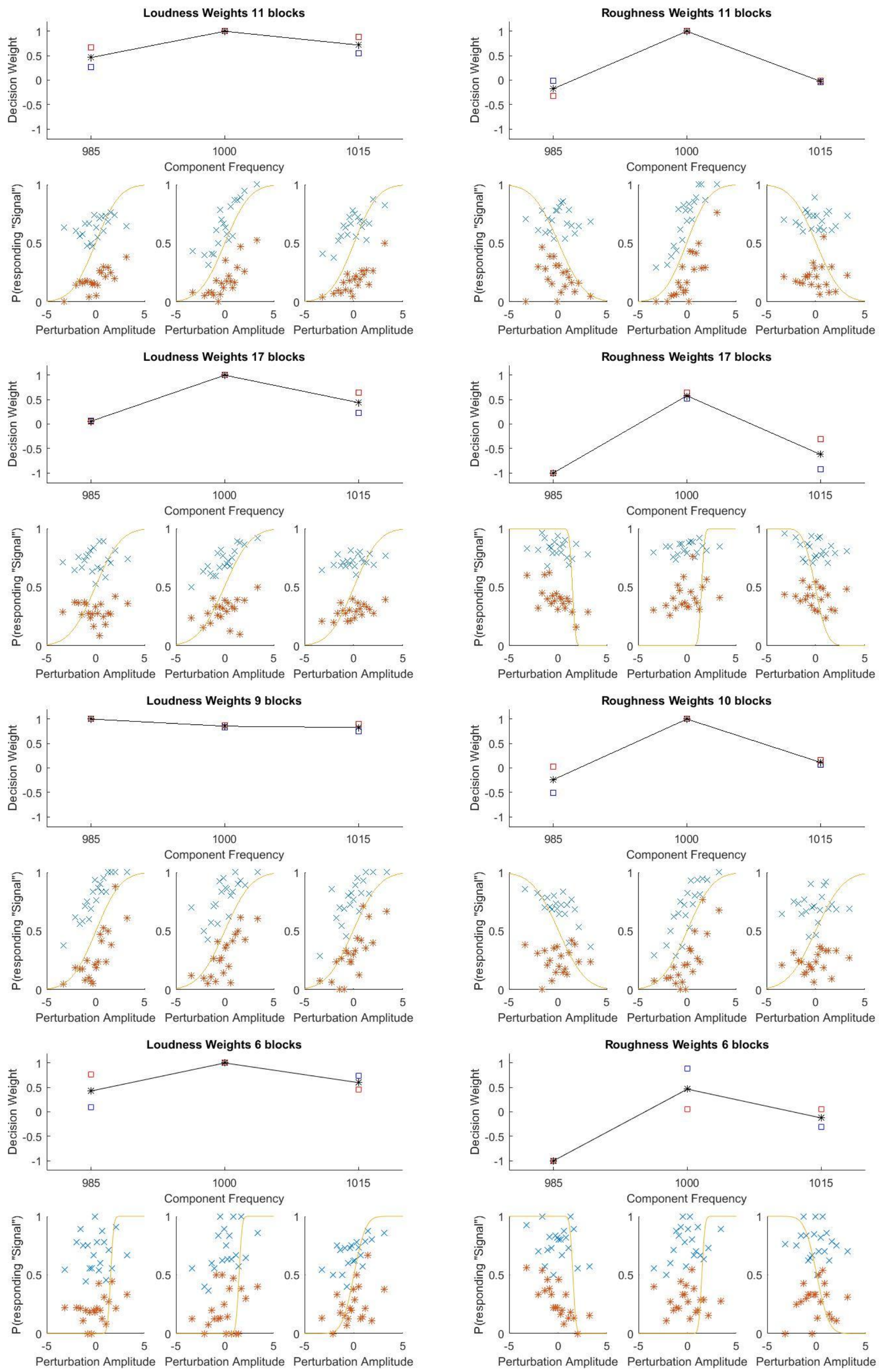

Figure 2.5: Loudness and Roughness conditions. Psychophysical weights for 3-tone complex. The left column is loudness weight profiles and the right is roughness profiles. Each row of weights and variance is a subject.

Individual differences and variations from model predictions are a consistently observed in these feature attending tasks, including in the original report of this paradigm in Southworth and Berg, 1995. The explanation of these differences is the subject of ongoing study, and is can be understood by a hybrid model, a mixed model, or variants in the weights assigned to the variables in equation 2.2. A hybrid model is proving to be the most difficult to develop, and would interpret the deviations from model predictions as the product of the acoustics features being correlated. For example, the hybrid model would propose that listeners may hear changes in roughness to also change in loudness. The mixed model approach to these variations provides that throughout the duration of data collection, a listener might alternate the calculation of their decision variable between the feature of instruction and other features. Last, the recent development of a weighting procedure for the coefficients from equation 2.2 assigns independent weights to F0 and 2F0. Simulations suggest that changes in the weights assigned to the fundamental frequency and other harmonics produce a change in the magnitude of the weights for the side components. These simulations only apply to the roughness model since that model relies on information from the fundamental.

Figure 2.6 demonstrates two listeners who were unable to consistently produce different weight profiles between the two conditions. The subject in the bottom row does not appear to produce a difference in weight profiles between the two conditions. In the upper panel, the subject changes the slope of the psychometric function for the lowest frequency sideband, but fails to do so for the high frequency sideband. Because the high frequency side band does not change sign, or even approach zero, this subject was classified as unable to perform separate listening strategies.

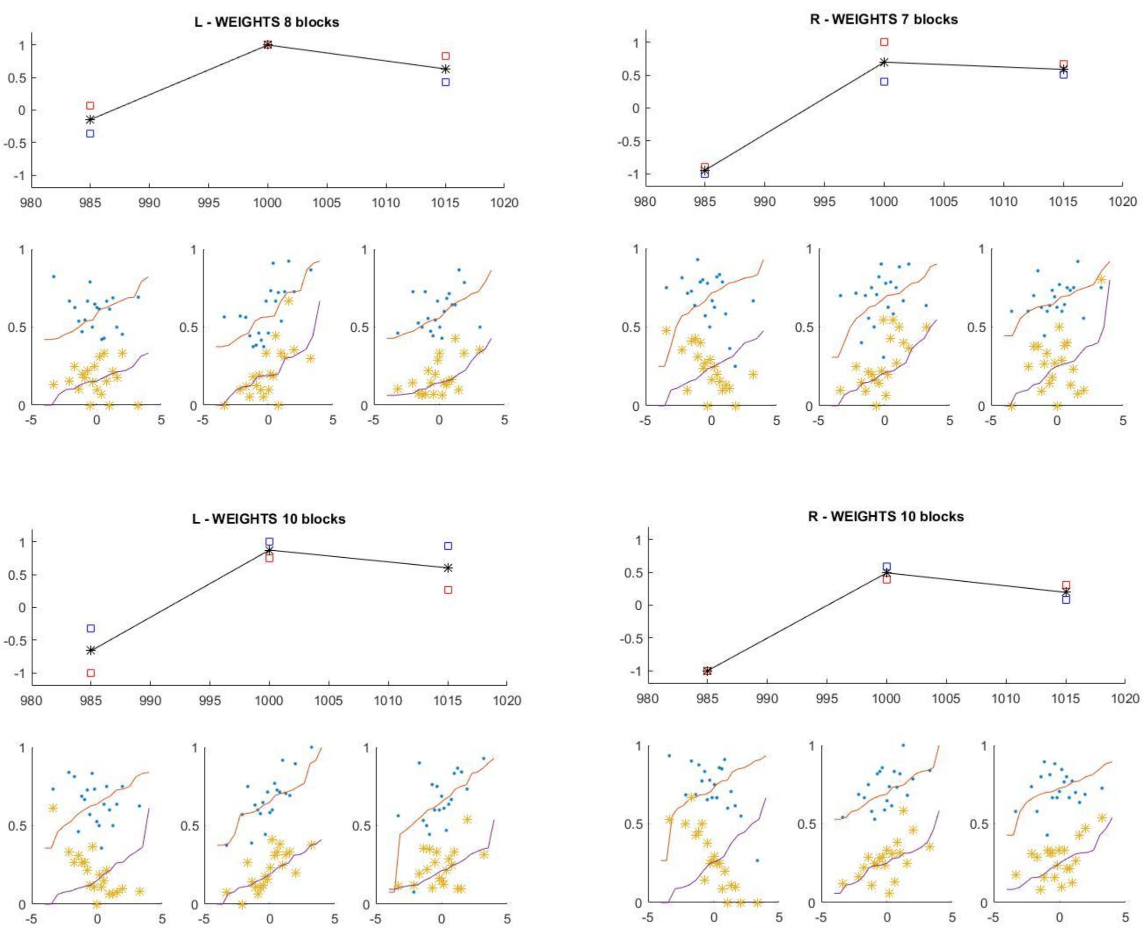

Figure 2.6: Subjects weights dissimilar than model predictions.

2.4.2 EEG Results

ERPs of a representative subject are plotted in figures 2.7 and 2.8 for loudness and roughness conditions, respectively. The upper plot in each figure corresponds to the ERP recorded from 122 electrodes. Lower panels of Figures 2.7 and 2.8 plot the variance across all electrodes over time, representing the topographical variance preceding and throughout the duration of the stimulus. Figures 2.7 and 2.8 were adopted from Cohen, 2015.

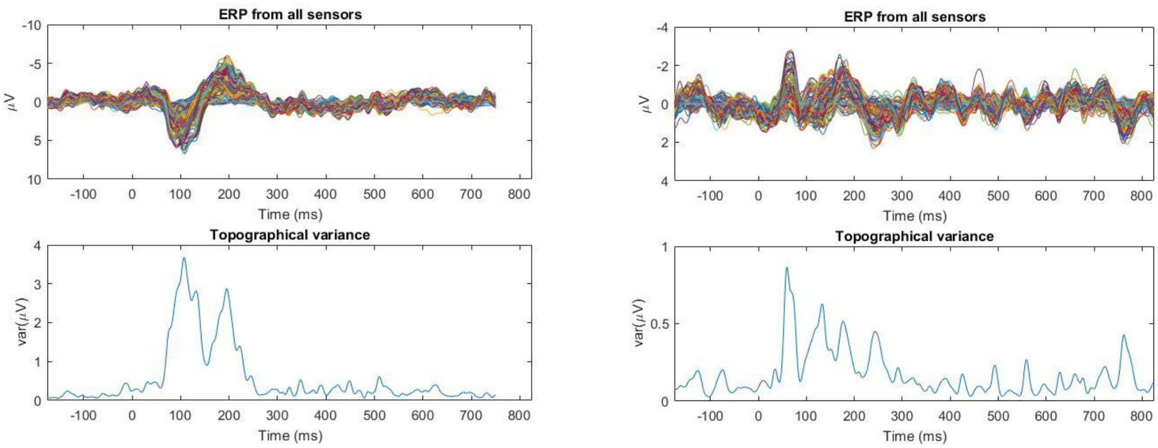

Figure 2.7: Loudness ERP. Grand ERP and topographical variance. ERP responses from all electrodes are plotted by μV across time. The variance across all electrodes is estimated for each time point.

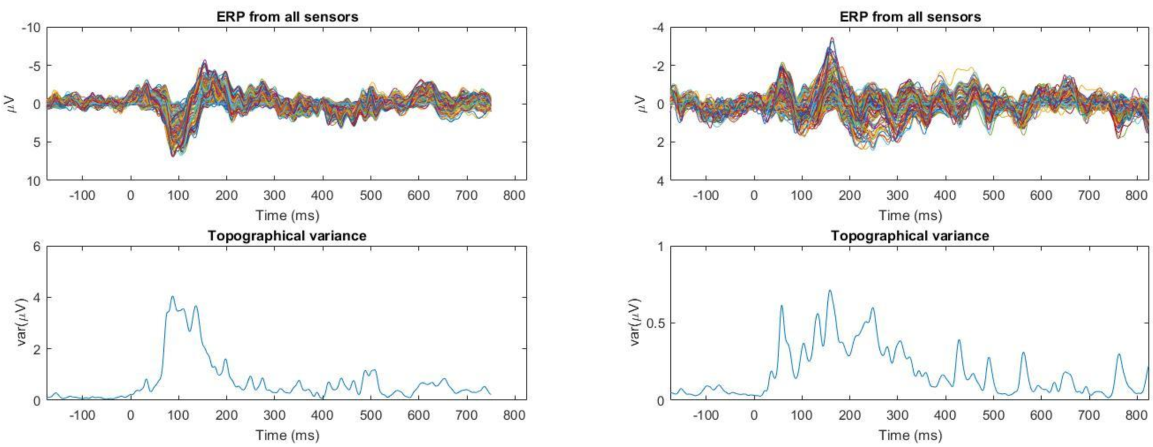

Figure 2.8: Roughness ERP. Grand ERP and topographical variance. ERP responses from all electrodes are plotted by μV across time. The variance across all electrodes is estimated for each time point.

The fast Fourier transform (fft) was applied to each electrode to recover the power of the ASSR associated with the stimulus envelope (Figure 2.9). Electrode data transformed by the fft was limited from 300 ms after stimulus onset to 800 ms (stimulus offset) to better recover the signal to noise ratio of the envelope. This method provides average frequency following information across the duration of the stimulus, but does not provide insight to the temporal dynamics of the ASSR.

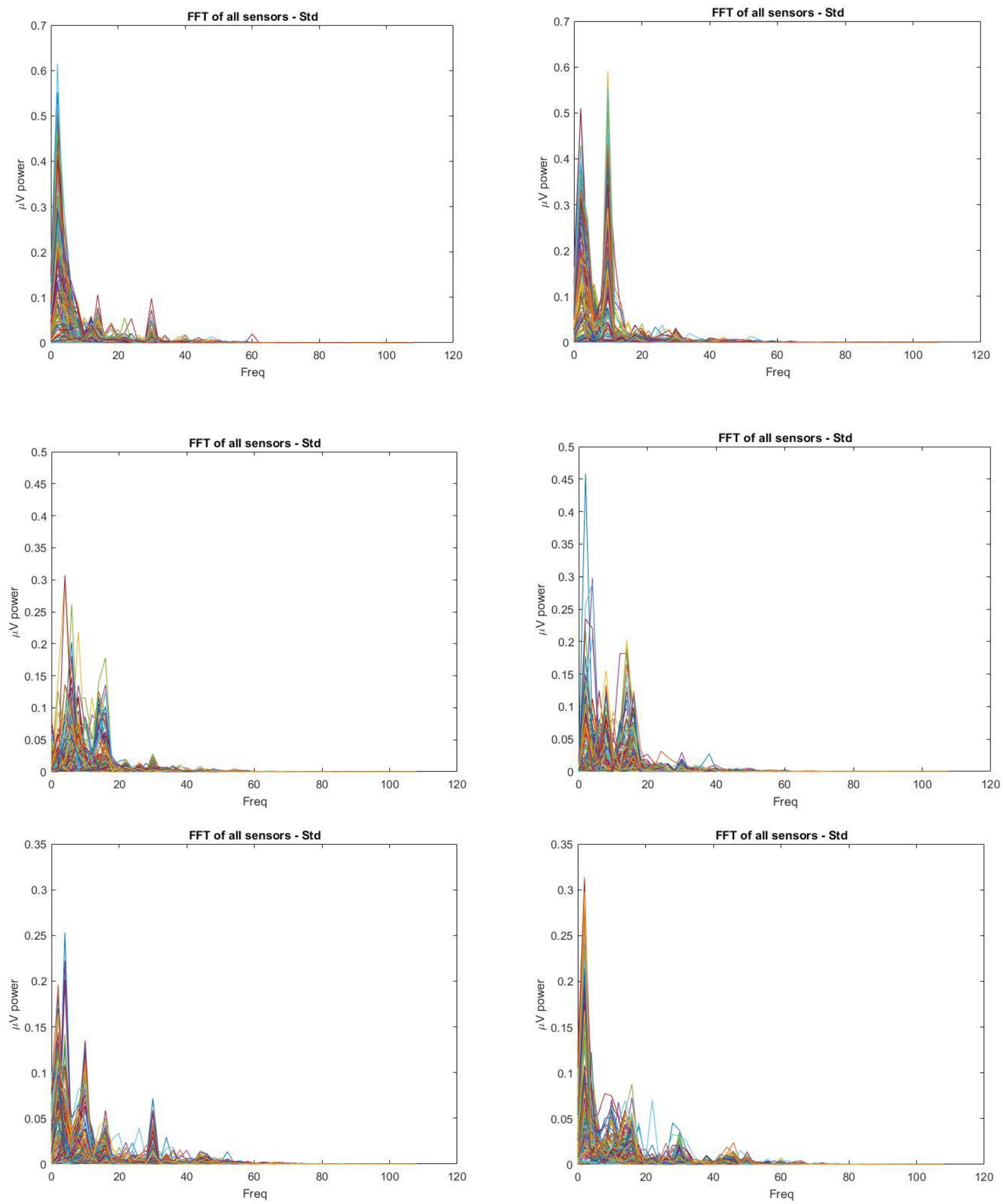

Figure 2.9: ERP Spectrum. The FFT of each ERP was calculated for both conditions from 300-800 ms. Subjects 1-3 are shown here.

The primary analysis considered the FFR, specifically examining the activity at 15 Hz and 30 Hz. The most salient effect of conditions across subjects was the difference in 30 Hz activity, which coincides with 2F0 of the stimulus envelope. To gain a clear understanding of the distribution of 30 Hz activity across electrodes, the 30 Hz coefficients from individual trials were subject to bootstrap re-sampling. By re-sampling the data with the bootstrap method, the influence that outlier trials have on the average 30 Hz activity is minimized. The distribution of the original 30 Hz power of all electrodes is plotted in Figure 2.10. Figure 2.11 shows the bootstrap transformed distribution for each condition. The null hypothesis that the 30 Hz power was rejected for every subject that displayed different weights $t(99)= 27.8, p<0.01$, $t(99)= 20.44, p<0.01$, $t(99)= 6.17, p<0.01$ and $t(99)= 8.3, p<0.01$.

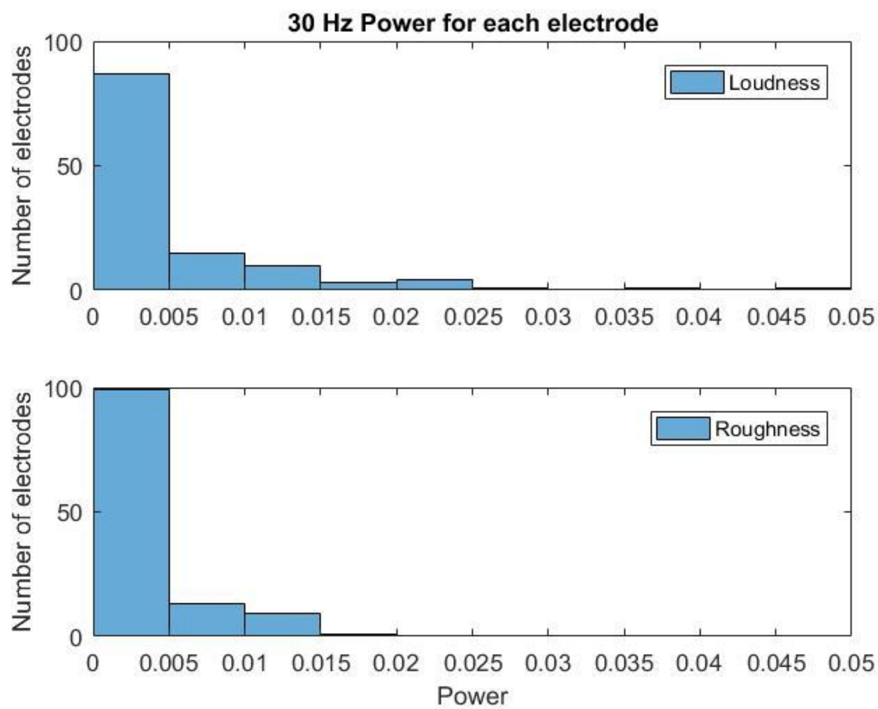

Figure 2.10: Pre-Bootstrap Distribution of 30 Hz Power.

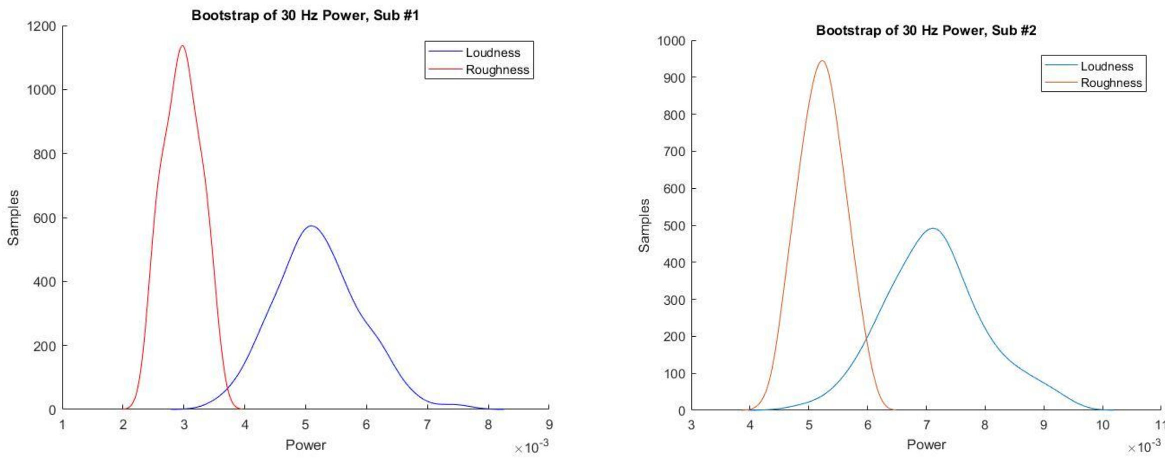

Figure 2.11: Bootstrap re-sampled 30 Hz distributions. Each distribution is attained from the full cap activity, across all electrodes from one participant. T-test determined significance at the 0.01 alpha level.

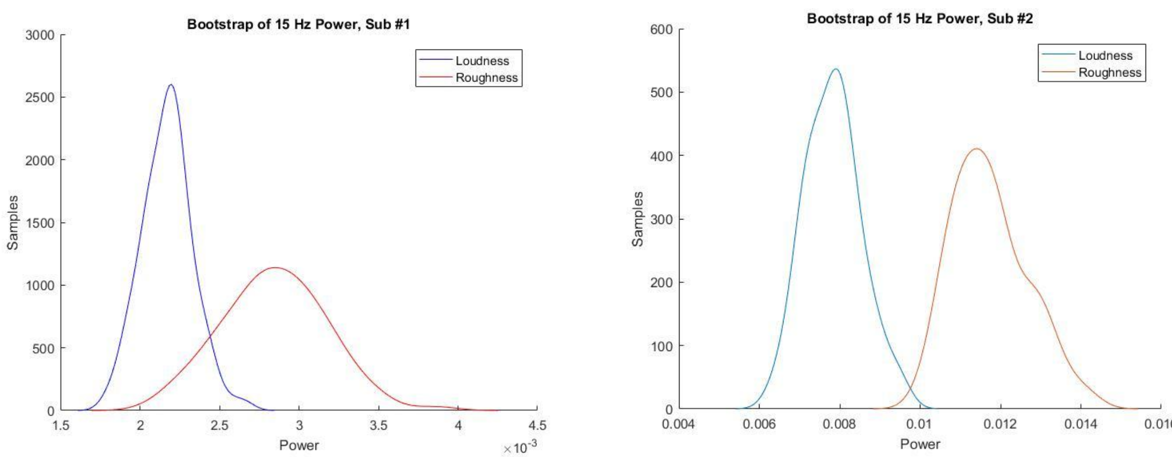

Figure 2.12: Bootstrap of 15 Hz Fourier Coefficients.

The 15 Hz Fourier coefficients were also bootstrap re-sampled across the mean electrode activity. Results from a t-test between the two distribution for Subject #3 failed to reject the null hypothesis that the two distributions were not different, $t(99)=1.6$, $p = 0.11$. It is noted that Subject #3 provided a profile of weights for the roughness condition that were only slightly negative.

2.5 Discussion

Psychophysical weighting procedures were applied to the three-tone complex for each condition's behavioral data: loudness and roughness. The model predictions for these two conditions are generally consistent with the respective profile of weights obtained. Individual differences between weight profiles are observed, and recent psychophysical developments are discussed in section Appendix B. The characteristic profile for each of condition/model is supported by the sign (positive or negative). Weights attained from the loudness condition are typically all positive in sign, with roughly equal weight being given to each component. Roughness weights are characterized by a large positive sign on the signal component and negative sign weights on the non-signal components.

The ASSR of the fundamental frequency appeared greater in power for the roughness condition than the loudness condition. Unexpectedly, a global difference between conditions emerged at the secondary fundamental frequency of twice the frequency of Δf. Subjects' ERP spectra revealed a consistent difference at 30 Hz both across subjects and between conditions. Based on both the hearing literature and the abbreviated duration experiments included in section 2.2, it could be expected that the onset of F0 or 2F0 activity would differ. However, this trend is not observed. Relative to judgments calculated from a decision statistic based on the envelope of a sound, loudness decision variables can be computed much earlier. The temporal characteristics of a stimulus are the primary statistic in the computation of a roughness decision variable. Abbreviated stimuli deprive the listener of this information and effect performance on temporal tasks, while listeners maintain performance on spectral tasks (Borucki & Berg, 2017).

Chapter 3

Automatic Attention: Psychophysical Enhancement of ETB

3.1 Introduction

Varying the parameters of complex tones can provide rich information about the different mechanisms that are employed by the human auditory system. In his ETB, band-widening experiment, Berg (2007) provided evidence that listeners automatically attend to either the temporal information or spectral structure. It appears that the ’attentional switch’ is the stimulus bandwidth. Other instances of a switch in the dominant process that is dependent on stimulus bandwidth exist, such as a change from within-channel to between-channel cues in comodulation masking release (CMR, Hall et al, 1984), but the change from a temporal process to a spectral process with the ETB paradigm is especially distinct. Empirical support that ETBs manifest a change in the dominant process from envelope discrimination to profile analysis is provided by both threshold functions and an analysis of decision weights (Berg, 2013). A particular strength of the paradigm is that both the envelope discrimination and profile analysis are well-developed quantitative theories. Many empirical and quantitative tools have been developed that may be useful in a search for neuro-correlates to auditory automatic attention.

New evidence for two distinct processes is presented here by showing that each is affected differently by manipulating the starting phases of the stimulus complex. ETBs are estimated under two conditions, one for which the starting phases of stimulus components are randomly sampled for every stimulus presentation and the other with the starting phase fixed at zero for all components. Thresholds for stimulus bandwidths less than the ETB are unaffected, whereas thresholds for bandwidths greater than an ETB improve by as much as 20 dB with a fixed-phase. An important consequence of this discovery is that a fixed phase stimulus enhances the precision in identifying the ETB for a particular listener. For the most part, the breakpoint in the threshold function is clear and unambiguous. This improvement in the paradigm will prove critical because an unambiguous estimate of the transition bandwidth is necessary for designing the EEG experiments reported in Chapter 4.

3.2 Methods

3.2.1 Listeners

Listeners were paid subjects ranging in age from 19-24, who were recruited via posted flyers on campus at the University California, Irvine. All listeners were subject to a hearing screening test to confirm that they had normal listening capabilities for frequency ranges of experimental stimuli.

3.2.2 Materials

Parameters for stimuli were adopted from previous ETB experiments. Single channel energy cues were degraded with a 20 dB level rove centered around 65 dB. Pitch cues were degraded with a 100 Hz frequency rove centered around 1000 Hz. Stimuli were presented for a duration of 500 ms and were onset/offset by a 5 ms cosine ramp. The number of spectral components varied across conditions, and consisted of an odd number of components between 3 and 25, linearly spaced with 40 Hz separation (Δf). The minimum, median, and maximum n is demonstrated in figure 3.1.

Stimuli were generated with Matlab R2008a running on a PC with Windows 7. The wave-forms were played through a two channel D/A converter (0202 USB 2.0 Audio Interface; E-MU Systems) at a 44.1 kHz sampling rate. These were passed through a manual attenuator for calibration. A TDT System II headphone buffer split the signals to both channels of the headphones. The sounds were delivered through Sennheisser HD414SL headphones. The subject was seated in a single-walled sound attenuating chamber (IAC). Feedback was presented on a computer monitor and responses collected with a standard computer keyboard. Subjects' responses were recorded with a keyboard, where they entered either a 1 or 2, corresponding with which of the two intervals consisted of the signal. Listeners had unrestricted response time.

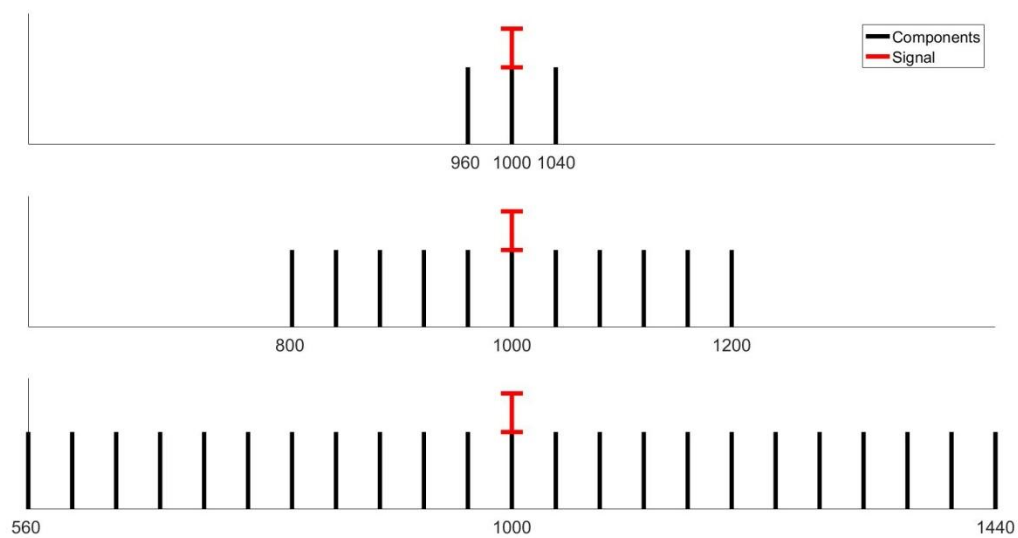

Figure 3.1: Bandwidth Manipulation. The components included in the minimum, median, and maximum bandwidths are plotted above. A threshold dependent signal was added to the center component. The center frequency of each complex was randomly assigned according to a uniform distribution between 950 Hz to 1050 Hz. Stimuli above represent the average center frequency at 1000 Hz.

3.2.3 Conditions

Listeners were instructed to select the signal interval in a two interval, forced choice task by entering 1 or 2 into a keyboard. Signals were an increase in amplitude of the middle spectral component, and were added in phase to that component. The signal is represented as $20*\log(a/\Delta a)$ where a is the amplitude of the standard and Δa is the change in intensity.

Thresholds were estimated for each block via a 2-down, 1-up procedure (Levitt 1971), which consisted of 50 trials and lasted approximately 2 minutes.

The primary experimental manipulation in this experiment concerned the phase of the individual components. Previous work applied a random phase to each of the individual components, while this experiment held all phases constant at zero.

An additional experiment was conducted to demonstrate the robustness of the transition bandwidth, independent of the number of components in the complex. By varying Δf, it is possible to manipulate the bandwidth of the overall stimulus of a complex with a constant n, or hold the bandwidth of the complex constant while also varying n. This experiment increased Δf to 80 Hz and reduced the number of components so that thresholds could be estimated for every other bandwidth from the original conditions. This method ascertains that the relationship between thresholds and components is dominated by the overall bandwidth of the stimulus.

3.3 Results

Thresholds are shown in Figure 3.2 as a function of bandwidth. Replications of the original experiment are represented by the solid symbols. All four listeners display asymptotic thresholds at wide bandwidths that lead to much uncertainty in determining the ETB. Data from the fixed phase condition are represented by open symbols. Removing the phase randomization and fixing phases at zero resulted in a dramatic decrease in thresholds. This change makes the ETB more apparent.

Error bars depict the standard error measurements across blocks. Most subjects do not demonstrate a difference in narrow bandwidth phase conditions; however error bars do demonstrate a large difference between estimated thresholds in wide bandwidth conditions. The relatively small error bars suggest accurate estimates of thresholds. The error bars associated with improved task performance tend to be greater than those with higher thresholds. This is likely due to threshold estimates being more susceptible to internal noise, but estimates are far enough apart to not reflect analogous processes.

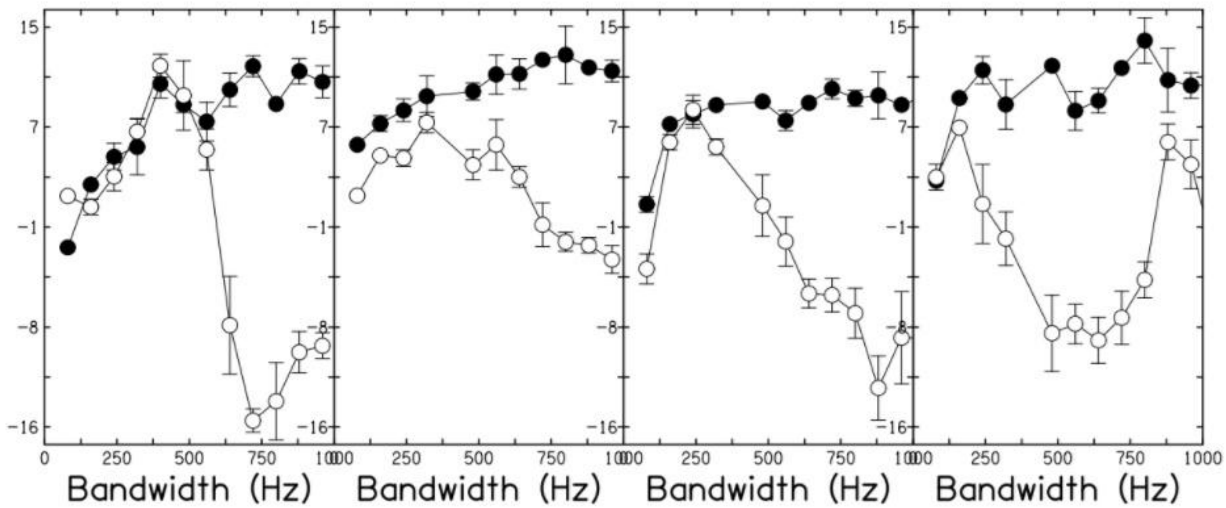

Figure 3.2: Thresholds calculated across stimulus bandwidth. 4 individual subjects are plotted. Each point represents the average threshold across blocks, for the respective number of components and phase manipulation. Filled circles are random phases, and open circles are fixed at zero phase. Error bars depict the standard errors.

The effect between the randomized phase and the fixed phase ranges from 10-20 dB less than the respective. However, subject 4 demonstrates a very large increase in thresholds for the three largest *n* conditions that is not currently understood.

A subset of subjects was selected to test the effect of the bandwidth of the stimulus by decreasing and increasing the density of the complex via frequency separation (Δf). Data are shown in Figures 3.2 and 3.3. For lower density conditions, thresholds were estimated at every other bandwidth of the original stimulus by doubling Δf (Figure 3). Because the breakpoint of the subject in Figure 3.2 is relatively narrow, thresholds were estimated at and above the breakpoint, but not below the breakpoint for the 80 Hz Δf condition (this condition).

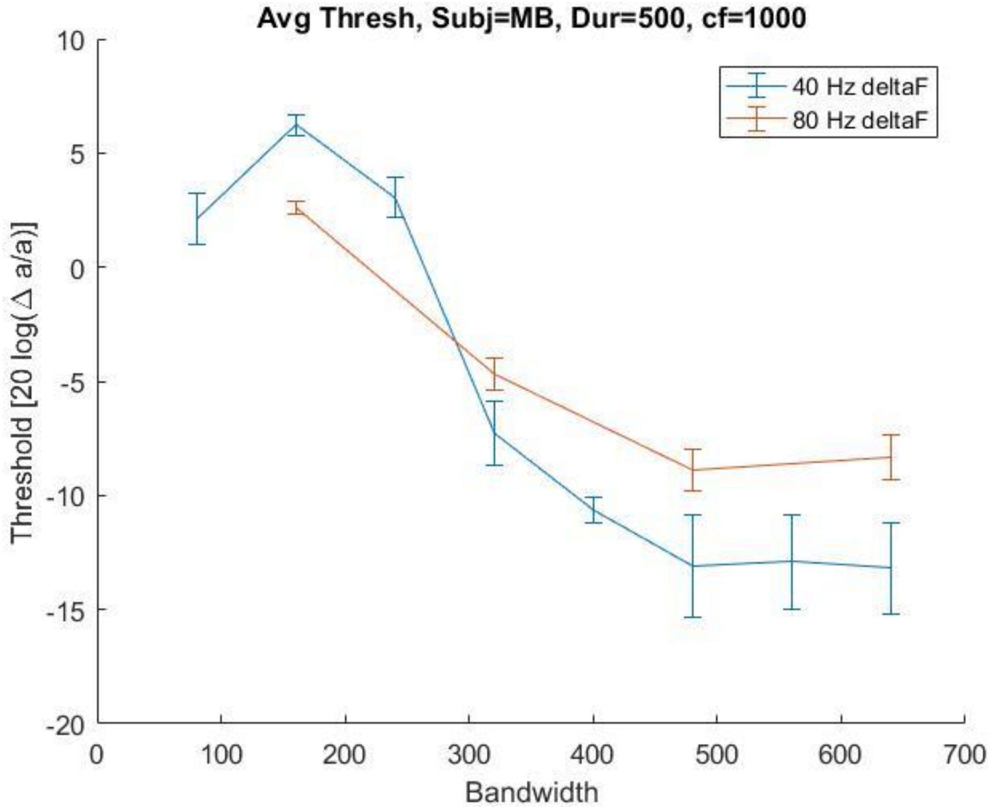

Figure 3.3: Constant bandwidth while increasing Δf. The bandwidth of each condition was held constant, but constructed by a lower number of components by increasing Δf from the baseline condition.

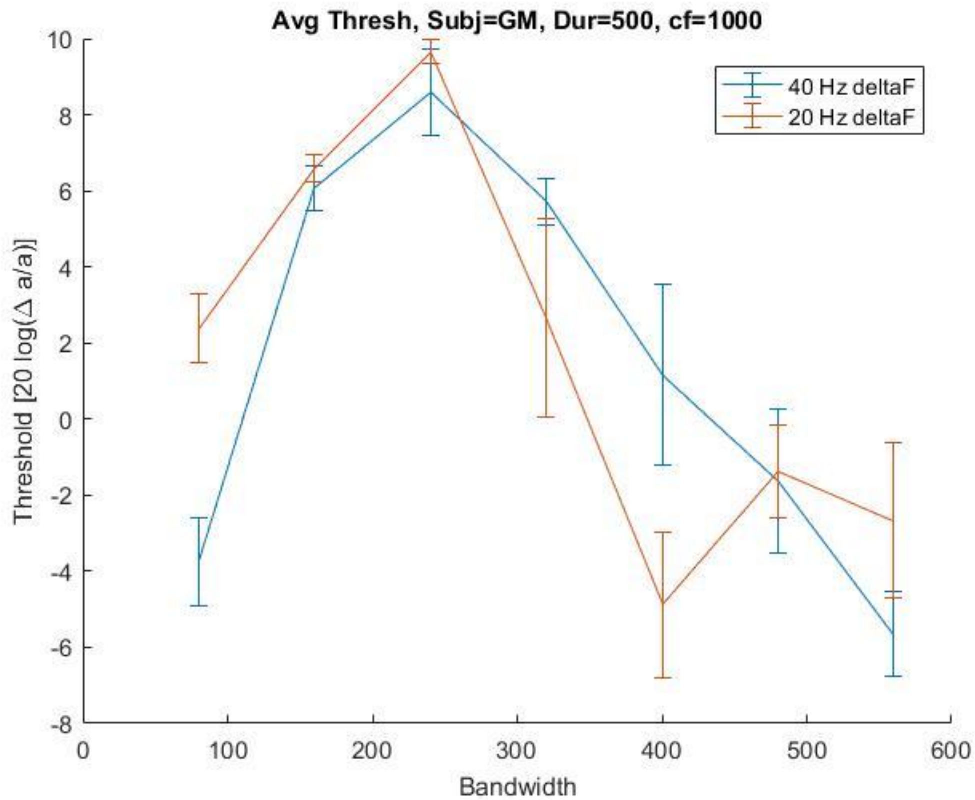

Figure 3.4: Constant bandwidth while decreasing Δf. The bandwidth of each condition was held constant, but constructed by a greater number of components by decreasing Δf from the baseline condition.

Thresholds were also estimated for a high density condition where Δf was decreased, thus requiring more components to construct equal bandwidths. The breakpoint between conditions remained constant, as shown in Figure 3.4. One noticeable difference between the two conditions is the decrease in thresholds for the narrowest bandwidth. Speculation regarding this difference is explored in the discussion below.

3.4 Discussion

Most subjects demonstrate the robustness of the ETB across conditions. The decrease in thresholds past the transition bandwidth in the randomized phase condition is accompanied by a dramatic decrease in thresholds for the fixed phase condition. This trend appears to be robust and dependent on the bandwidth of the stimulus as variations in Δf do not erode the effect for greater bandwidths (Figure 3.4 & Figure 3.5). While the general trend continues, there is however a minor change in the thresholds for conditions composed of smaller *n* and greater Δf for the listener in Figure 3.4. The magnitude of the effect is reduced compared to identical bandwidths in the 40 Hz Δf condition. It is noted, that the difference in thresholds is minor, as some conditions exhibit overlapping error bars. This minor change may be caused by the reduction in the number of components. For smaller and larger *n*, the thresholds can still be explained to have a dependency on the signal to noise ratio. It is likely that the smallest *n* condition produced a threshold lower than baseline (40 Hz Δf) because the 80 Hz Δf condition consists of fewer non-signal components. Thus, downward trends in thresholds for increasing *n* conditions are likely due to an increase in the signal to noise ratio. For these larger bandwidths, the decrease in thresholds is reduced in magnitude because the number of non-signal components of the profile is reduced. The increase in *n* past the breakpoint continues to improve performance because comparisons are being made to the non-signal average. This model was originally proposed by Green and is formally represented by:

$$X_s - \frac{\sum X_i}{n} \qquad \text{Equation 3.2}$$

where X_s is the signal component, X_i are non-signal components, and *n* is the number of non-signal components. This model compares the signal to the average level of the standard.

Profile analysis is considered to be a spectral process which analyzes the shape of the power spectrum, presumably calculated from the energy output from each filter of a peripheral auditory filter bank. It is thus counter-intuitive that phase manipulations would have an effect for bandwidths greater than the ETB and not for narrower bandwidths when a temporal process is dominant. Moreover, the results are inconsistent with the absence phase effects profile analysis experiments (Green, 1983). One difference is that Green and his colleagues used a spectrally sparse stimulus in which any given auditory filter is expected to pass energy from no more than tone, in contrast to the spectrally dense stimuli used here for which an indeterminate number of tones fall within a given auditory filter. According to a hypothesis proposed by Mathis and Miller (1947), phase effects occur only if two or more tones are processed by the same filter. Thus, no differences would be registered for fixed-phase and random phase conditions for spectrally sparse stimuli. However, the strong phase effects observed for spectrally dense stimuli suggest that the output of an auditory filter is not strictly energy and that some undetermined temporal factor must play a role in listening performance.

Chapter 4

Automatic Attention – EEG

4.1 Introduction

Auditory listening strategies can be dictated by the listener to an extent, while some experimental manipulations can force the listener to adopt alternative strategies. The transition between listening strategies was demonstrated in a bandwidth widening experiment (Berg 2007). The number of non-signal components was increased in a detection task that added a signal to the central component. The observed initial increase in thresholds is understood to follow as the increase in the number of non-signal components degrades the signal to noise ratio of the hypothetical decision statistic (i.e. equation 3.1). However, a breakpoint was observed where the thresholds eventually decrease from the listeners' improvement in performance. Provided that this non-monotonicity was observed under a consistent change in stimulus characteristics, it is evident that listeners are employing another mechanism in the computation of their decision variable. The likely process responsible for this threshold function is profile analysis.

Transitions can be demonstrated when listeners thresholds peak before decreasing by several dB. The difficulty in studying automatic attention can partially be attributed to the subjectivity of the internal state of a listener. Transition bandwidths force subjects to attend to different properties of an acoustics stimulus, namely the temporal and spectral characteristics. Traditional transition bandwidth experiments degrade the signal to noise ratio of the stimulus by increasing the number of components, n, forcing the subject to adopt a listening strategy that is dependent on the overall profile of the complex. Compared to thresholds estimated original ETB experiments, this experiment produced a decrease in thresholds by nearly 20 dB for multiple subjects by substituting the randomization of the phase of each component for fixed phases. The increase in transition salience strengthens support for the existence of a separate process that listeners are unknowingly integrating into the calculation of their decision variable. This provides evidence of a mechanism that dynamically allocates attention.

Transition bandwidths occur when listeners adapt their decision strategies as the bandwidth of a stimulus is increased (Berg 2007). The estimated transition bandwidth is defined by the peak threshold across the number of components and appears to be independent of component separation (Δf) and central frequency (Turner, dissertation 2010). Evidence for the adaptation of a dynamic, stimulus driven, decision strategy is supported by the weight profiles attained for multiple *n* conditions (Berg 2013). The estimation of the weight that each component contributed to the listener responding signal used methods developed by Berg (1989). Amplitude perturbations were independently sampled from a normal distribution for each spectral component.

Beginning with a narrow-band, 3 component complex, thresholds are estimated at increasing bandwidths. The number of components, *n*, is increased by odd numbers through 23 components. The listener's task is to identify the signal interval in a two-interval, forced choice task. Previous studies randomly sampled the phase of each component independently (check range of phases). The phases of the spectral components were fixed in all conditions for this experiment. Provided the enhancement of the transition bandwidth from holding the phase of individual components constant, such stimuli parameters are adapted for this experiment. Furthermore, the fixation of phase increases computational efficiency of phase-locked responses to recover information related to the envelope following response.

Thresholds will be estimated for each *n* condition, while the scalp potential of each subject is recorded with EEG. It is intuitive to expect scalp recordings modulated by a waveform similar to the envelope of the sound, but it is speculated that the strength of similarity between these waves will be stronger for small *n* conditions. Justification for this expectation founds from previous experiments that demonstrate the envelope of the waveform to be the primary predictor of performance for these conditions. For conditions with larger *n* components, the power of the F0 is expected to increase proportionally to the number of components.

4.2 Methods

4.2.1 Listeners

The same listeners from Chapter 3 were used for this study. Listeners came in for 2 days of training (Chapter 3), and a third day for EEG collection.

4.2.2 Materials

Listeners engaged in a two-interval forced choice task to estimate thresholds for each *n* condition, and were instructed to identify the interval which contained a signal. Stimulus parameters were identical to those on Chapter 3. Responses were collected with a Cedrus response box which allowed for the ergonomic collection of responses to minimize motor related, EEG artifacts. Auditory feedback was provided to listeners in the form of a 4 kHz tone for incorrect responses and a 5 kHz response for correct responses. Feedback was only provided after stimulus offset, and responses that preceded stimulus offset were omitted. A 1 second gap separated feedback offset from the onset of the next trial. The inter-stimulus interval was uniformly distributed between 300 and 700 ms to minimize the effects of stimulus habituation. Threshold estimates were computed identically from methods in Chapter 3.

4.2.3 Conditions

Two *n* conditions were selected for each side of the peak threshold (ETB). Selection of the specific number of components was determined on an individual basis, and by the criteria that the number of components produced the least different thresholds, compared to the respective counterparts on the opposite side of the ETB. In total, 4 *n* conditions were selected for each subject.

4.2.4 EEG Materials

Stimulus artifacts in the EEG recordings were minimized by the usage of piezoelectric insert earphones which omitted sound from distant from the scalp (typically clipped to collar), via soft plastic tubes. Standard headphones would not suffice for these recordings as the electromagnetic properties of stimulus generation would be introduced to scalp recordings.

EEG recordings were collected with a 128 channel Compumedics Quik-Cap and Compumedics Quik-Gel conductive gel. Recordings were sampled at 1 kHz via Neuroscan Synamps and Neuroscan software. Noisy electrodes were identified during recording sessions by the author, and were identified by either being abnormally large in amplitude, small in amplitude, or unlike surrounding electrodes. These electrodes were fixed between blocks. For noisy trials that were recorded, artifact rejection techniques were applied.

4.3 Results

Psychophysical thresholds were estimated for four bandwidths, two on each side of the breakpoint from psychophysical training. Thresholds are plotted in Figure 4.1, where each panel corresponds to a different subject. Consistent with previous results, higher bandwidths *n* conditions yield lower threshold estimates. Because of the consistent individual differences in breakpoint bandwidth, the bandwidth of the four conditions is variable between subjects. To minimize the differences between subjects, the two large *n* conditions were held constant at 560

Hz and 640 Hz for three of the four subjects. This was not possible for the remaining listeners because of the relatively large bandwidth of that subjects' breakpoint.

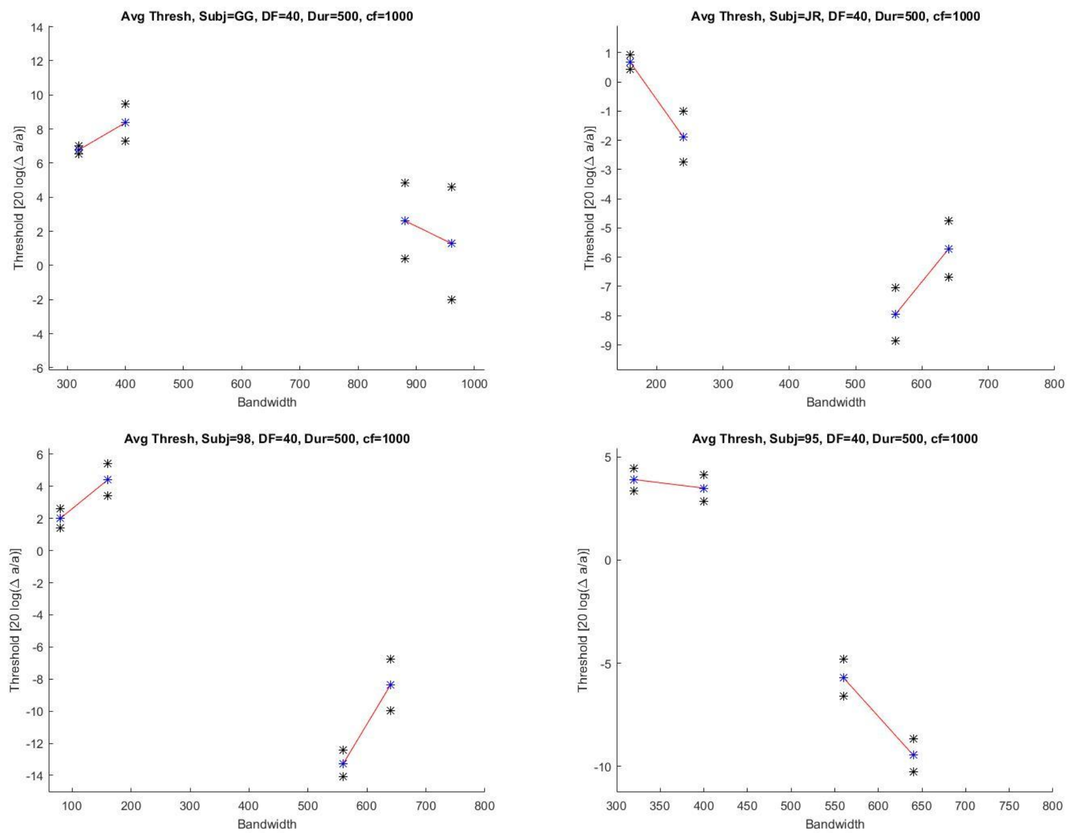

Figure 4.1: Psychophysical threshold estimates from EEG recordings.

Figure 4.2 demonstrates the frequency activity from all EEG electrodes in the most extreme *n* conditions from each side of the breakpoint. The fft was computed for the ERP between 300 ms and 500 ms. The motivation for this duration was to recover a better signal to noise ratio of the FFR due to broadband activity caused from the onset of ERP activity. Stimuli lasted for a duration of 500 ms.

Upon visual inspection, a noticeable peak changes in power between the two conditions plotted in figure 4.2. The condition in the right panel has greater 40 Hz power.

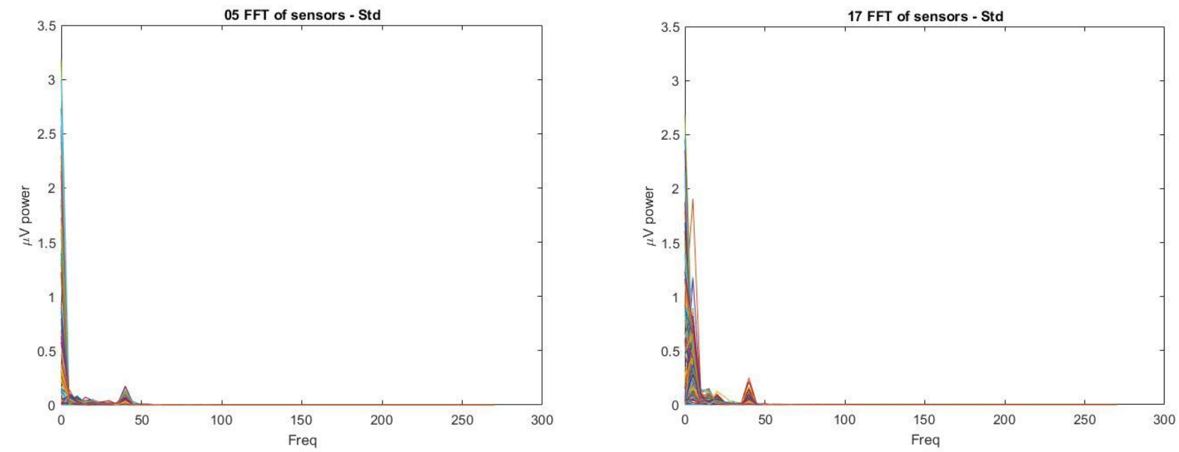

Figure 4.2: FFT of all sensors. FFTs were computed on each electrode from 300 – 500 ms.

The difference between these two conditions was investigated by bootstrap re-sampling the FFR at the Δf across conditions, shown in Figure 4.3. The attained distributions reveal an increase in 40 Hz power between the 5 *n* and the 17 *n* conditions. This effect is likely due to an increase in the SNR of the EEG recordings caused by an increase in the bandwidth of the stimulus. Such an increase in bandwidth would yield more activated neural fibers, thus increasing the power of the activity at 40 Hz.

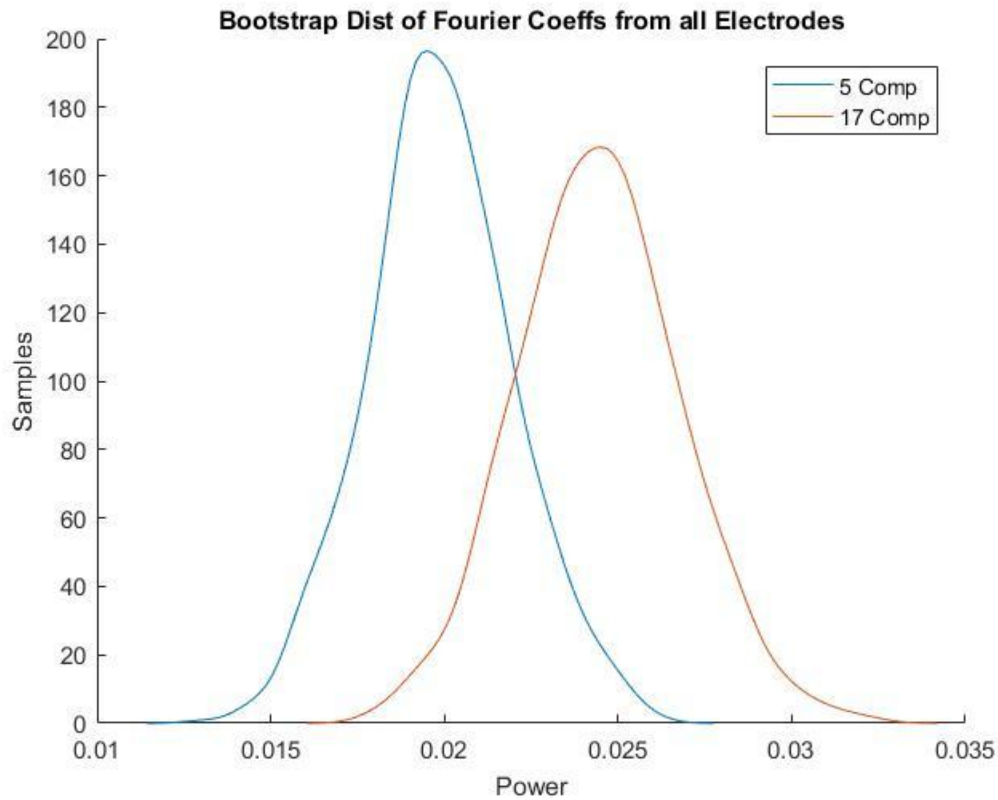

Figure 4.3: Bootstrap re-sampling of 40 Hz Power.

4.4 Discussion

Two stimulus bandwidths from each respective side of the breakpoint for the individual subjects' ETB constructed the four conditions that each subject was exposed to during EEG sessions. The estimated thresholds remained comparable between psychophysical and EEG sessions. Subjects scalp potentials displayed increasing power at the 40 Hz region as the number of components are increased. It is unclear whether this increase in power is independently associated with the characteristics of the stimulus, or if the characteristics of the stimulus are combined with endogenous activity.

It is hypothesized that the 40 Hz envelope of the stimulus is the dominant cue in conditions with low *n*. Therefore, it would be expected that the EEG recordings from those conditions would yield enhanced 40 Hz activity compared to conditions where the stimulus characteristic dominating listener's performance is the spectral profile.

Chapter 5

Conclusion

Evidence for two separate mechanisms of attention has been presented throughout the various experiments in this dissertation. Dynamic mechanisms of attention allocate resources for the appropriate computations given a specific tasks demands. Evidence has been proposed that supports these mechanisms are dynamic both within a constant stimulus and between stimuli that vary along a limited dimension.

The second chapter provided measurements of both psychophysical and biological data that support existing literature that specify the selective attention mechanisms. These mechanisms rely on task relevant information in the computation of a decision statistic, and enhance the cortical representations of that information.

Transition bandwidth experiments were proposed in chapter 3 and demonstrate the dynamic allocation of the mechanisms responsible for attention as the characteristics of a stimulus are altered. The characteristics that are altered force the computation of the decision variable from envelope information at lower *n* conditions, to profile analysis at large *n* conditions.

Bibliography

Berg, B. G. (1989). Analysis of weights in multiple observation tasks. *J. Acoust. Soc. Am.*, 86(5)

Berg, B. G. (2007) Estimating the transition bandwidth between two auditory processes: Evidence for broadband auditory filters. *J. Acoust. Soc. Am.*, 92(121)

Berg, B. G. 2013 A decision weight analysis of transition bandwidths. *J. Acoust. Soc. Am.*, 07(133)

Berg, B. G., Green, D. M. (1990) Observer efficiency and weights in a multiple observation task. *J. Acoust. Soc. Am.*, 88(149)

Berg, B.G., Nguyen, Q. T., and Green, D.M. (1992). Discrimination of narrow-band spectra. i: Spectral weights and pitch cues. *J. Acoust. Soc. Am*., 92(4):1911–1918.

Borucki, E., Berg, B. (2017) Isolating spectral cues in amplitude and quasi-frequency modulation discrimination by reducing stimulus duration. *Hearing Research*, 17(348)

Cherry, E. C. (1953). Some experiments on the recognition of speech, with one and two ears. *J. Acoust. Soc. Am., 25(5):975–979.*

Cohen, M. (2014) Analyzing Neural Time Series Data. *MIT Press.*

Dai. H. (2000) On the relative influence of individual harmonics on pitch judgment. *J. Acoust. Soc. Am.* 107:953–959.

Green, D. M. (1983). Profile analysis: A different view of auditory intensity discrimination. *American Psychologist* 83(38).

Green, D. M. (1992). The number of components in profile analysis tasks. *J. Acoust. Soc. Am.* 91, 1616–1623.

Green, D., Berg, B., Dai, H., Eddins, D., Onsan, Z., and Nguyen, Q. (1992). Spectral shape discrimination of narrow-band sounds. *J. Acoust. Soc. Am.*92, 2586(1992)

Hafter, E. R., Sarampalis, A., Loui, P., (2008). Auditory perception of sound sources, *Springer*.

Hall, J. W., Haggard, M.P., and Fernandes, M. A. (1984a). Detection in noise by spectrotemporal pattern analysis, *J. Acoust. Soc. Am*. 76, 50- 56.

Levitt, H. (1971). Transformed up-down methods in psychoacoustics. *J. Acoust. Soc. Am.*, 49(2):467–477.

Näätänen, R. (1990). The role of attention in auditory information processing as revealed by event-related potentials and other brain measures of cognitive function. *Behavioral and Brain Sciences*. 13(2).

Perrin, Pernier, Bertrand, and Echallier (1989). Spherical splines for scalp potential and current density mapping. *Electroencephalogr Clin Neurophysiol*. 72(2):184-7.

Southworth, C. and Berg, B. G. (1995). Multiple cues for the discrimination of narrow-band sounds. *J. Acoust. Soc. Am.*, 98(5).

Turner, M., (2010). *COSS analysis for adaptive tracks and some problems related to the estimated transition bandwidth*. PhD thesis, University of California, Irvine.

Zwicker, E., Flottorp G., and Stevens, S.S. (1957). Critical bandwidth in loudness summation *J. Acoust. Soc. Am.* 29, 548-557.

Appendix A

Deprived Temporal Information Produces Intensity Judgments

To further develop the psychophysical procedures from Berg (1995), auditory stimuli were abbreviated in duration, so that listeners were deprived from making judgments based on envelope information. Several psychophysical models quantify the perception of roughness judgments by calculating decision variables derived from envelope statistics (Green, 1983, Berg, 1995). This manipulation forces listeners to make decisions based on the overall intensity of the sound. Pitch judgments were confined by both the lack of envelope information and the narrowband nature of the stimulus.

The effect of stimulus duration on listener performance serves an experimental control for multiple stimulus characteristics. It is suggested that suggest that shorter duration stimuli warrant an increment in amplitude for detection, as tones greater in duration can be detected at lower amplitudes (Watson and Gengel, 1969). Watson and Gengel observed that the temporal integration of a stimulus is dependent on the frequency of the signal.

Humans have the ability to selectively attend to a sound while other competing sound sources are present, as well as attend to various features of a single sound (Cherry, 1957, Berg, 1989). Developments in psychophysical measurements have allowed researchers to ascertain which feature a subject is attending to, while others are present. Such features include loudness, timbre, and pitch. The spectral weights produced are generally as expected, based on previously existing psychophysical models. However, individual differences may best be explained by considering the possibility that these features do not occur in isolation.

To better understand the variation produced across individuals, isolating these features and simulating possibly mixtures of features may produce a more holistic explanation. This work has been initiated to better understand roughness weights, by normalizing stimuli intensity and employing a frequency –rove (Tan, 2017). Depriving the listener of that information forces the listener to attend to the envelope of the stimulus, producing weights characteristic of roughness. This experiment attempts to ascertain the stimulus characteristics that produce loudness weights, by abbreviating the stimulus to a short duration.

Degrading temporal cues via shortening stimulus duration is an intuitive effect and has been well documented with amplitude modulation tasks (Viemeister 1979). Experiments employing amplitude modulation and quasi-frequency (AM/QFM) provide an example of how the spectral cues can be isolated via reduced stimulus duration. The abbreviation of duration in an AM task caused listeners' thresholds to increase roughly 20 dB, and perform nearly at chance (Borucki and Berg, 2017). The AM of the stimulus increased from 50 to 800 Hz, where subjects were instructed to detect changes in the depth of the modulation. A non-monotonicity was observed for higher rates of modulations, where thresholds decreased after an initial increase. It is speculated that the increase in listener performance at higher frequencies is caused by distortion products interacting with low frequency components of the complex. Support for this hypothesis was demonstrated via the abbreviated duration of the stimulus which has no effect on thresholds in this modulation range. This follows because such an interaction involves the spectral properties of the stimulus, which are unaffected by stimulus duration.

Previous work suggests that decisions that are calculated based on envelope information require the accumulation of that information over time in order to calculate a decision variable (Green 1983, TMTF 1979). In this experiment, the accumulation of that information will not be possible due to the abbreviation of the stimulus, forcing listeners to make judgments based on intensity information. Psychophysical weights will be derived that can either support or reject this hypothesis.

General Methods

Participants

Participants were students and researchers at the University of California, Irvine and ranged in age 19 to 29. Prior to beginning data collection, all listeners were screened to be normal listeners, as defined by thresholds equal to or less than 20 dB SPL for the frequency regions to be presented.

Methods

Thresholds were calculated via a two-down one-up procedure, calculating estimates via Levit, 1971. Subjects were presented with a two-alternative forced choice task, with the signal randomly assigned.

Psychophysical weights were calculated from the varying perturbation level of each component as a function of the probability of the listener correctly identifying the signal.

Materials

Both conditions involved instructing the listeners to attend to the acoustic feature of loudness.

Stimuli were constructed and presented via Matlab. Complex tones were constructed in the spectral domain, so that amplitude perturbations could be added to the stimulus with more computational efficiency.

The conditional manipulation for this experiment was stimulus duration, where Condition 1 was presented for 500 ms and Condition 2 was presented for 30 ms.

Results

Responses from each subject were measured by the probability of being correct across the spectrum of amplitude perturbations for each component, producing three psychometric functions. These functions are summarized as a weight profile that represents the sign of the slope and the variance of the best-fit psychometric function.

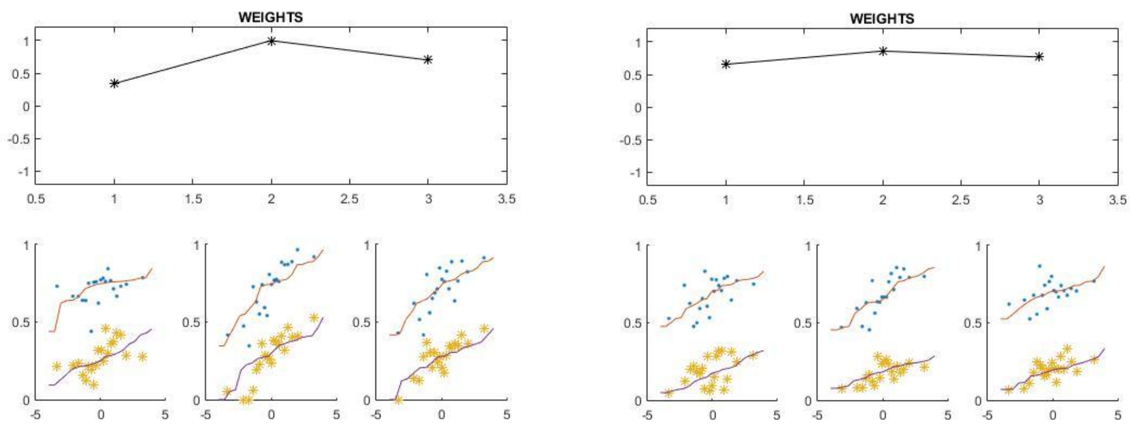

Figure 1. 500 ms and 30 ms Loudness Judgments. Top: summary weights for the fit for each psychometric function, per each component. Bottom: each x-axis represents the perturbation spectrum for each component, plotted as a function of the listener responding signal for each perturbation level.

Discussion

Weight profiles suggest that depriving listeners of envelope information forces the usage of a decision variable calculated from the intensity of the stimulus. The weights of the side components were increasingly dependent on the amplitude of the perturbation when duration of the stimulus was limited.

This psychophysical manipulation serves as a hypothesis test to ascertain the characteristic of the stimulus which contributes to the specific pattern of weights associated with loudness. The stimulus isolates the perceptual state of the listener, and forces them to attend to what they perceive as loudness.

Future psychophysical methodologies may adapt this manipulation as a training procedure. It is unknown whether or not the deprivation of stimulus characteristics produces a training effect that changes the profile of weights compared to when relevant characteristics were available. Historically, psychophysical experiments employ a pre-stimulus cue to either manipulate or control the attention of the observer (Hafter et al, 2007). Given the stimulus manipulations that influence the attended acoustic cue of the listener, a pre-stimulus cue consisting of the respective manipulation could potentially reduce the internal noise present in the observer.

Appendix B

Roughness: Alpha-Beta Weight Simulations

Individual differences in spectral weight patterns reflect variance in listening strategies per making use of the various stimulus statistics available. Timbre judgments dominated by the envelope of the sound rely on the frequency separation of the components and the associated fundamental harmonics. Simulations demonstrate that applying different weights to these modulation channels can alter the magnitude of spectral weights, while retaining the sign of each weight, as well as the feature-specific profile.

The development of applying weights to the modulation rates which dominate the timbre decision variable allows for an estimation procedure to classify listeners as either “alpha” or “beta” listeners. Either classification refers to the first or second harmonic of the envelope of a sound that is constructed with three sinusoidal components that are symmetrically separated around a central frequency (CF). This advancement allows for both hypothesis testing in future psychophysical experiments, and accounts for the variability in weight profiles across listeners.

The results from the alpha-beta simulations produce minor variations in the profile of spectral weights, similar to that seen from individual differences. This recent advancement can be immediately applied to results obtained from the recordings of scalp potentials from listeners that are participating in the timbre discrimination task in the previous chapter. Although it is not necessarily within the resolution of EEG, it is possible that the envelope following response of the cortex could be influenced by the weights assigned to the various rates of modulation. For instance, the weight profile of an alpha listener may be associated with neural activity that enhances alpha frequency activity. This is not to be confused with alpha rhythmic activity. Alpha frequency activity refers to the envelope following response of the cortex for the respective alpha modulation rate. Alternatively, a given alpha or beta listener might suppress the activity of their non-dominant modulation pattern.

The acoustic feature of timbre is commonly described as the roughness/smoothness of a sound, and is largely influenced by the Hilbert envelope of that sound. The envelope of a sound is typically comprised of resolved harmonics that modulate the amplitude over time. Previous studies have demonstrated that this acoustical cue can be isolated from other cues (Green et al. 1992, Richards 1992). Green (1992) demonstrated this by randomly varying the overall level of the sound, degrading the listeners' ability to rely on intensity information. A model to describe the data was proposed that explains the data assuming the listener is calculating their decision variable from information present in the power spectrum of the Hilbert envelope.

Southworth and Berg (1995) reported that the weight profiles which listeners produced when attending to the timbre of a sound changed from other feature conditions, but were subject to individual differences. All but one subject did not conform to the model predictions that were proposed by Green (1992). The decision rule that is proposed by this model predicts that the listener will identify a signal in interval one when:

$$\Sigma (\eta i-\tau i)zi1 > \Sigma (\eta i-\tau i)zi2, \qquad (1)$$

where the observation vectors of ηi and τi represent the Fourier coefficients computed from the Hilbert envelope of wither the standard or signal respectively. The z coeffieient is a weighting variabel This model treats all of the respective Fourier coefficients with the same weight. However, in order to advance this model and account for individual differences, Berg has proposed a procedure that dynamically weights the Fourier coefficients computed by the model.

Traditional applications of this model involve a complex of tones that have given amplitudes and a particular frequency separation. The number of components and frequency separation dictate the Fourier coefficients computed from the Hilbert envelope, but in a simple three-tone experiment these coefficients are limited to Δf and twice Δf. These two coefficients are the primary focus of this section and are referred to alpha and beta, respectively.

The estimation of weights on separate harmonics was conducted by conducted by Dai (2000) where a random perturbation was applied to the frequency of the harmonic on an individual trial basis. This produce allowed for the calculation of a perceptual weight across harmonics. Results demonstrated that the frequency region around 600 Hz dominated the listeners' responses instead of a particular harmonic number. The weighting procedure employed by Dai is not to be confused with the weights that are assigned to the harmonics in the α-β simulations. Where Dai attainted weights for the respective harmonics used by listeners, here α-β weights are a manipulation of the decision variable calculation that is dominated by the first and second harmonics.

Findings from Dai support earlier work that ascertained a cutoff point where listeners change their decision criteria from relying on higher order to lower order harmonics (Plomp 1967). Plomp conducted a pitch discrimination task by decreasing the fundamental frequency of the lower harmonics from a baseline F0, and increasing F0 for higher harmonics. The experimental manipulations for this experiment were the fundamental frequencies and the point of distinction between lower and higher harmonics. The findings suggest that the fundamental frequency influences which harmonic the listener uses when making comparisons with other sounds. Fundamental frequencies strongly impacted listeners' decision strategies for F0 at and above 1400 Hz, and as F0 was decreased, the listeners relied on higher order harmonics.

Method

To better understand the dynamics of the multiple harmonics available for listeners, simulations adopted the model from Green (1992) that best explain the data for listeners attending to timbre (see equation #1). The Fourier coefficients α-β were assigned a fixed weight and a variable weight, which alternated dependent on the condition.

Discussion

The consistent symmetry of the weight profiles do not address the sometimes asymmetric profiles that are produced by some listeners. The relationship between the primary and secondary harmonics appears to influence the magnitude of the weights of each side component. Listeners that weigh their decision variable heavily on α variable produce less negative weights compared to listeners that rely heavily on β weights.