UC Irvine
UC Irvine Electronic Theses and Dissertations

Title
Molecular Psychophysics of Selective Auditory Attention

Permalink
https://escholarship.org/uc/item/3cc6d2wr

Author
Tan, Alison Y.

Publication Date
2015

Peer reviewed|Thesis/dissertation

eScholarship.org
Powered by the California Digital Library
University of California

UNIVERSITY OF CALIFORNIA,
IRVINE

Molecular Psychophysics of Selective Auditory Attention

DISSERTATION

submitted in partial satisfaction of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

in Psychology

by

Alison Y. Tan

Dissertation Committee:
Professor Bruce Berg, Chair
Professor Charlie Chubb
Professor Fan-Gang Zeng

2015

© 2015 Alison Y. Tan

DEDICATION

For my family, friends and mentors.

"*We shall not cease from exploration. And the end of all our exploring will be to arrive where we started and know the place for the first time.*" - T.S. Eliot

"*The problem with the human observer is that he or she is human.*" - S.S. Stevens

"*The single most salient characteristic of the auditory sense is its analytic ability.*"- David M. Green

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

I would like to express my deepest gratitude to my advisor, Bruce Berg, who has continually provided encouragement and guidance in the pursuit of scientific truths. I would also like to extend my appreciation to my committee members, Charlie Chubb and Fan-Gang Zeng who have provided time, critical insight, and guidance into this work.

I extend a special thanks to additional members of my advancement committee, Kourosh Saberi and Sheryl Tsai.

I would also like to express many thanks for the invaluable guidance from Michael Lee who made Bayesian insights possible.

Thanks to my fellow lab members, Michael Bellato and Matthew Richardson who have provided wonderful conversation and feedback.

I would also like to thank friends from UCI for many things: Tyler Dean for the music, Karina Rubin for the wit, Samhita Dasgupta, Catherine Yang, and Alicia Brayboy for the support.

Finally, I want to thank my partner-in-crime, Justin, for the humor, love and endless read-throughs.

Financial support was provided by the University of California, Irvine and NSF Grant BCS-0764003.

CURRICULUM VITAE

Alison Y. Tan

alisonytan@gmail.com

EDUCATION

Doctor of Philosophy in Psychology **2015**
University of California, Irvine *Irvine, CA*

Masters of Science in Cognitive Neuroscience **2014**
University of California, Irvine *Irvine, CA*

Bachelor of Arts in Cognitive Science, Minor in Biology **2010**
University of California, Irvine *Irvine, CA*

AWARDS AND GRANTS

Associated Graduate Student Travel Grant **2015**
Department Travel Grant *Irvine, CA*

Summer Fellowship **2010-2012**
University of California, Irvine *Irvine, CA*

UROP Grant **2009-2010**
Access Mechanisms of Long-term Musical Memory *Irvine, CA*

UROP Grant **2009-2010**
Syllabic Structure of words affecting phonological Similarity *Irvine, CA*

Departmental Honors in Psychology **2008-2010**
University of California, Irvine *Irvine, CA*

RESEARCH EXPERIENCE

Doctoral Research **2012–2015**
University of California, Irvine *Irvine, CA*
Advisor: Dr. Bruce Berg

- Investigated training effects for cue switching for narrowband stimuli.

- Developed psychophysical tools to improve listeners use of cues
- Tested and updated quantitative models of auditory processes
- Obtained observer weights, performance and efficiency measures for dichotic presentation of sample discrimination paradigm
- Investigated temporal processing differences in diotic vs. dichotic presentation with varying inter-tone intervals and number of tones
- Managed software and hardware acquisition and upgrades
- Mentored and assisted training of 9 undergraduates in the lab

Undergraduate Research **2009–2012**
University of California, Irvine *Irvine, CA*

Advisor: Dr. Kourosh Saberi

- Investigated perceptual access to musical memory for pitch through both aural and/or vocal-motor systems
- Determined that musicians can accurately voice isolated musical tone with standard deviation of 3 semitones, supporting a two system model of pitch retrieval
- Collected behavioral data using psychophysical techniques (i.e.: testing listeners for absolute pitch and vocal recordings)

Advisor: Dr. Mary Louise Kean **2008–2010**

- Assisted in large scale research on determining the extent of phonological similarity effect (PSE)
- Implemented a syllabic model of word structure to address the structural difference between rhyme (e.g. sun fun run) and overlap (e.g. cat cab can) words
- Prepared literature reviews, developed proficiency at SPSS for data entry and analysis, assisted in data collection via Matlab, presented data at UCI symposium (UROP) and Departmental Honors Symposium

RESEARCH INTERESTS

- Mechanisms of auditory attention and sound segregation
- Selective attention for individual listening strategies in difficult listening environments

- Computational modeling of individual differences for diotic vs. dichotic listening at the peripheral and cortical levels
- Temporal processing differences in cocktail party environments for normal human subjects compared to hearing impaired
- Training effects using classic psychoacoustic paradigms to investigate auditory learning

TECHNICAL AND PERSONAL SKILLS

Programming Languages: Proficient in: Matlab, LaTeX, Basic R

Industry Software Skills: Matlab (Advanced), JASP (Advanced), SPSS (Advanced), Adobe Photoshop and Illustrator (Advanced), MS Office products (Advanced), and Apple Office Products (Advanced).

Other Technical Skills: Cantonese - English Bilingual

TEACHING EXPERIENCE

Introduction to Psychology **Summer 2014**
University of California, Irvine *Irvine, CA*

Introduction to Psychology I (Teaching Assistant) **2010-2012**
Introduction to Psychology II: Learning and Memory (Teaching Assistant) **Winter, 2011**
Language in the Brain (Teaching Assistant) **Spring, 2011**
Advanced Experimental Psychology (Teaching Assistant) **2015**
Probability and Statistics (Teaching Assistant) **Fall, 2013**
Human Memory (Teaching Assistant) **Winter, 2013**
Writing about Memory (Teaching Assistant) **Winter, 2012**
Hearing (Teaching Assistant) **2013–2014**
Linguistics (Teaching Assistant) **2014**
Neuropsychology: Frontal Lobes (Teaching Assistant) **Spring, 2015**

REFEREED CONFERENCE PUBLICATIONS

Tan, A.Y., Berg, B.G. (2015). Measures of ear lateralization in dichotic sample discrimination task. [Abstract]. *ASA Meeting Papers.* Pittsburgh, PA., May, 2015. Paper id: 1pPP36

CONFERENCE PRESENTATIONS

Tan, A.Y. & Berg,B.G.. (2015). Measures of ear lateralization in a sample discrimination task. *Acoustical Society of America.* Poster presented at 168th Meeting: Acoustical Society of America, Pittsburgh, PA, May, 2015.

Tan, A.Y. & Kean, M.L. (2010). Syllabic Structure of Rhyme and Overlap Words Affecting Phonological Similarity Effect.*Poster presented at 17th UROP Meeting.* University of California, Irvine, May, 2010

Tan, A.Y. & Saberi, K. (2010). Access Mechanisms in Long-Term Musical Memory. *Poster presented at 17th Meeting: UROP Meeting.* University of California, Irvine, May, 2010.

WORKS IN PROGRESS

Tan, A.Y. & Berg, B.G. (2015). Measures of selective attention in a dichotic sample discrimination task. *In progress.* to be submitted to Journal of Acoustical Society of America.

ABSTRACT OF THE DISSERTATION

Molecular Psychophysics of Selective Auditory Attention

By

Alison Y. Tan

Doctor of Philosophy in Psychology

University of California, Irvine, 2015

Professor Bruce Berg, Chair

Mechanisms underlying selective auditory attention currently lack a comprehensive theoretical framework for considering many empirical phenomena. For instance, there exists little scientific vernacular for discussing differences in dichotic and diotic performance under conditions with equivalent information. The following work applies an empirical paradigm that adds noise to the stimulus as a means to study attention to individual components of the stimulus complex. The first part uses a sample discrimination task that presents tones to alternating ears while manipulating parameters such as informativeness, inter-tone interval, and binaural differences in loudness. Only 2 of 9 listeners showed an optimal listening strategy. Most, instead, show a left-ear advantage and are unduly influenced by loud tones that carry little information. The second part examines selective attention to the basic dimensions of sound - pitch, timbre (e.g., roughness), and loudness. Methods were developed to manipulate the stimulus and observe the effect of specific cue. A theoretical model associated with each cue yields a different pattern of decision weights in a discrimination task. Listeners are instructed to attend to target cues and estimated weights quantify the degree of selective attention to that cue. Some listeners matched theoretical predictions, whereas others appear to use a combination of cues, rather than a single cue. I show that there are clear changes to estimated weights that directly reflect perceptual changes to an isolated cue when a second and then a third cue wis introduced. Findings from both sets of studies

introduce novel ways to study auditory selective attention and provide initial building blocks for a theoretical framework.

Introduction

Natural auditory environments often contain many complex and concurrent sounds. In a context where there are many competing sound sources, the human auditory system can still selectively attend to a source of interest while ignoring the rest. The mechanisms underlying selective auditory attention, however, are not yet fully understood; there is no comprehensive framework that relays how attention influences perceptual ability.

The primary goal of this research is to develop measurements to study and explore the limitations of auditory selective attention using molecular psychophysics. Chapter 1 begins by defining the individual listening patterns in a dichotic sample discrimination task in which tones with differing levels of informativeness are sent alternating to either the left or right ear. Most listeners showed better performance in the left ear among while only 2 out of 9 listeners can be considered optimal such that they can listen to the most informative target tones in difficult and easy conditions. The chapter then explores whether there are temporal differences in dichotic and diotic listening and how these manipulations can improve listener performance.

Chapter 2 examines listeners' ability to attend to specific cues, specifically pitch, timbre ("roughness"), and loudness. The perceptual experiences of these dimensions of sound are objectively analyzed, and I determine that different stimulus manipulations result in decision weights that reflect perceptual differences. While some listeners closely adhere to the model, others show deviations from the model which suggest a combination of listening cues that are reflected in their decision weights. The methods developed here generate quantifiable measures for different dimensions of sound and help create objective dialogue about individual perception.

Chapter 1

Measures of Selective Attention in a Sample Discrimination Task

1.1 Overview

The purpose of these experiments is to address individual differences in selective attention. A sample discrimination task was used to investigate the ability of listeners to selectively attend to information from one ear while ignoring information from the other ear. On each trial, seven 60 ms-tones were drawn from a normal distribution with means of 1000 or 1100 Hz. Half of the trials used "low-mean" stimuli in which the means of the distributions for both even and odd tones that were drawn were 1000Hz; the other trials used "high-mean" stimuli in which the means of both distributions were 1100 Hz. The task was to judge whether the stimulus on a given trial was a high-mean or a low-mean stimulus. The informativeness between even and odd tones is manipulated as well. Even numbered tones were the most informative (d'=2) and presented to one ear and the less informative, odd numbered tones (d'=0.5) were presented to the other ear. The tone sequence was presented alternating from ear to ear.

In all task conditions, the frequencies of tones 2, 4 and 6 were drawn from a Gaussian distribution with standard deviation 50 Hz. This allowed even tones to have less variance and provide more information than odd tones. The frequencies of tones 1, 3, 5 and 7 were drawn from a Gaussian distribution with standard deviation 200 Hz; thus making the odd tones less informative. In easy conditions, informative and uninformative tones were presented at 70 dB, 50 dB, respectively. In difficult conditions, informative and uninformative tones were presented at 50 dB, 70 dB, respectively.

Psychophysical measures like decision weights, efficiency measures and performance level (d') all showed an unexpectedly wide range in a listener's ability to selectively attend to the most informative tones. Estimates of d' range from 2.4 to 0.7. These measures are a step towards developing a more quantifiable method to investigate auditory selective attention. Some listeners displayed high efficiency estimates in all conditions while others showed a marked ear preference, achieving high efficiency only when the informative tones were presented to a particular ear. Another group of listeners showed a distinct inability to selectively attend in any condition. This latter group is most affected by the intensity manipulation.

1.2 Introduction

The ability to attend to one ear while ignoring information from the other ear is a useful listening tool in any acoustic environment. Intuitively, one can imagine that an individual can listen using their right ear equally as well to one could to their left ear. Dichotic listening experiments have been a preferred method in the last half of the century to investigate how listeners attend to stimuli in different channels (Cherry, 1953; Broadbent, 1954; Hugdahl, 2011). Many variations of dichotic presentation have been developed such as the

CV-syllables paradigm (Studdert-Kennedy and Shankweiler, 1970) in which pairs of syllables are presented and the subject reports which syllable was perceived after each immediate trial. Or a cued attention task in which listeners are given a cue that predicts the location of a subsequent target stimuli for which the listener can respond (Ofek and Pratt, 2004) while other experiments of dichotic presentation, inform listeners beforehand to focus their attention on stimuli arriving at a specific ear (Bryden et al., 1983).

Berg (1990) estimated decision weights using a task where a sequence of tones are presented to both ears with frequencies drawn from normal distributions. In the diotic sample discrimination task, sensitivity indices for even and odd tones were d'=2 and d'=1, respectively, so even tones were more informative and reliable. Berg found that listeners placed greater emphasis on more intense tones even if they came from a less reliable samples.

This robust effect, known as level dominance, has been studied in various contexts including: long temporal gaps between tone bursts (Turner and Berg, 2007), tones sampled with different overall level and large level perturbations (Lutfi and Jesteadt, 2006), and when identifying sound sources that form basic auditory objects (Lutfi et al., 2008). Additional cases like using wide-band noise sequences (Oberfeld and Plank, 2011) and detecting the order of statistical changes in a sound stream (Richards et al., 2013) also yielded the same strategy in which listeners attended to the loudest stimuli in the display.

In this study the effect of level dominance is utilized to manipulate task difficulty. We use a dichotic sample discrimination task with seven tones where the highly informative even-tones are sent to the target ear in every condition. Task difficulty is manipulated by sequentially alternating the tone bursts and the intensity of the tones sent to each ear, with 20 dB level difference between even and odd tones. In difficult conditions, informative and less informative tones are presented at 50 dB, 70 dB, respectively and vice versa for easy conditions.

1.3 Experiment 1: Dichotic sample discrimination task

1.4 Methods

1.4.1 Listeners

Nine students from the University of California, Irvine participated in a dichotic sample discrimination task. All subjects ranged in age from 18-28 years and were screened for normal hearing. Listeners displayed less than 20 dB hearing loss for pure tones ranging from 0.5-8 kHz with the exception of one listener AN who showed 20 dB HL at low frequencies 0.125 - 0.5 kHz for her right ear only. Further testing with DPOAEs indicated she has some outer hair cell loss at low frequencies.

1.4.2 Procedure

On each trial, subjects were presented a sequence of seven tones drawn with the same mean. The tones were 60 ms in duration and alternate from the from one ear to the other without pause, essentially an inter-tone interval of 0 ms. Listeners were instructed to attend to and report from one ear (i.e: the right or left ear). On any given trial, tones could were either drawn from the *low mean* distribution (μ_L = 1000 Hz) or a *high mean* distribution (μ_H = 1100 Hz). The task was to judge whether the stimulus on a given trial was a *high-mean* or a *low-mean* stimulus.

The frequencies of the odd tones (first, third, fifth and seventh) were sampled from a distribution with standard deviation of 200 Hz ($d' = 0.5$). This has greater variance and lower reliability while the even tones (second, fourth, and sixth) were sampled from a distribution with a standard deviation of 50 Hz which has less variance and higher reliability ($d' = 2$).

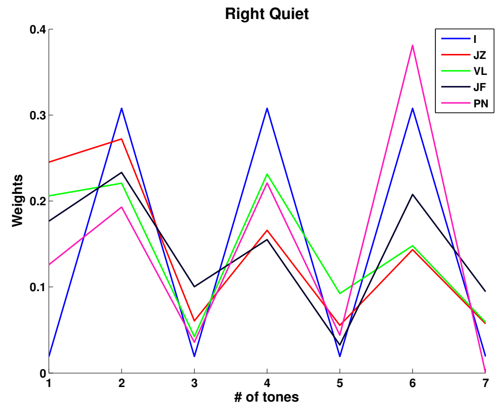
After presentation of tone sequence, listeners indicated whether the even tones were sampled from the high mean or low mean distribution. Task difficulty is manipulated by presenting odd-and even-numbered tones at different intensities. In easy conditions Right Loud (RL) and Left Loud (LL), informative and less informative tones are presented at 70 dB, 50 dB, respectively. In difficult conditions, Right Quiet (RQ) and Left Quiet (LQ), informative and less informative tones are presented at 50 dB, 70 dB, respectively. Listeners were provided immediate feedback on their accuracy after each trial.

Stimuli were generated with Matlab R2008a running on a PC with Windows 7. The waveforms 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. Listeners had unrestricted response time.

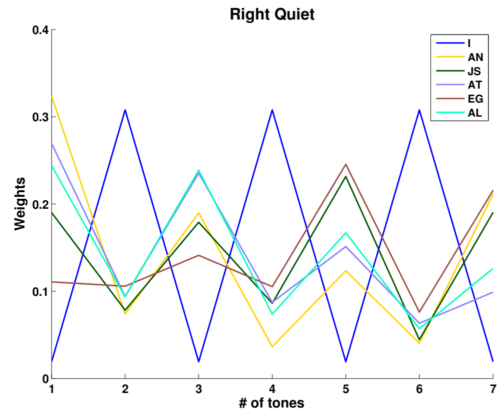
1.5 Results

Decision weights are estimated for each subject across conditions (Berg, 1989). This method assumes that observations are combined to produce a decision statistic of the weighted sum, $\sum a_i(x_i + \epsilon_i)$ where $x_i, i =$ 1,2,...n is the tone frequency for the ith observation, a_i is the weight associated with x_i and ϵ_i represents internal noise that is normally distributed with a mean of zero and a variance of $\sigma^2_{internal}$. Figure 1.1 shows estimated weights as a function of temporal position for each subject. Optimal weights for each condition are shown by the blue dotted line (see Appendix A). Listener results show a wide range of individual differences for estimated decision weights in the difficult conditions (RQ, LQ), while easy conditions Right

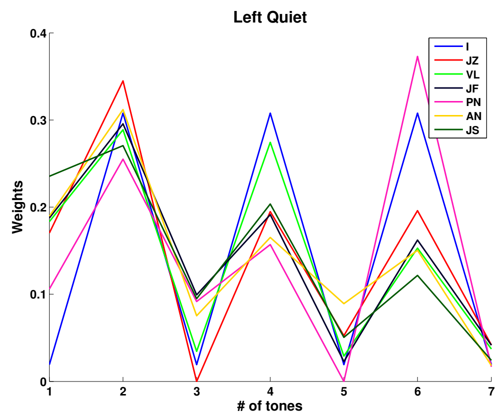
Loud and Left Loud (RL, LL) led to more optimal performance for most listeners.



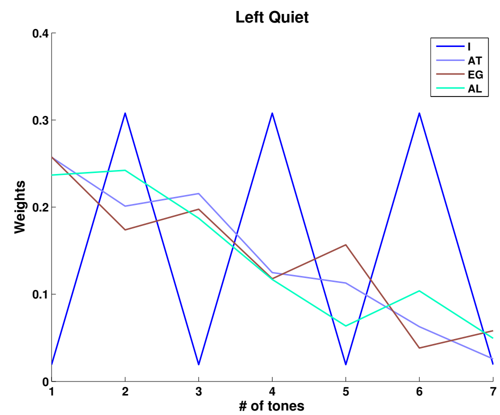
(a)



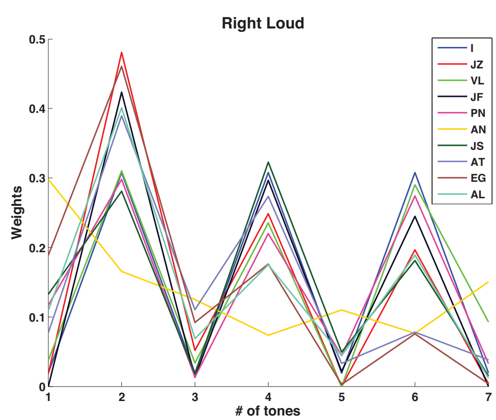
(b)



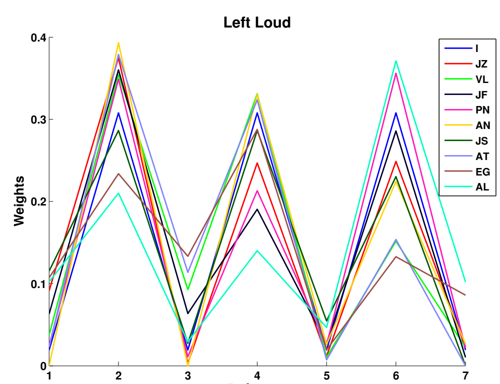
(c)



(d)



(e)



(f)

Figure 1.1: Weights for each condition grouped by listening strategy. Ideal weights are shown in blue with peaks at the even tones. 1a and 1c show Right-Quiet Able and Left-Quiet Able listeners who listen to the most informative tones, respectively. 1b and 1d demonstrate listeners who are listening to the tones in the non-target ear (Right Quiet, Left Quiet). All individuals except for AN perform optimally in Right Loud and Left Loud conditions (1e,1f).

Listeners in Figure 1.1a and 1.1c represents listeners who are able (Right-Quiet Able, Left-Quiet Able) to attend to the most informative tones for RQ and LQ conditions. While Figure 1.1b and 1.1d show listeners who are unable to attend to the more informative tones for RQ and LQ. It is of interest that some listeners who were able to attend to the most informative tones in RQ were not able in LQ and vice versa. Figure 1b shows the subset of listeners who weighted the odd tones greater in the left ear than the more informative tones in the right ear. Figure 1e shows a subset of listeners in the LQ condition that had a primacy effect, placing greater weight on the initial tone instead of the more informative tones.

All of the participants had better performance in the RL and LL conditions (d'), except for AN. Due to hearing loss at low frequencies for the subject's right ear, AN can be observed in the panel Figure 1.1c to be the only listener with a non-optimal weighting strategy in that condition.

Additional results from the mean efficiencies for all conditions across listeners are listed in Table 1.1. The equations for calculating efficiency follow Berg (1990). Mean performance and efficiency measures are grouped here by listener weighting strategy, as shown in Figure 1.1. The listeners in Right-Quiet Able are grouped by the subject list shown in Figure 1.1a and Right Quiet are listeners grouped from subjects in Figure 1.1b. The averaged d' for Right-Quiet Able and Left-Quiet Able were as good as those for the easier conditions (RL, LL).

Table 1.1: Mean efficiency estimates for subjects grouped by listening strategy

	d'	η_{wgt}	η_{noise}
Right-Quiet Able	2.26	0.36	0.316
Right Quiet	1.01	0.118	0.418
Left-Quiet Able	2.29	0.430	0.297
Left Quiet	1.64	0.211	0.379
Right Loud	2.08	0.592	0.297
Left Loud	2.22	0.698	0.238

The first measure η_{wgt} is the loss in efficiency from listeners using non-optimal weights,

$$\eta_{wgt} = \frac{(d'_{wgt})^2}{(d'_{ideal})^2} = \frac{\sum \hat{a}_i^2 \sigma_i^2}{\sum a_i^2 \sigma_i^2}, \tag{1.1}$$

where d'_{ideal} is the ideal performance with $\hat{a}_i$ being the ideal weights for an optimal observer (Appendix A) and σ_i is the standard deviation at the ith temporal position. D'_{wgt} represents an ideal observer who has a non-optimal weighting strategy where a_i represents the observed weights from the listener. This efficiency measure closely represents the differences observed among the RQ, LQ and RQ-Able and LQ-Able groups. The mean estimate for η_{wgt} across individuals in the Right-Quiet Able and Left-Quiet Able conditions was greater by a factor of 2, 0.36 and 0.43, respectively, than listeners in Right Quiet and Left Quiet. The second measure η_{noise},

$$\eta_{noise} = \left(\frac{d'_{obs}}{d'_{wgt}}\right)^2, \tag{1.2}$$

where d'_{obs} is the listener's observed d' and η_{noise} represents additional loss in efficiency from factors like internal noise and Table 1.1 shows that η_{noise} is stable across conditions. Lastly, the mean estimate η_{obs} which relates observer performance to ideal performance shows a similar pattern to η_{wgt}. Right-Quiet Able and Left-Quiet Able are better by half compared to Right Quiet and Left Quiet group.

Main Effects

A 2 x 2 within-subjects repeated measures ANOVA for level (quiet or loud) and ear (right or left) was performed for each individual on the basis of their d' values for all blocks across each condition (i.e.:RQ, RL, LQ, LL) (see Table 1.2). The effect size r^2 is left blank for individuals that did not have an interaction between ear and level.

Table 1.2: ANOVA results for each individual

	Main effect of ear			Main effect of level			Interaction		
Subject	F	p	r^2	F	p	r^2	F	p	r^2
JZ	4.502	0.042	0.127	212.601	0.042	0.440	0.462	0.502	-
VL	5.926	0.035	0.372	0.441	0.522	-	5.912	0.035	0.372
JF	0.496	0.491	-	0.001	0.970	-	0.298	0.592	-
AL	26.715	0.001	0.7276	43.649	0.001	0.8136	30.639	0.001	0.7539
AN	51.47	0.001	0.799	39.061	0.001	0.750	30.006	0.001	0.698
JS	6.819	0.021	0.3275	104.455	0.001	0.8818	15.828	0.001	0.5306
AT	0.011	0.919	-	42.079	0.001	0.7928	26.230	0.001	0.7045
PN	3.147	0.101	-	2.673	0.128	-	0.558	0.469	-
EG	6.633	0.022	0.321	62.549	0.001	0.817	0.127	0.727	-

All subjects but two (JF and PN) showed an ear effect, with significantly better performance in the left ear (Table 1.2). Optimal listeners in this experiment would be expected to show no main effects since an optimal listener would not be affected by ear or level of presentation. From these analyses, only 2 listeners, JF and PN show optimal listening. Subjects JZ, AN, JS, EG, AL, and AT showed a main effect for loudness with better performance in the loud condition for both right and left ears (Figure 1.2).

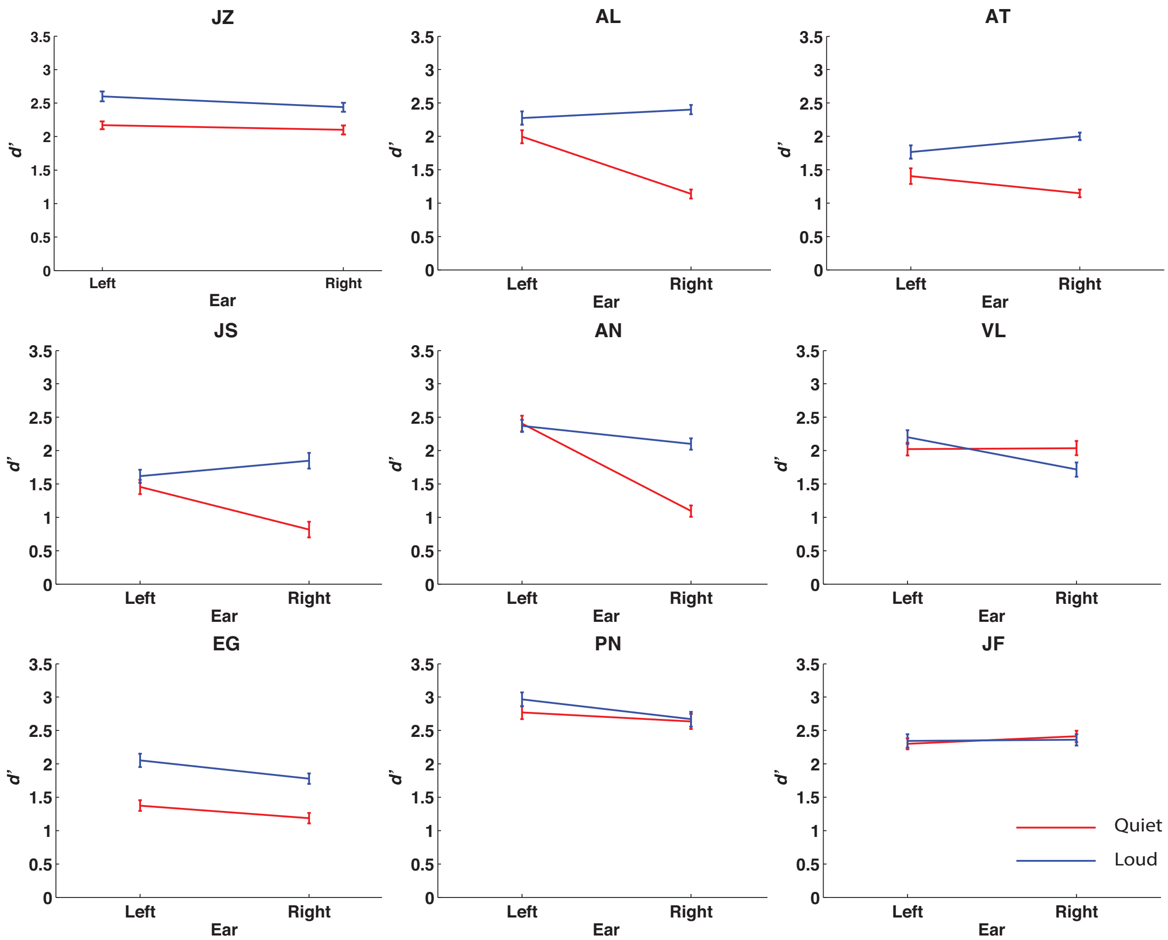


Figure 1.2: Individual d' values for Quiet and Loud conditions for both ears with error bars. Listeners PN and JF show no main effect of level or ear.

Interactions

Five subjects (VL, AL, AN, JS, and AT) showed a significant interaction between ear and level. For these listeners task difficulty exaggerated performance in the right ear such that performance was worse for hard conditions and better for easy conditions.

VL showed poor performance moving from the left to right ear for the loud condition but stable performance for the quiet condition. JS, AT and AL showed that moving from the left to the right ear resulted in significantly worse performance for the quiet condition and better performance for the loud condition. AN did worse for both quiet and loud conditions moving from the left to right ear.

JZ, EG, PN and JF showed no interaction between ear and level. PN and JF are the only two listeners who were not affected by ear or level of presentation and had no interaction between ear and level. These two listeners exhibit optimal listening behavior because they performed well for both loud and quiet conditions in the left and right ear.

1.5.1 Bayesian Interpretation

The standard statistical analysis of the GLM presented has a number of well-known limitations, stemming from a basic inability to represent uncertainty about inferences, or quantify evidence for and against hypotheses on the basis of data (e.g.,Morey et al., 2015; Wagenmakers, 2007). To address these deficiencies, and provide a complementary alternative statistical analysis testing the same hypotheses and data, we conducted a Bayesian analysis using the JASP program (Love et al., 2015).

JASP provides a number of Bayesian measures for evaluating the evidence for hypotheses in terms of data, including Bayes factors, which are the Bayesian gold standard (Kass and Raftery, 1995). Since our interest is in a small number of hypotheses—the presence or absence of two main effects and their interactions—we focused on the posterior model probabilities $\mathrm{P}(M \mid data)$. These probabilities effectively summarize the information provided by all of the Bayes factors between all pairs of models, under the (reasonable, in our case) assumption that the models considered exhaust the theoretically interesting possibilities.

Specifically, we considered five models— null, ear, level, ear and level, and interaction—that correspond to the various meaningful possibilities of the presence or absence of main effects and interactions. The null model corresponds to the possibility there is no effect of level being quiet or loud on the left or right ear. The ear model corresponds to a main effect of ear being left or right. The level model corresponds to a main effect of whether level was loud or soft. The ear and level model corresponds to both of these main effects applying independently. The interaction model corresponds to an interaction between the right and left ear with the quiet or loud level manipulation.

Our analysis assumes each of these models is a priori equally likely, and thus the posterior model probabilities reflect the evidence provided by the data for and against each. These posterior probabilities automatically take into account the goodness-of-fit and complexity of each of the models. They also naturally lie on the meaningful scale of probabilities, calibrated by betting. The posterior model probabilities for each subject analyzed individually, along with other output from the JASP program, are shown in Table 1.3-1.11. Probabilities less than one-millionth have been denoted by '-'.

JS, AL, AN, and AT had a posterior probability of the interaction model of 67% or greater. That is, of the five models, the data provide evidence that makes it 67% certain this is the best model for AT and 99% certain for JS, AL and AN. On this basis, with reference to Figure 1.2 to understand the direction of the effect that leads to the difference, we conclude that moving from the left to right ear results in better performance for the easy loud condition, while performance in the difficult quiet condition worsens significantly. This result suggests a shared pattern among all 4 listeners in which the performance in the quiet versus loud level, moving from the left to right ear, emphasizes the task difficulty level such that the quiet condition gets harder and the loud condition gets easier.

Table 1.3: Bayesian model for AL

Model	$P(M)$	P($M\mid$ data)	BF_M	BF_{10}	$\% error$
Null	0.2	-	5.21e-9	1.00	-
Ear	0.2	-	-	1.0	-
Level	0.2	2.86e-5	1.14e-4	21.9e3	0.802
Ear & Level	0.2	2.94e-4	0.001	22.6e3	2.36
Interaction	0.2	1.00	12.4e3	7.67e8	15.6

Table 1.4: Bayesian model for AN

Model	$P(M)$	P(M | data)	BF_M	BF_{10}	$\% error$
Null	0.2	-	-	1.00	-
Ear	0.2	-	-	44.2e3	1.44
Level	0.2	-	-	10.4	0.690
Ear & Level	0.2	1.73e-4	6.90e-4	6.90e6	6.03
Interaction	0.2	1.00	23.0e3	4.00e10	1.78

Table 1.5: Bayesian model for JS

Model	$P(M)$	P(M | data)	BF_M	BF_{10}	$\% error$
Null	0.2	-	-	1.00	-
Ear	0.2	-	-	0.717	1.11
Level	0.2	2.34e-4	9.36e-4	22.6e3	2.72
Ear & Level	0.2	2.98e-4	0.001	28.8e3	2.46
Interaction	0.2	0.999	7.52e3	9.68e7	4.45

Table 1.6: Bayesian model for AT

Model	$P(M)$	P(M | data)	BF_M	BF_{10}	$\% error$
Null	0.2	-	-	1.00	-
Ear	0.2	-	-	0.285	1.791
Level	0.2	0.256	1.38	45.8e4	1.77
Ear & Level	0.2	0.073	0.316	13.1e4	2.42
Interaction	0.2	0.67	8.14	1.20e6	10.3

VL also had a posterior probability for the interaction model of about 49 %, but also a 28 % posterior probability for the ear model. Thus, there is uncertainty as to which of these two models provides a better account of VL's performance, and it makes sense to consider how their performance would be interpreted in each case. Under the interaction model, VL shows a decrease in performance in the easy condition moving from the left to right ear, while performance in the difficult condition improved slightly. This is unexpected given that the loud condition is easier for most subjects. Under the ear model, VL shows better performance in her left ear than the right ear.

Table 1.7: Bayesian model for VL

Model	$P(M)$	P(M | data)	BF_M	BF_{10}	$\% error$
Null	0.2	0.101	0.450	1	-
Ear	0.2	0.275	1.52	2.72	1.42
Level	0.2	0.036	0.148	0.353	0.952
Ear & Level	0.2	0.102	0.455	1.01	3.14
Interaction	0.2	0.486	3.79	4.81	3.38

EG had posterior model probabilities of 63% for the ear and level model, and 21% for the interaction model. Once again, this result expresses the uncertainty that is a natural feature of Bayesian analysis. The data do not unequivocally support one model over all others, but are consistent with two of the models. The ear and level model is most likely, but there is also some lesser evidence for the interaction model. Under the ear and level model, better performance is seen in the left ear for both conditions, while under the level model better performance is seen in the loud rather than quiet conditions for both ears.

Table 1.8: Bayesian model for EG

Model	$P(M)$	P(M | data)	BF_M	BF_{10}	$\% error$
Null	0.2	-	-	1.00	-
Ear	0.2	-	-	1.11	0.813
Level	0.2	0.160	0.764	1.02e6	1.735
Ear & Level	0.2	0.626	6.70	3.99e6	8.53
Interaction	0.2	0.213	1.08	1.36e6	2.53

JF had a posterior model probability of 61% for the null model. This provides strong, but not completely conclusive, evidence for the lack of any effect of level on her right or left ear. In other words, the most likely account is that the listener is stable in performance for quiet or loud conditions in both ears. This further shows listener JF can selectively attend to both her right and left ear equally well.

JZ had posterior probabilities of 49% for the level model and 38% for the ear and level model. The level model shows the loud condition resulted in better performance for both ears. The

Table 1.9: Bayesian model for JF

Model	$P(M)$	P(M | data)	BF_M	BF_{10}	$\%\,error$
Null	0.2	0.608	6.197	1.0	-
Ear	0.2	0.188	0.925	0.309	1.03
Level	0.2	0.144	0.671	0.236	0.816
Ear & Level	0.2	0.045	0.187	0.073	1.87
Interaction	0.2	0.016	0.066	0.027	2.54

Table 1.10: Bayesian model for JZ

Model	$P(M)$	P(M | data)	BF_M	BF_{10}	$\%\,error$
Null	0.2	-	-	1.0	-
Ear	0.2	-	-	0.578	1.11
Level	0.2	0.493	3.89	46.7e4	0.973
Ear & Level	0.2	0.386	2.52	36.6e4	1.41
Interaction	0.2	0.121	0.549	11.4e4	2.86

ear and level model include that performance is slightly better in the left ear for both quiet and loud conditions.

PN had the largest posterior probability of 39% for the ear model, but also had 19% posterior probabilities for the null model and the ear and level model. Thus, the data are quite ambiguous as to the best account of performance. Under the ear model, the left ear has a better performance than the right ear for both loud and quiet conditions. This can be seen in Fig 1.2.

Table 1.11: Bayesian model for PN

Model	$P(M)$	P(M | data)	BF_M	BF_{10}	$\%\,error$
Null	0.2	0.190	0.94	1	-
Ear	0.2	0.391	2.57	2.06	1.01
Level	0.2	0.091	0.401	0.479	1.77
Ear & Level	0.2	0.195	0.968	1.02	1.99
Interaction	0.2	0.132	0.611	0.696	31.1

These analyses show that at least half of the listeners in this group are under the interaction model where there is better performance in the left rather than the right ear. While there are clear individual listening patterns that exist here, there is overwhelming evidence that listening ability between the two ears is not equal for all listeners, as one would expect. Two additional listeners, EG and PN, under the ear model showed a better performance in the left rather than right ear, while JZ was most affected by the loud level manipulation. Only one listener, JF, was found to be under the null model, indicating that she had the ability to selectively attend to both ears equally well while others showed a marked preference for the left. Future work could amass enough subject data to create a Bayesian hierarchical model that could predict listener weights with minimal data collection hours. This would be extremely beneficial, specifically, in a clinical setting as a way to effectively screen listeners.

1.6 Experiment 2: Effect of longer inter-tone interval

1.7 Introduction

The goal of this manipulation is to improve performance for listeners that gave the most weight for the uninformative tones at 1, 3, 5, and 7 rather than the informative, event tones in the Right Quiet and Left Quiet conditions. The ITI selected here is 300ms due to findings from (Berg, 1990) which showed that there is no difference in listener performance when tones are between 50 ms - 200 ms ITI. Listener AN is noted to have mild hearing loss at low frequencies in her right ear that explains the poor performance in both RQ and RL observed in Experiment 1. It was of interest to observe if a longer ITI could improve AN's performance.

1.8 Methods

Differences in temporal processing are examined here with listeners selected by performance from Experiment 1 to participate in Experiment 2. Listener AT, AL and EG shared similar listening patterns for condition RQ and LQ (see Figure 1.1). Both showed peak decision weights for odd tones instead of even tones in the RQ condition while both listeners' weights in LQ showed a precedence effect in which greater weight was placed for the first 2 tones and a downward flattened slope for the rest of the tones, indicating listeners were unable to attend to weight tones 3-7 optimally. Listener AN only showed non-optimal weighting for RQ condition.

Listeners were presented a sequence of seven tones drawn from the same distribution as follows from Experiment 1. The task is still to judge whether the stimulus on a given trial was a high-mean or a low-mean stimulus. In this follow-up experiment, the inter-tone interval was presented at 300 ms.

1.9 Results

Decision weights from AT and AL shifted from the initial peak weights at the odd numbered tones observed in Experiment 1 to peak weights on the even tones when the tones were separated with a 300 ms ITI. Both listeners' performance measure reflected the optimal weight shift with d' values that increased by a factor of 2. This finding shows evidence contrary to (Berg, 1990) using a diotic presentation that indicated a presentation rate between 50 ms - 200 ms that had no effect on shifting the weights to back to optimal.

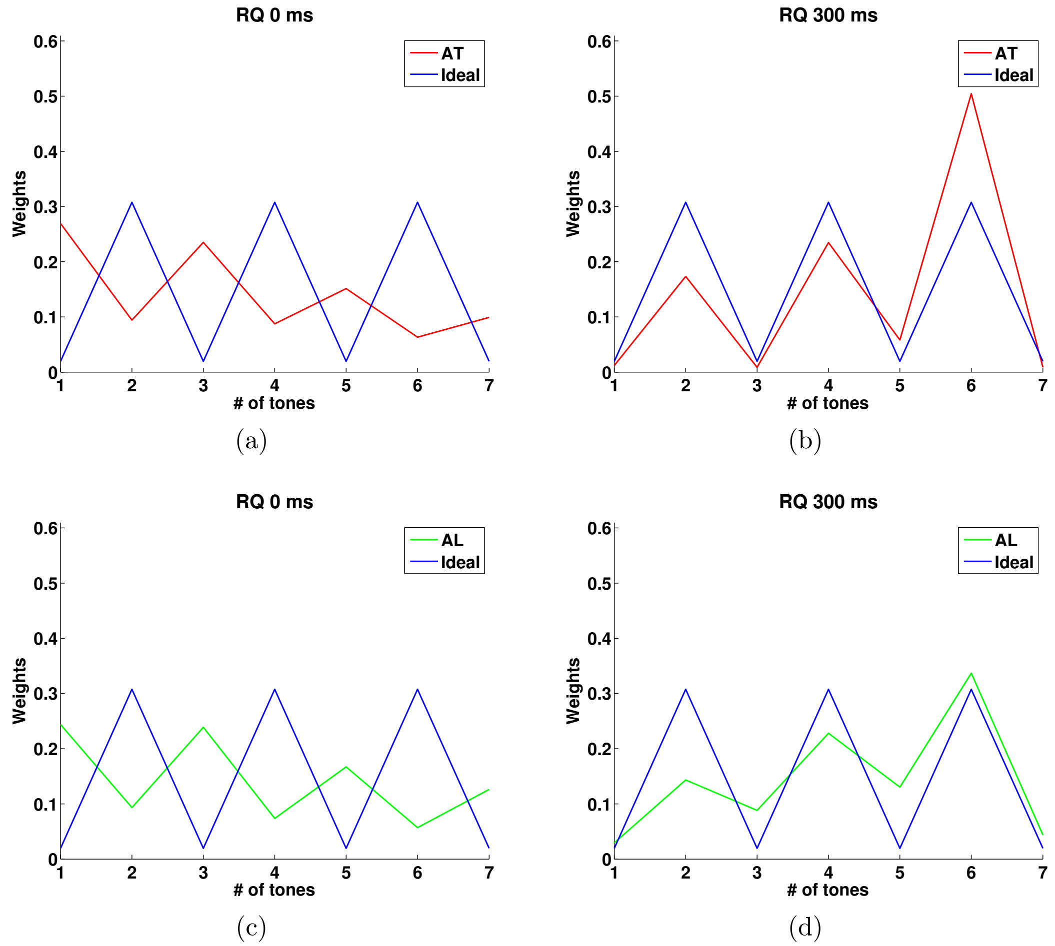
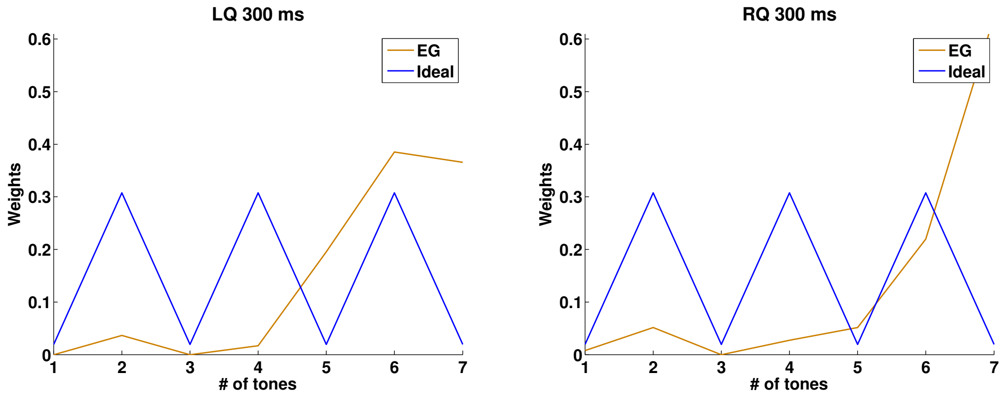


Figure 1.3: Weights for condition RQ at 0 ms and 300 ms for subjects AT and AL.

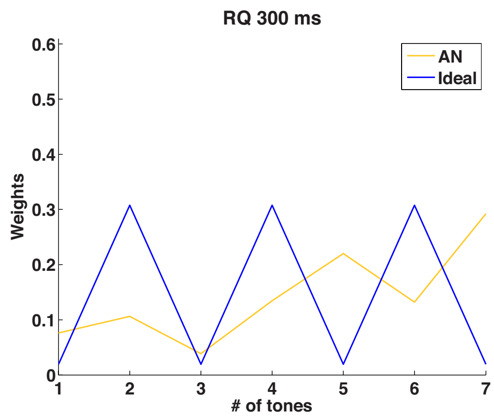
However, estimated weights observed from EG and AN did not reflect the same weight shift as AT and AL. Figure 1.4 shows that when ITI is 300 ms for both LQ and RQ conditions, EG's weights were worse than in Experiment 1 as seen in Figure 1.1 b and d. EG's d' for LQ and RQ at 300 were, 1.18 and 0.735 respectively.

Listener AN was noted previously to have mild hearing loss at low frequencies. It was of interest to see if a longer ITI would help AN in this particular case. However, Figure 1.5 does not show marked improvement from Experiment 1.



(a) (b)

Figure 1.4: Weights for condition LQ and RQ at 300 ms for subject EG.



(a)

Figure 1.5: Weights for condition RQ at 300 ms for subject AN.

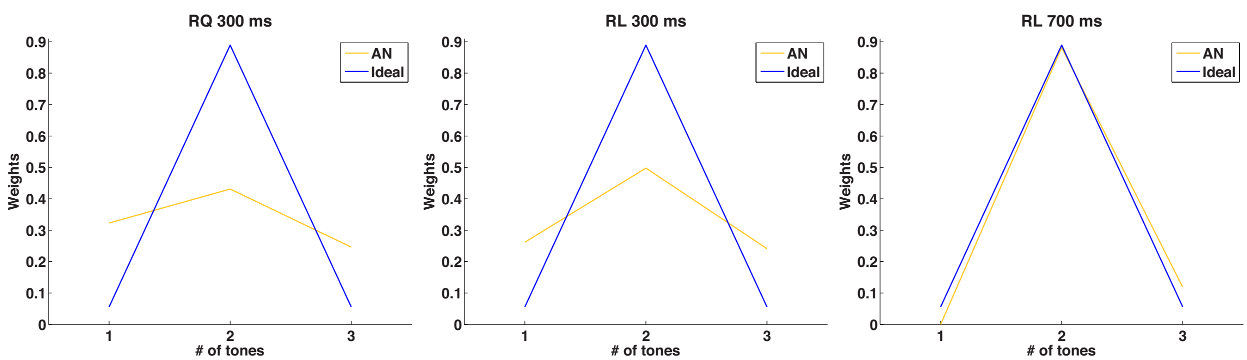
1.10 Experiment 3: Effect of fewer tones and longer inter-tone interval

1.10.1 Methods

In this experiment, fewer tones were employed for 2 listeners who had difficulty in the Experiment 2. Since AN and EG did not show improvement with a longer ITI for seven tones, it was of interest to determine if seven tones was too much of a cognitive load. (Turner and Berg, 2007) found that in instances with large recency effect in which listeners have most weight on the last presented tone, it's possible that using fewer tones may reduce cognitive load to improve performance. Two listeners, EG and AN, from Experiment 1 were recruited to do a follow-up experiment based on their initial performance in Experiment 1.

1.10.2 Results

Findings show that reduced cognitive load does not improve subject weights as evidenced by the d' values for EG and AN. Figure 1.6 shows that even with a 300 ms ITI and reducing the number of tones did not improve AN's weights. Only when ITI was at 700 ms ITI for the easy condition RL did AN's weights approach optimal weights. This suggests that reducing the number of tones and increasing ITI is not enough to improve a listener who has low frequency hearing loss.



(a) (b) (c)

Figure 1.6: Weights for condition RQ and RL at 300 ms and RL at 700 ms, with 3 tones for subject AN.

Figure 1.7 shows decision weights from EG that indicate better performance for LQ at 300 ms ITI than RQ. Reducing the number of tones and increasing the ITI to 300 ms was not enough for EG to fully weight the central component. The ITI was increased to 500 ms before EG's weights were comparable to the optimal weights.

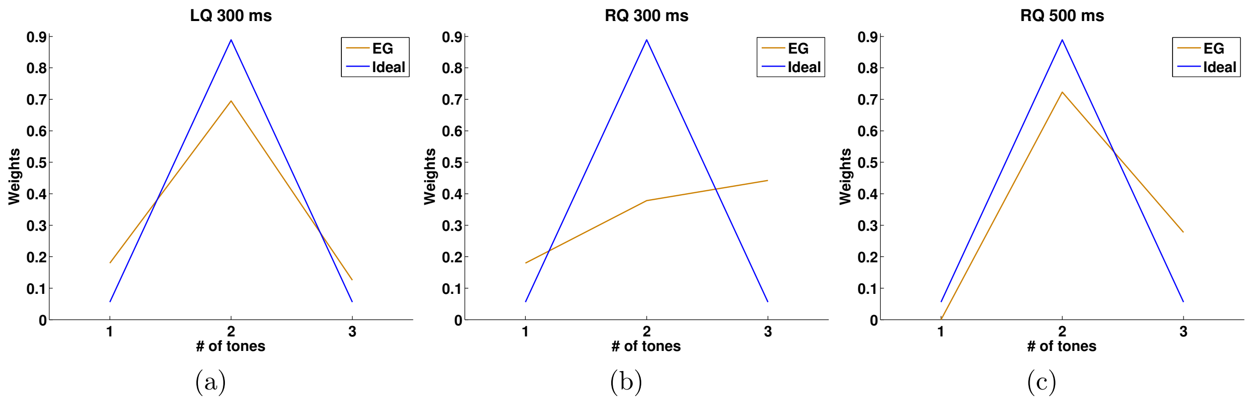


Figure 1.7: Decision weights for condition LQ and RQ at 300 ms with 3 tones for subject EG.

1.11 Experiment 4: Dichotic vs Diotic listening

In this experiment, comparisons between dichotic and diotic listening are examined with three listeners recruited from Experiment 1, JZ, JF, and AT. Due to the differences listening patterns found in Experiment 1 and the improvement of performance with a longer ITI, it is of interest to determine if there are temporal differences that exist between diotic listening and dichotic listening..

1.11.1 Methods

Listeners were selected based on performance in Experiment 1. The dichotic condition (RQ or LQ) that had the lowest d' was selected as the condition that would be tested with longer ITI relative to the diotic listening. The diotic listening conditions involve presenting quiet (DQ) and loud (DL) tones to both ears simultaneously. Unlike conditions in dichotic listening, both ears receive all seven tones.

1.11.2 Results

Listeners show varying ability to attend to the most informative tones in diotic presentation. Increasing the ITI, however, improves performance across all listeners, with most listeners being able to attend to the informative tones once the ITI is increased to 500 ms. More interesting, however, are the differences between dichotic and diotic listening as the ITI is increased. All listeners show improvement at shorter ITIs in dichotic listening which implies a faster processing mechanism.

JZ

The first panel in Figure 1.8a shows that JZ placed greater weight for the even-numbered tones more in the RQ condition at 0 ms. This suggests that JZ is able to attend to the more informative tones in this condition. The second panel in Figure 1.8a, shows peak weights for the odd-numbered tones in DQ condition at 0 ms. This suggests that JZ was unable to attend to the more informative tones. The third panel in Figure 1.8a shows a precedence effect where JZ placed the most weight on the initial tone for the DL condition at 0ms. JZ's performance was better in the dichotic conditions than the diotic listening conditions at 0 ms. Additionally, both decision weights and performance measures (d') for RQ, DQ, and DL at 300 ms were similar.

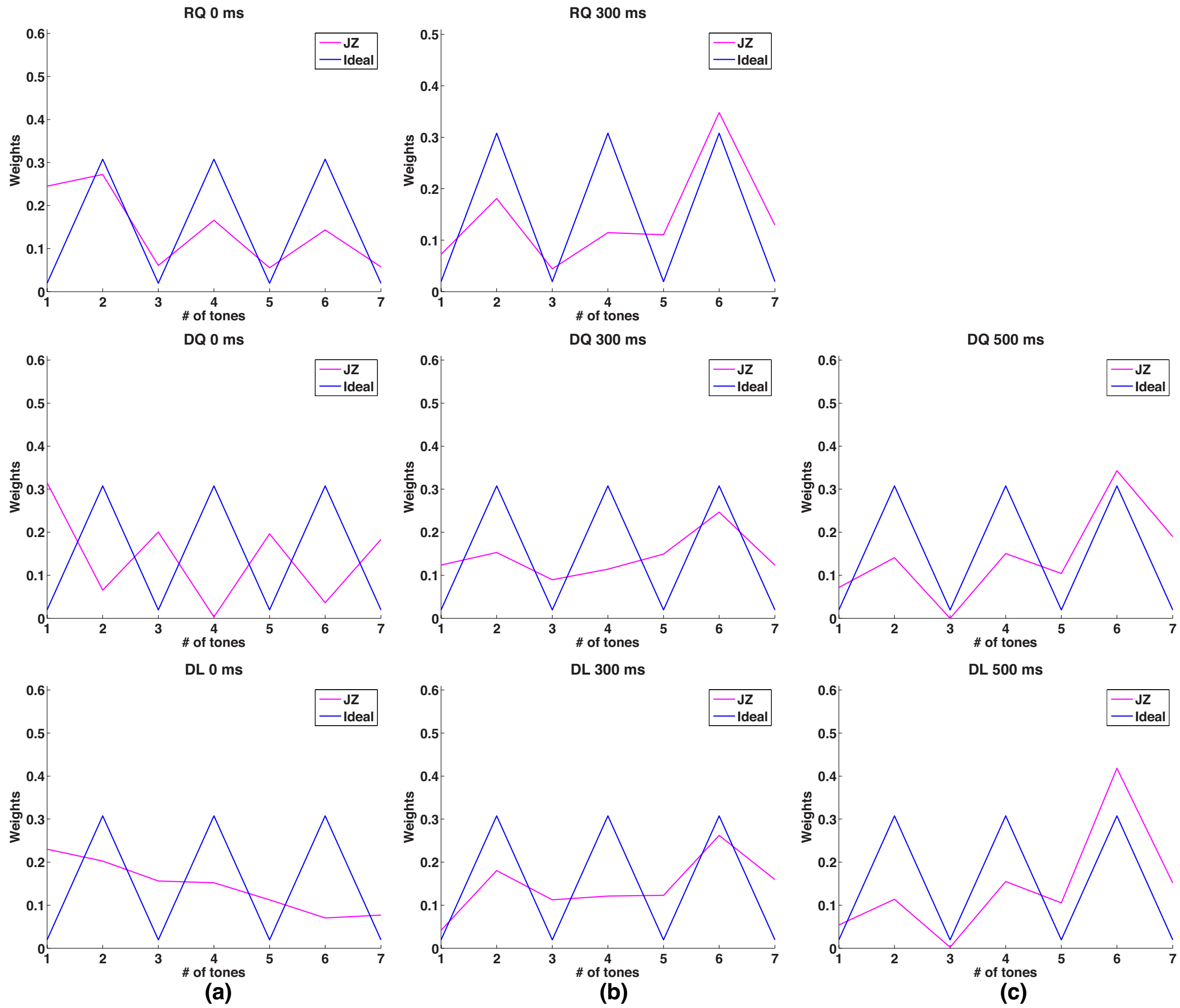


Figure 1.8: Listener JZ's decision weights. Top row represents dichotic condition RQ at 0 ms and 300 ms. The second and third row represent diotic listening conditions, diotic quiet (DQ) and diotic loud(DL) at 0 ms, 300 ms, and 500 ms.

Because JZ can attend to the target tones that are more informative in the dichotic condition at 0 ms but not for the diotic listening conditions at 0 ms, this suggests that temporal differences exist. This is also reflected in performance measures shown in Table 1.12. DQ and DL conditions only showed improvement when ITI was increased to 500 ms as shown in Figure 1.8c.

Table 1.12: JZ Performance measure d' in dichotic and diotic listening

	Dichotic		Diotic	
Time(ms)	Right Quiet	Right Loud	Quiet	Loud
0	2.12	2.42	0.769	1.99
300	2.00	-	2.04	2.18
500	-	-	1.92	2.13

AT

Figure 1.9a shows that AT weighted odd-numbered tones more than the informative, even-numbered tones in RQ or DQ and DL at 0 ms. This implies AT was unable to attend to the more informative tones in dichotic or diotic listening. However, when ITI is increased to 300 ms, AT showed weights shifted back to the more informative tones in RQ while DQ and DL resulted in an apparent recency effect with the most weight on the last, informative tone. Performance (d') was better in the dichotic RQ condition at 300 ms than DQ and DL as evidenced by Table 1.13. When ITI increased to 500 ms, decision weights were observed on the even tones for dichotic and diotic listening conditions with a better d' for RQ as seen in Table 1.13.

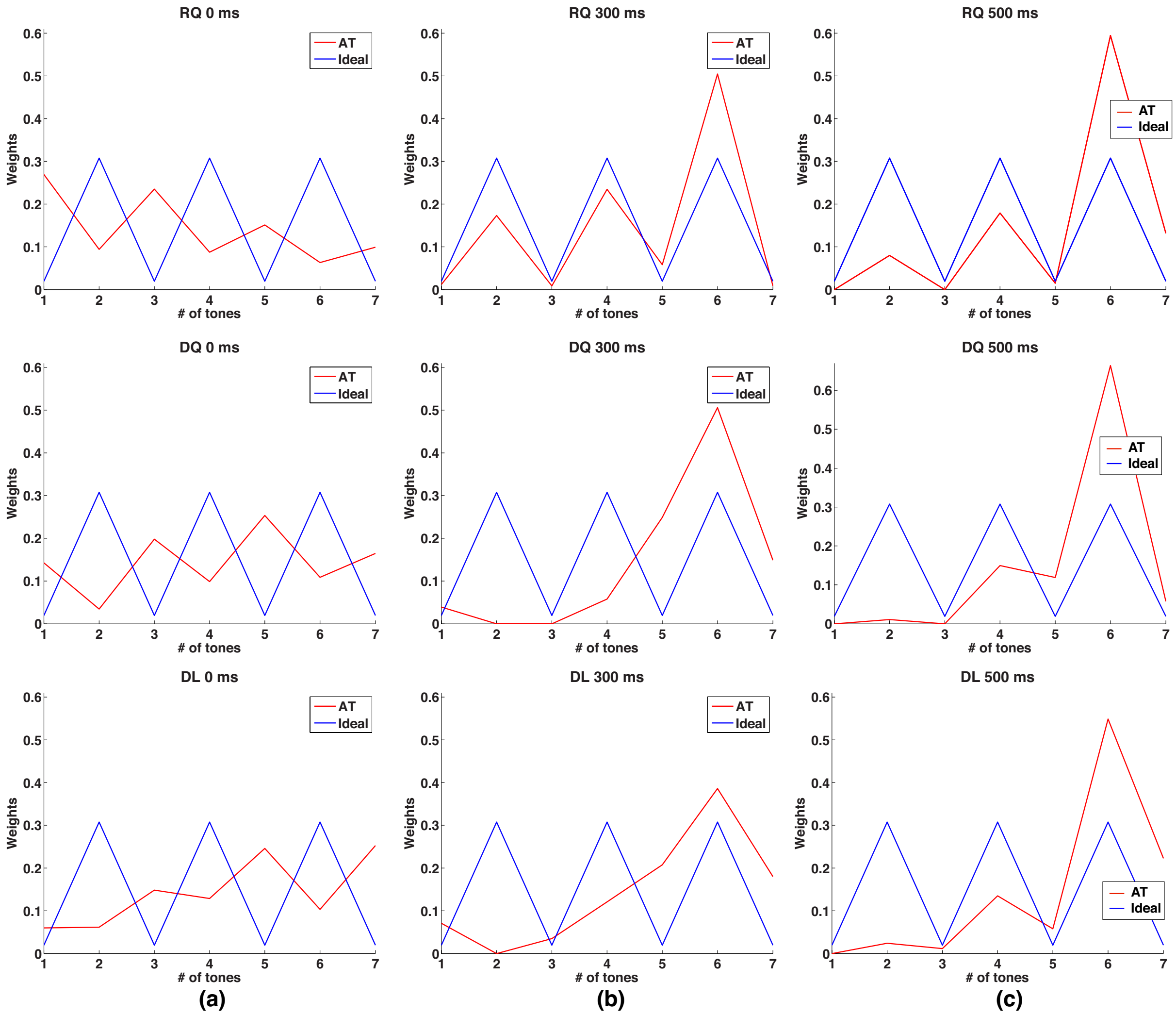


Figure 1.9: Listener AT's decision weights. Top row represents dichotic condition RQ a 0, 300 and 500 ms. The second and third row represent diotic listening conditions, diotic quiet (DQ) and diotic loud(DL) at 0 ms, 300 ms, and 500 ms.

Table 1.13: AT Performance measure d' in dichotic and diotic listening

	Dichotic		Diotic	
Time(ms)	Right Quiet	Right Loud	Quiet	Loud
0	0.95	2.02	1.10	1.32
300	1.90	-	1.33	1.31
500	1.64	-	1.38	1.42

JF

JF has shown from Experiment 1 that she is able to attend to the more informative tones in easy and hard conditions equally. Figure 1.10a shows that JF's decision weights are better at LQ than DQ for an ITI at 0 ms. This is reflected in listener's performance measure, d' as noted in Table 1.14. This suggests that even for a superior listener, performance in dichotic listening was better than diotic listening conditions when the more informative tones were quiet (i.e.: harder task). Also, when the more informative tones in the diotic condition are loud (DL), performance was then comparable to that of dichotic LQ condition.

Table 1.14: JF Performance measure d' in dichotic and diotic listening

	Dichotic		Diotic	
Time(ms)	Left Quiet	Left Loud	Quiet	Loud
0	2.38	2.02	1.88	2.35
300	2.36	-	2.39	2.12
500	-	-	-	-

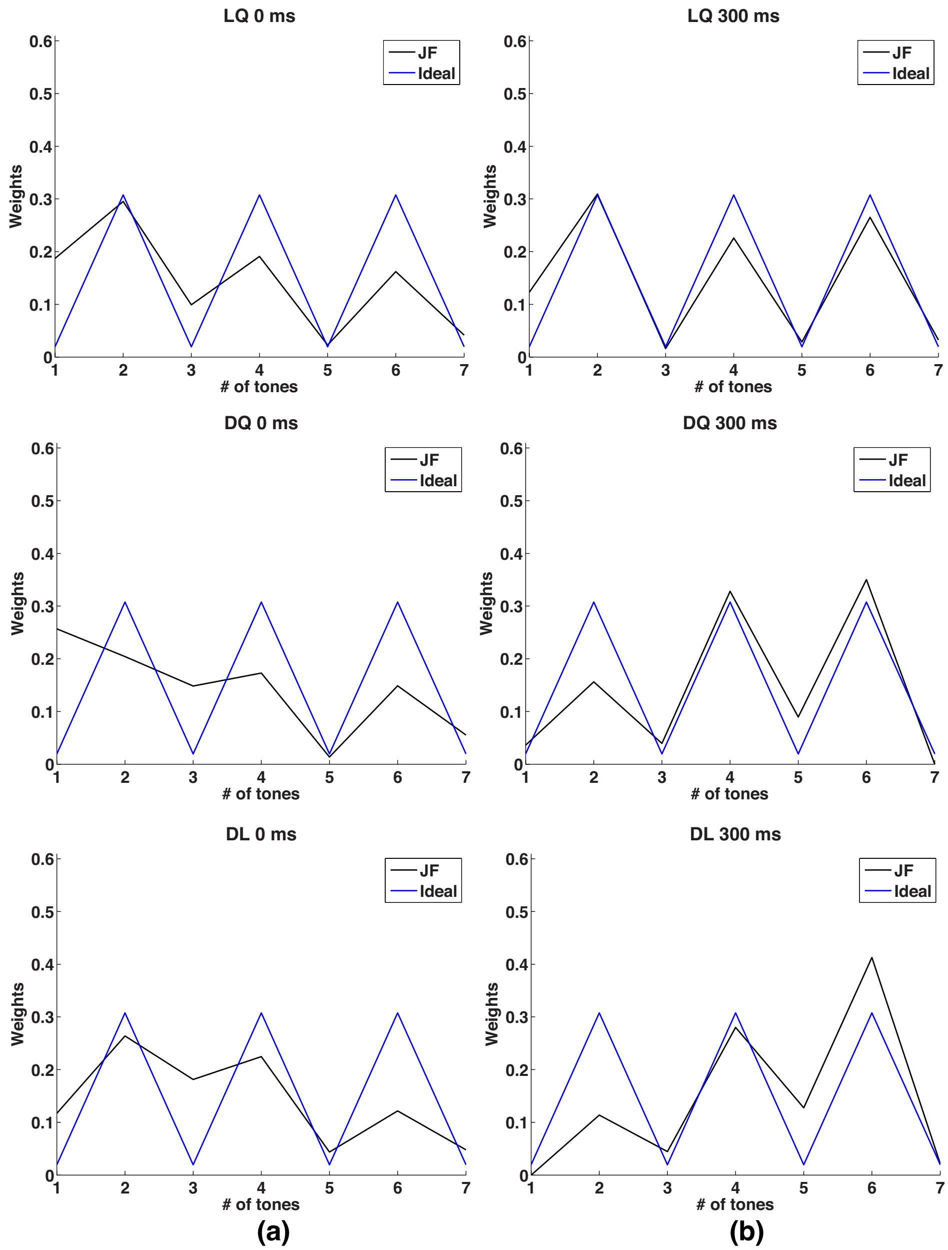


Figure 1.10: Listener JF's decision weights. Top row represents dichotic condition LQ a 0 and 300 ms. The second and third row represent diotic listening conditions, diotic quiet (DQ) and diotic loud(DL) at 0 ms and 300 ms.

1.12 Discussion

There is a limited framework with which to discuss auditory selective attention. Results from this study show a wide range of individual differences for decision weights, efficiency measures and performance (d') in listener ability to attend to the most informative tones. The findings show that one's ability to selectively attend to information in a target ear varies across individuals, which is unexpected. The ANOVA results showed that only two listeners showed no bias. The other seven listeners show a distinct ear preference while five listeners out of the seven demonstrated an interaction between ear and level. These non-optimal listeners achieved better performance (d') and decision weights only when the informative tones were presented to a particular ear.

Using decision weights, d' and efficiency (η_{wgt}), we developed a psychophysical approach for assessing how well listeners are able to attend to a specific ear. Estimated decision weights in Figure 1.1 help to visualize individual listening patterns. For example, Figure 1.1b and 1.1e show five listeners using non-optimal weighting strategies, i.e., show a bias towards loudness or ear. These listeners gave greater weight to the odd tones as noted by the peaks at 1, 3, 5, and 7, instead of attending to the more informative, even tones. Conversely, Figure 1.1a shows four listeners placing greater weight on the more informative, even tones. These listeners are able to attend to the target ear.

The results from both Bayesian analyses and GLM ANOVAs corroborate findings that most listeners exhibit a left ear superiority, i.e., they perform better when attending to the left ear. This suggests the presence of an ear advantage for the left ear to tonal stimuli. This contrasts with the right ear advantage (REA) first reported in Kimura (1961) where the right ear outperformed the left ear for verbal stimuli. Although REA has been further supported for verbal stimuli (Ahonniska et al., 1993; Berman et al., 2003; Studdert-Kennedy and Shankweiler, 1970) definitive psychophysical evidence of an ear advantage has been largely unexplored. Currently, there are limited methods available to study what gives rise to a left ear advantage in relation to tonal stimuli and the interactions it has with selective attention.

Zatorre and Belin (2001) proposed a hemispheric network model in which the right hemisphere is more involved in the spectral processing while the left hemisphere is responsible for rapid temporal processing. This suggests that the core auditory cortex responds bilaterally to temporal changes and hemispheric specialization. This would support the idea of a right ear advantage for verbal stimuli and a left ear advantage for tonal stimuli.

Other findings from this chapter highlight patterns of individual differences in listening capability. Different groups of listeners exist and not all are equally capable. A closer investigation in dichotic listening showed that for some listeners benefited dramatically from an increase of 300 ms ITI. Even more surprising were the findings for AN and EG, listeners who did not improve with a longer ITI. These listeners, in Experiment 3 and 4, did not show improvement with an increased 300 ms ITI. Only when the number of tones was decrease did they yield a better d' and improved weights. While AN's improvement with three tones required the longest ITI, it can be partially explained by hair cell damage as noted by DPOAE tests. For EG, improved performance differed from the left to right ear. While the left ear (LQ) improved with three tones and an ITI increase of 300 ms the right ear did not show comparable improvement until ITI was at 500 ms.

In a greater context, this chapter also finds evidence that temporal differences exist between dichotic and diotic listening. Interestingly, Turner and Berg (2007) found that with a diotic sample discrimination task, the ITI that resulted in listener weights and d' to converge for quiet and loud tones are much longer between 500 - 700 ms, with a large recency effect present. Here, we see that most listeners clearly converge before or at 500 ms with decision weights that indicate much less ambiguity that those found in Turner and Berg (2007). Perhaps this difference is due to the length of the stimulus (five tones versus seven tones) and the additional information that might be available to the listeners with seven tones. Future studies might explore convergence as a factor of tone length as a measure of information content.

By no means is the pretense here a novel model of auditory attention. However, the results provide a new methodology for assaying whether or not listeners are able to selectively attend to a target stimulus and the temporal differences in listening presentations. This presents a way to discuss a listener's ability to attend in a psychophysical paradigm in a concrete way. Future work using Bayesian modeling could predict listener strategies for ear preference, level effects, and ability to isolate information in noisy environments.

Chapter 2

Manipulation of cues in narrow band stimuli

2.1 Overview

When listeners discriminate spectral changes in narrow-band sounds, different cues are used by subjects that have distinct patterns of decision weights (Southworth and Berg, 1995). The cues that listeners use are described here as different dimensions of sound: pitch, timbre (or what is referred to as roughness here), and loudness. Southworth and Berg (1995) showed that listeners can selectively attend to a specific cue out of the three available cues when the same stimuli is presented. Chapter 2 investigates the effects of stimulus manipulation that can effectively switch listener perception between these three listening cues and develop a procedure for individuals who have difficulty with certain cues. Individuals demonstrate use of two listening strategies and psychophysical measures are developed to observe the effects for listeners who do not match model predictions. It is of interest then to establish listeners can attend to each of the three cues and how to manipulate the stimulus. This investigation also broaches the topic of perceptual consciousness by providing a means to describe a perceptual experience by observing the pattern of decision weights.

2.2 Introduction

This study investigates the influence of selective attention on the auditory system using narrow-band spectra. Specifically, the extent with which a listener can attend to a particular cue in a discrimination task where there are competing cues. It has been shown that individuals exhibit different patterns of listening for different dimensions of sound, like pitch, timbre and loudness. These dimensions of sound are referred to as cues. Essentially, listeners can attend to one cue over others and change between cues. The advantage here is that distinct single channel models for pitch, roughness and level show that cues are available. This is reflected in patterns of listening strategies that indicate individuals are in fact changing their listening strategy (Southworth and Berg, 1995).

While it is difficult for listeners in auditory research studies to verbally describe different dimensions of sound in a quantifiable way, this method broaches the discussion of how the auditory process is translated into perceptual experience. This provides insight into the mechanisms responsible for how individuals orient their listening and has implications for the dialogue of auditory attention.

Previous research has shown that listeners can discriminate cues in narrowband sounds relating to level (Zwicker et al., 1957), pitch (EWAIF; Feth, 1974; Feth and O'Malley, 1977) and timbre or "roughness" (Green et al., 1992; Richards, 1992). While studies will often isolate one cue and limit other cues to avoid influencing the target cue– for example, degrading the level cue by using a roving-level procedure as to limit a strategy that would detect overall level of the stimulus (Gerald Kidd et al., 1989)– findings from Southworth and Berg (1995) showed that individuals could discriminate between three perceptual cues using the same tone configuration. This is a notable finding because it suggests that listeners have distinct decision weights with the simultaneous availability of these cues (cf. Ahumada and Lovell, 1971). Specifically, these decision weights map onto different perceptual experiences using the same stimulus configuration. The different decision strategies observed are based on verbal direction to subjects to attend to one cue over others. This aspect of selective attention is of interest especially in the context of using cues that are simultaneously available.

Berg et al. (1992) estimated decision weights for three-tone spectra with a roving-level procedure to degrade cues based on overall level and revealed a distinct pattern of weights dependent on bandwidth. The signal component for all conditions was an intensity increment added to the center frequency at 1000 Hz. Three ranges of of bandwidth –wider, intermediate, and narrower– were used in the study and the nonsignal components were symmetric about the center frequency. Three distinct patterns of weights were observed when the bandwidth was changed from very wide to very narrow.

This suggests that different auditory mechanisms may have optimal ranges of processing. An asymmetric pattern of weights at intermediate bandwidths suggest the use of pitch cues. Instances of bandwidths that are wider than a critical band would require across channel comparisons (i.e.: profile analysis; Green, 1988). Specifically, listeners were able to discriminate narrowband spectra predominately on the basis of pitch, with the signal sounding higher in pitch, between tones with a $\Delta f = 40$ and $\Delta f = 80$ Hz apart. When tones were separated by 10 Hz, however, listeners showed symmetric weight patterns and reported that the signal sounded smoother relative to the standard, which are consistent with the model. Using model predictions corroborates patterns of weight are distinct for pitch or roughness (envelope differences) for these narrowband stimuli.

While Berg et al. (1992) used a roving level procedure, Southworth and Berg (1995) made the level cue available. Figure 2.1 adapted from Southworth and Berg (1995) shows how decision weight estimates changed as listeners were instructed to listen to the three different cues. Because stimuli lie within a critical band, the effect of differences in the patterns of weight estimates is what is of interest rather than a change in relative weight for a specific component.

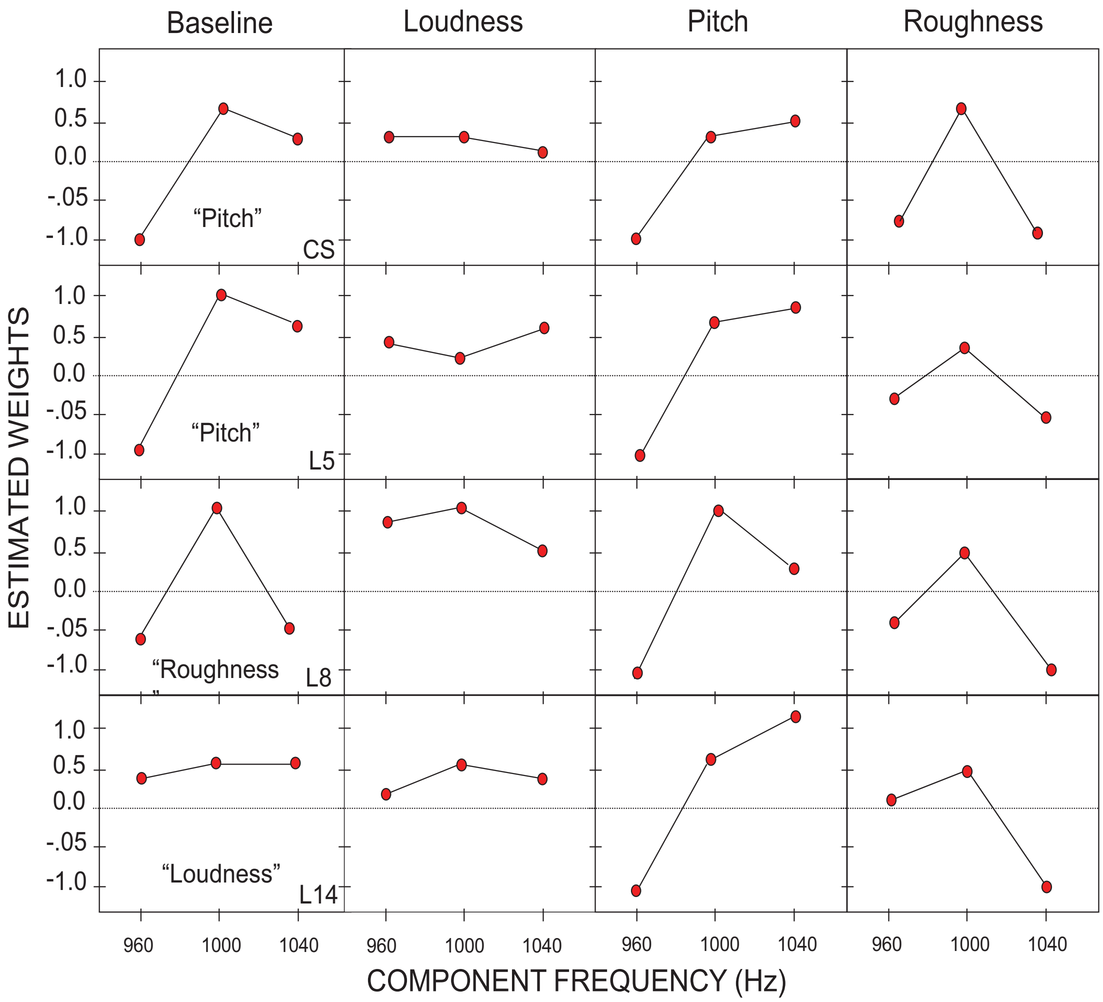


Figure 2.1: Adapted findings from Southworth and Berg (1995). Panels across represent each participant. Estimated decision weights from four listeners that reflect distinct patterns for pitch, loudness and roughness. Listeners were not given specific instructions to a particular cue and were asked to indicate what interval contained the signal. From the baseline condition, the listeners' preferred cue is represented.

Although Southworth & Berg showed that listeners could discriminate between three different cues, they also found listeners who had difficulty with certain cues that resulted in decision weights that did not follow the weight pattern predicted by the model for the roughness discriminations. This can be observed from adapted findings from Southworth and Berg (1995) as well as the current findings from this study in Figure 2.7. Two explanations were possible; either listeners have difficulty with the cue or the envelope model is inconsistent with the listener strategy when attending to the cue. Southworth and Berg resolved this issue with a post-hoc condition that showed when information from other cues (i.e.: pitch and level) are degraded, the observed pattern of weights from the listeners corresponded more closely with the predicted by the model.

It is of interest here to identify what stimulus manipulations can directly affect listener perception. With the ability to isolate and introduce one cue at a time, this allows a window in discerning how listeners' weights change as different cues are introduced and how this translates to the degree of difficulty listeners may have with different perceptual cues.

2.2.1 Three Models of Narrow-band Sound Discrimination

Three models based on absolute level, envelope-weighted average instantaneous frequency (EWAIF) and power spectrum of the envelope are used here to represent level, pitch and roughness discriminations respectively. Using computer simulations for each model, a decision statistic is calculated for each trial. Estimates of thresholds are from reversal points of 5000 simulated trials.

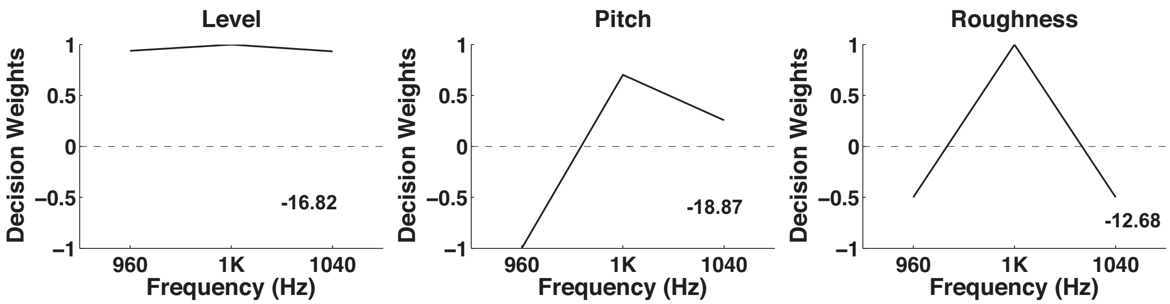


Figure 2.2: Model weights calculated from computer simulations for level, roughness and pitch cues. Threshold estimates for each model are shown in the lower right-hand corner of their respective panels and can be considered as the ideal threshold for the level model. The pitch model is more sensitive with a steeper filter and a more sensitive decision statistic may exist for the envelope. Optimal thresholds are relative and depends on the performance of a well-specified model.

The level model calculates the total power over the stimulus complex by integrating the spectrum $f(t)^2$ over the duration T_d,

$$\int_0^{T_d} f(\mathrm{t})^2 dt\,. \tag{2.1}$$

The power of each stimulus component is positively correlated to the decision variable. The signal is added to the second component of the first interval. An incremental change in the intensity of any of the components is perceived as louder and increases the probability of reporting a signal. Therefore, all the weights are positive as shown in the leftmost panel of Figure 2.2. In simulations, the level model responds signal in the first interval when,

$$\sum_{i=1}^{3}(x_{i1} + L_1) > \sum_{i=1}^{3}(x_{i2} + L_2)\,, \tag{2.2}$$

where, x_{i1} and x_{i2} represent the perturbations of the ith component in the first and second intervals. L_1 and L_2 are the levels of the tone-complexes presented in the first and second intervals. All three components in the level model are positive in sign and equal in magnitude.

The estimated threshold for the level model is -16.82.

Discriminations made on the basis of pitch are approximated using Feth's envelope-weighted average of the instantaneous frequency model (Ewaif; Feth, 1974; Feth and O'Malley, 1977) where

$$\frac{\int_0^T \mathrm{instf}(t)\mathrm{env}(t)dt}{\int_0^T \mathrm{env}(t)dt}\,, \tag{2.3}$$

env(t) is the envelope of the temporal waveform and instf(t) represents the instantaneous frequency of the wave form at time t. The EWAIF value is a calculation of the spectral pitch from a complex waveform. However, using the EWAIF equation directly will not provide differences in mean EWAIF value between signal and standard. Berg et al. showed that when two non-signal components are symmetric about the signal component of 1000 Hz, the mean predicted EWAIF value for the pitch of the signal and the standard remain unchanged at 1000 Hz, even when the signal is added to the standard. To account for observed data, the stimulus is passed through the skirt of an off-frequency filter before the EWAIF value is calculated. Off-frequency listening yields better performance and accounts for the asymmetric weights that were observed (Southworth and Berg, 1995). It suggests that information about pitch can be acquired in the skirt of the filter (EWAIF-OFL; Leshowitz and Wightman, 1971; Eckstein, 2011).

The EWAIF-OFL model first attenuates each component of the three-tone complex by using a filter centered below the complex with a 70 dB per octave, linear skirt. The interval that has the higher calculated EWAIF value from (2.3) is then selected as the signal. This decision statistic results in an asymmetric weight pattern with a negative sign for the 960 Hz component and positive signs for both the center component at 1000 Hz and at 1040 Hz. The estimated threshold for the model is -18.87.

The decision statistic for the third model is based on changes in the shape of the power spectrum of the envelope such that the differences in the envelope are perceived as perceptual changes in roughness' or smoothness of the narrow-band stimulus (Green et al., 1992). The spectrum of the Hilbert envelope is extracted for both the signal and the standard. Green's model takes the mean difference (η_i - τ_i), where η_i and τ_i are the expected values of the power spectrum for signal and the standard, respectively. A modification was made to Green's model since they used equal amplitude stimuli where mean power of the spectrum of the envelope was essentially equal to the standard deviation σ of the power at each envelope frequency. Since perturbations are added in order to collect weights in this current study, the amplitudes of the current stimuli have more variance and is better accounted for with an adjustment take the mean difference and then divide by the variance. The resulting decision rule for choosing signal in interval one is

$$\sum_{i=1}^{2}\left(\frac{\eta_i - \tau_i}{\sigma_i}\right) z_{i1} > \sum_{i=1}^{2}\left(\frac{\eta_i - \tau_i}{\sigma_i}\right) z_{i2}\,, \tag{2.4}$$

where η_i and τ_i are the power spectrum coefficients at Δ_f and $2\Delta_f$ for the intermodulation frequencies of the complex for the standard and the signal respectively. z_{i1} and z_{i2} represent the observed vectors on the normalized envelope power and σ_i is the variance of the standard. The greatest magnitude is on the center component and is positively signed with the 2 side components at 960 and 1040 Hz negatively signed at -0.5. Estimated threshold using this envelope model is -12.68. Investigations also revealed that there were no tonal interactions between components that went beyond the models described here (see Appendix B).

2.3 General Methods

2.3.1 Listeners

Eight students from the University of California, Irvine participated as listeners for these tasks as research volunteers and for research credit. The author (AT) participated as unpaid volunteer. All listeners were audiometrically tested and showed less than 20 dB hearing loss between 0.125 to 8 kHz. Listeners age ranged from 20-26 years. Naive listeners were at least 20 hours in narrowband spectral discrimination tasks. AT has had previous experience listening to narrowband sounds.

2.3.2 Stimuli

The standard comprised of three equal-amplitude sinusoids centered at 1000 Hz with side components that were manipulated either with a Δf at 40 Hz apart or 10 Hz apart depending on the type listening cue that was targeted. The signal was added in phase to the central component of the complex. The phases of the three components were randomly sampled from a uniform distribution with a range of 2π radians.

In order to estimate spectral weights, a small perturbation is added to the level of each component on each presentation. The perturbation is sampled randomly and independently from a normal distribution with a mean of zero and standard deviation of 1 dB. The two intervals were separated by a 500 ms gap.

Stimuli were generated with Matlab R2008a running on a PC with Windows 7. The waveforms 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 fixed 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. The overall presentation level was 65 dB SPL.

2.3.3 Procedure

On each trial, subjects were presented with a two alternative forced choice, two-interval forced choice with the stimuli comprising a three-tone complex for both the signal and standard. Listeners had unlimited time to respond during each trial via keyboard. Each block had 70 trials. Stimulus had an onset and offset shaped by a 20 ms linear ramp. A two-down, one-up, adaptive tracking procedure yielded 70.7% correct responses Levitt (1971), with the step size decreasing to 2 dB after the third reversal. A threshold estimate for each 70-trial block was calculated from the average signal level at each reversal point, excluding the first three reversals.

Listeners are asked to discriminate the signal on the basis of three cues: level, pitch and roughness. For the level condition, listeners are instructed to select the stimulus that was perceived to be louder. In the pitch condition, listeners are instructed to select the stimulus that is perceived to be higher in pitch. In roughness condition, listeners are asked to select the stimulus that sounds smoother.

Decision weights are from maximum likelihood estimates for the mean and variance of a conditional-on-a-single-stimulus (COSS) function (Berg, 1989, 1990). Each COSS function represents the response probability as a function of the intensity of one of the three components. This results in a pair of COSS functions for each component, when the signal was in interval one and when the signal was in the second interval. The weights reported are an average of these two independent weight estimates. Since the weights converge to a theoretical limit, rms difference between sets of weights are used and rms value ≤ 0.1 is used. The rms difference,

$$rms = \sqrt{\frac{\sum_i^3 (w_{i1} - w_{i2})^2}{3}}, \tag{2.5}$$

between the sets of weights is used as a reliability index where w_{ij} is the estimated weight for component i when the signal is added to interval j.

2.3.4 General Strategy

A baseline was initially collected to determine listener's preferred cue. Then, manipulations of the stimulus began with isolating the roughness cue. Pitch cues were degraded by using a Δf = 10 Hz, and a frequency shift where the center frequency was drawn uniformly from 1000 Hz $\pm$ 10Hz. Level cues were degraded by equalizing the level of the signal and the standard. Listeners determined which interval contained the smoother stimulus. After listeners completed this condition, the level cue was reintroduced by removing the equalization of the signal to the standard. Listeners again attended to the roughness cue. Then, with both roughness and level cues available, listeners were asked to attend to the level cue.

The third part of this experiment reinstated pitch cues at Δf = 40 Hz. Listeners attended to each of the three cues: level, roughness, and pitch.

2.4 Pilot Experiment

Listeners began with the baseline condition Δf = 40 Hz to establish if there was evidence of a preferred listening cue. Three listeners, (JS, AP and TT) were provided feedback but were not instructed to listen to a specific cue. Listeners completed between 25-30 hours in the baseline condition.

2.4.1 Results

Decision weights and thresholds were estimated for listeners in the baseline condition. Preliminary data indicated that although they were given feedback, "correct" or "incorrect", participants showed weights that appeared to be a combination of two cues. Figure 2.3 the first panel under Baseline, only 1 listener (AP) shows a clear pitch cue preference while the other 2 listeners have observed weights that did not show a clear preferred cue. Future subjects in the following case studies were not required to do a baseline.

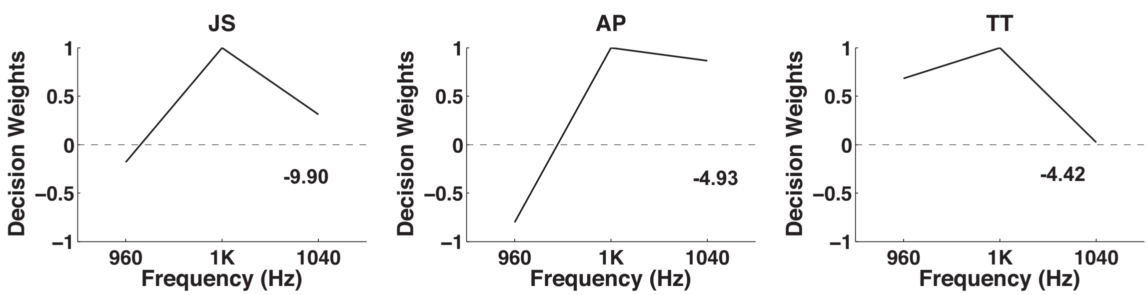


Figure 2.3: Decision weights for the baseline condition when Δf was 40 Hz apart for the baseline.

2.5 Manipulations with a Δf =10 Hz

Preliminary evidence from Southworth and Berg (1995) showed that listeners had the most difficulty with the roughness cue. Combined with the ambiguous baseline obtained from the pilot results, beginning with the Roughness-Only condition was ideal. Using EWAIF-OFL calculations, expected EWAIF-OFL differences at Δf =10 Hz are less than a JND with μ_{signal} = 993.54 Hz and $\mu_{standard}$ = 992.69 Hz (σ_{signal} = 1.33, $\sigma_{standard}$ = 0.99) compared to Δf = 40 Hz with μ_{signal} = 974.32 Hz and $\mu_{standard}$ = 970.81 Hz (σ_{signal} = 5.48, $\sigma_{standard}$ = 4.07). This suggests that reducing the bandwidth also reduces pitch cue as well as the frequency shift. The roughness cue is further isolated by equalizing the overall intensity cue such that level was forced to be equal in the first and second stimuli on any given trial. In listening conditions where the level cue was available, this meant that level was allowed to be randomly influenced by the perturbation that was introduced on that trial.

Five listeners, including the three listeners from the pilot study, were measured at a narrower bandwidth with a frequency separation of Δf = 10 Hz. Listeners completed at minimum 40 blocks for this condition without feedback.

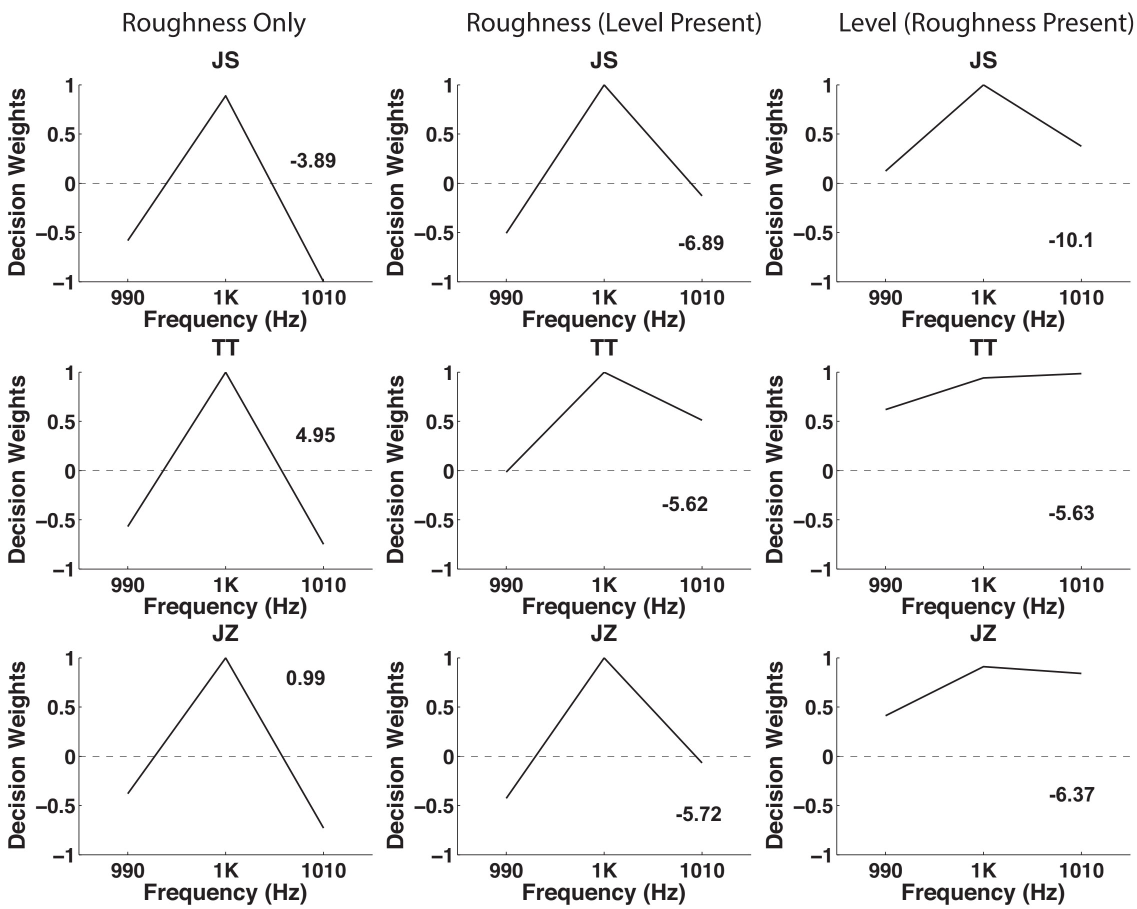


Figure 2.4: Decision weights for listeners with Δf = 10 Hz apart. Each row represents an individual listener in different conditions: roughness only, roughness with level present, and level with roughness present

2.5.1 Results

The first column in Figure 2.4 and in Figure 2.5 shows estimated weights that closely resemble the model when roughness is the only available cue for listeners. The second column of both Figure 2.4 and Figure 2.5 show that when the level cue is present and listeners were asked to attend to the roughness cue, the weights are affected such that changes in sign–from negative to positive, are observed– for the two side components at 990 and 1010 Hz which follows the model predictions for the level cue. However, overall symmetry maintains a roughness cue influence.

The shape of the weights seen here are a marked change from when listeners attended to Roughness Only. This is observed again when listeners attend to the level cue while roughness is present. Listeners attending to level in the presence of a roughness cue show weights that matched the model with equal magnitude between the three components. More importantly, it serves to show that perceptions of what constitutes a signal are quantifiably different.

It is important to note that listeners are observed here to be affected by the presence of other cues. For example, the level cue is seen in Figure 2.4 and Figure 2.5 to affect perception of roughness in a way that suggests either a hybrid model–which would posit that listeners can attend to more than just one cue on a trial-by-trial basis or that a mixed-model in which listeners proportionally use only one cue on each trial.

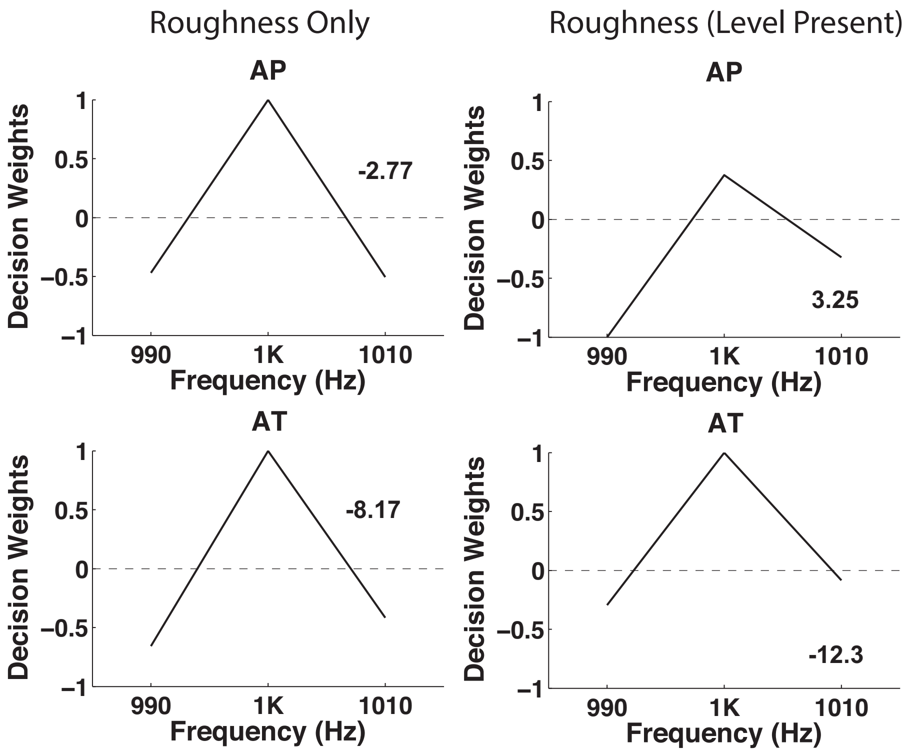


Figure 2.5: Decision weights for listeners with Δf = 10 Hz apart. Each row represents an individual listener in a different condition: roughness only, roughness with level present, and level with roughness present. AP and AT did not complete the Level with roughness present conditions.

2.6 Manipulations at Δf = 40 Hz

Five listeners, JZ, AT, JS, AP, TT who completed all three listening conditions from the previous study at Δf =10 Hz were asked to continue on at Δf = 40 Hz to listen for all three cues, pitch, level, and roughness. Two additional listeners, JF, who completed the roughness only condition prior to this experiment, and CL participated. Listeners began with the roughness cue first, then pitch, then level. The weights associated with the pitch cue, according to the model, indicate that the lowest frequency, 960 Hz should have the greatest magnitude and is negatively signed while the signal component and the higher frequency 1040 Hz component is positively signed. Roughness weights are symmetric about the signal component and has the greatest magnitude and is positively signed. The non-signal components are negatively signed -.5. Level weights are represented by signal and non-signal components being equal in magnitude at 1 with each component positively signed.

2.6.1 Results

Figure 2.6 shows weights from one listener, AP, who had weights for all three cues that followed the model. This indicates that she was able to separately attend to all three cues.which indicated she was able to attend to all three cues.

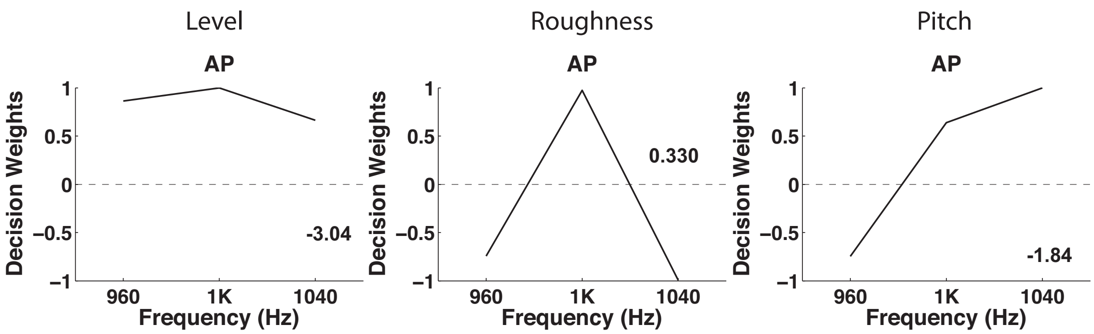


Figure 2.6: Decision weights for subject AP when Δf = 40 Hz apart. Listener shows weights that match the model for all three cues of roughness, level, and pitch.

Figure 2.7 and Figure 2.8 show overall patterns that closely resemble the model for subjects JZ, AT, JS and CL when listening to the level cue and pitch cue. Figure 2.8 shows that all listeners but TT are able to attend to the level cue. One important observation is that when listening to the level cue, estimated weights for listeners JZ, AT and JS show a clear asymmetry that is inherent to the pitch model, i.e., that the largest magnitude is on the lowest frequency component.

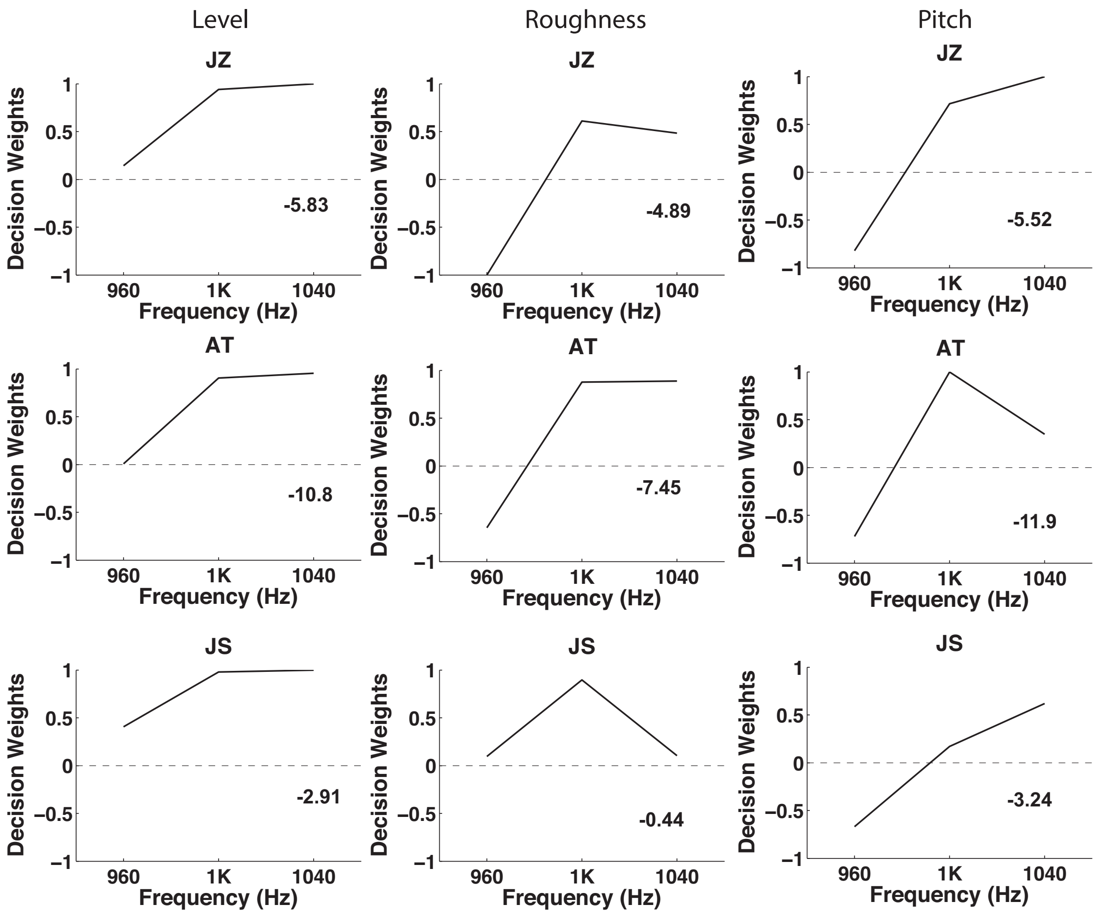


Figure 2.7: Decision weights for baseline condition when $\Delta f = 40$ Hz apart.

When attending to roughness, the second column of Figure 2.7, listeners AT and JZ show weights that resembled pitch rather than the roughness cue. This indicates that these listeners were perhaps unable to distinguish between the two cues. JS showed symmetrical weights with the greatest magnitude at 1 for the signal component. This is indicative of roughness, however each component was positively signed which reflects some influence of the level cue. This is further evidence that some listeners are unable to attend to a single cue.

Another listener, CL showed mirrored asymmetrical weights for Roughness that is typically observed when listeners are using off-frequency listening and attending to pitch cues (Berg et al., 1992).

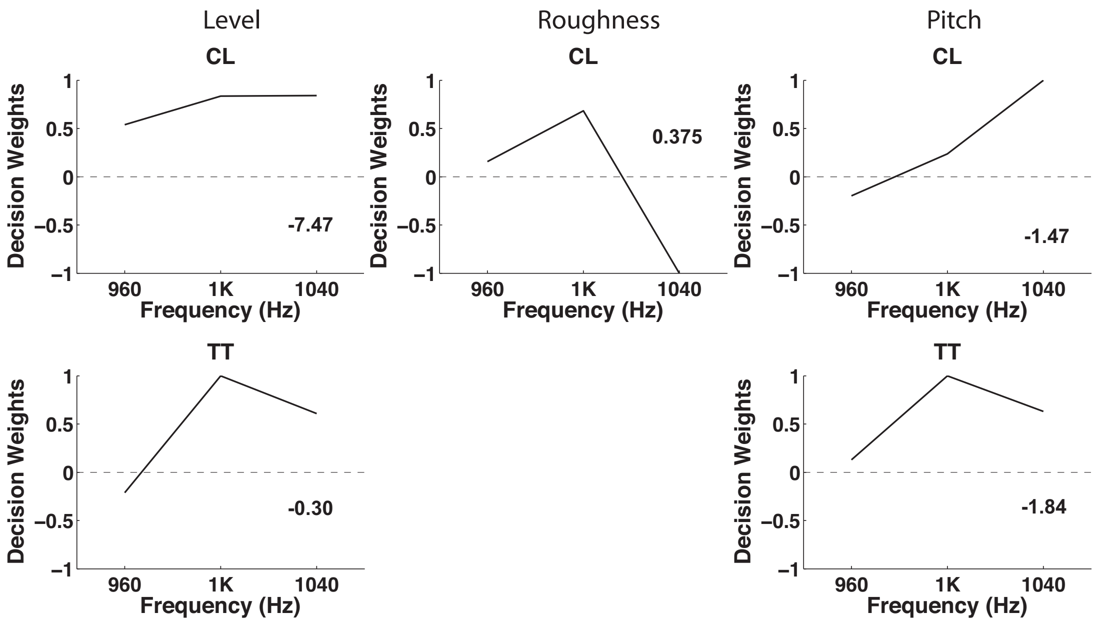


Figure 2.8: Decision weights for baseline condition when $\Delta f = 40$ Hz apart.

Estimated weights observed from listener JF in Figure 2.9 indicated the ability to attend to the Roughness cue at $\Delta f = 40$ Hz. However, when attending to pitch, JF's weights appear to be a combination of two cues with an influence from roughness cue due to the greatest magnitude being on the signal component.

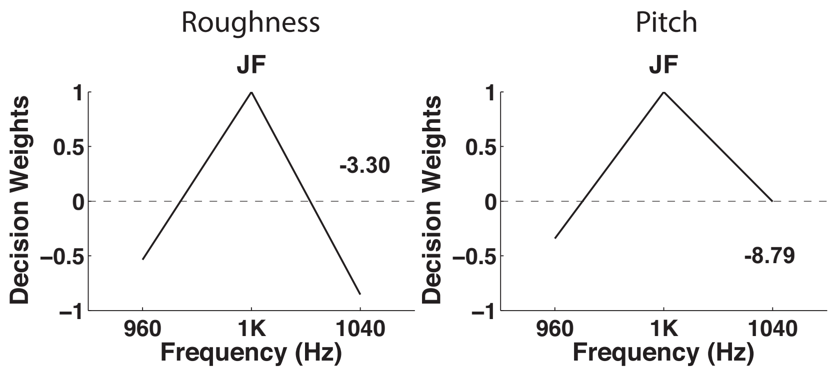


Figure 2.9: Decision weights for listener JF at Δf = 10 Hz for roughness cue (level present) and Δf = 40 Hz apart for pitch and roughness cues. JF can is the only listener able to attend to roughness cues at when Δf = 40 Hz and at Δf = 10 Hz for roughness when level is present.

2.7 Discussion

A psychophysical measure that can directly relay a subject's perceptual experience is a useful tool, especially in the context of understanding a listener's ability to attend to a specific cue. The three-tone narrow band spectra used in these experiments offer the advantage of isolating a specific cue. We can then observe how the decision weights change as a result of a secondary available cue. This allows a quantifiable way to observe how one cue influences the other and to what degree there are a mixture of cues. In some ways, these experiments go toward creating an objectively guided dialogue about individual perception.

In the context of Δf = 10 Hz, when roughness was the only cue available, Figure 2.4 and Figure 2.5 showed that listener weights reflected that of the model predictions and that they were able to listen to the roughness cue in isolation. This means that non-signal weights were negatively signed with the greatest magnitude at the center component. However when asked to attend to the roughness cue in the presence of level, the perception of roughness is influenced in such a way that the sign for the non-signal components change from negative to positive and the symmetry changes for the non-signal components as well. This implies that many listeners can not attend only to roughness when multiple cues are present. The weights clearly show that level is impacting the listeners perception of roughness.

When listeners are asked to attend to level in the presence of roughness, Figure 2.4 shows that listener weights matched the level model, with equal magnitude for each component. Only JS had weights that indicated potential use of a mixture of cues for this condition.

When pitch cues are reintroduced in the $\Delta f = 40$ Hz condition, overall weight patterns show that some listeners are able to give different weights for level and pitch that match the model. However, most listeners in the level condition show positive sign for each component but have an asymmetry that is inherent to weights for pitch. General patterns at $\Delta f = 40$ Hz are more indicative of listeners using a mix of cues. These findings strongly prompt future directions to determine if listeners are using some proportion of two cues separately or a use of hybrid cue on a trial by trial basis.

There is, however, evidence that a listener can produce all three weights. AP was able to produce decision weights that followed closely for each mode. This provides evidence that all three cues exist, and provides motivation to further examine training effects in other listeners.

These findings provide a framework for listeners to be trained to attend to a specific cue. This could have direct impact on cochlear implant (CI) research since it allows for a quantified assessment of the subject's perceptual experience of a particular cue. Pitch perception, for example, has been challenging to simulate correctly for CI users. In CI users with one implanted ear, there is still access to residual hearing. Due to the presence of residual hearing, there is a mismatch introduced between hair cells that are stimulated acoustically and nerve fibers that are stimulated electrically. Reiss et al. (2014) found that hybrid CI listeners had difficulty associating a pitch percept evoked by single CI due to an electro-acoustic mismatch that developed overtime. Initially, CI users were given an acoustic reference tone in the non-implanted ear, but over the course of several months, their pitch perception shifted to match that of the non-implanted ear. The methodology used in this chapter could be applied to CI research such that perceptual weight changes for CI users can be monitored and visually described.

Bibliography

Ahonniska, J., Cantell, M., Tolvanen, A., and Lyytinen, H. (1993). Speech perception and brain laterality: the effect of ear advantage on auditory event-related potentials. *Brain and Language*, 45:127–146.

Ahumada, A. J. and Lovell, J. (1971). Stimulus features in signal detection. *J. Acoust. Soc. Am.*, 49:1751–1756.

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

Berg, B. G. (1990). Observer efficiency and weights in a multiple observation task. *J. Acoust. Soc. Am.*, 88(1):149–158.

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.

Berman, S. M., Mandelkern, M. A., Phan, H., and Zaideld, E. (2003). Complementary hemispheric specialization for word and accent detection. *NeuroImage*, 19(2003):319–331.

Broadbent, D. (1954). The role of auditory localization in attention and memory span. *Journal of Experimental Psychology*, 47(3):191–196.

Bryden, M., Munhall, K., and Allard, F. (1983). Attentional biases and the right-ear effect in dichotic listening. *Brain and Language*, 18:236–248.

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

Eckstein, V. (2011). *A Novel Model for Pitch Perception and Functional Localization of Attentionally Modulated Pitch and Loudness Perception*. PhD thesis, University of California, Irvine.

Feth, L. L. (1974). Frequency discrimination of complex periodic tones*. *Perception and Psychophysics*, 15(2):375–378.

Feth, L. L. and O'Malley, H. (1977). Two-tone auditory spectral resolution. *J. Acoust. Soc. Am.*, 62(4):940–947.

Gerald Kidd, J., Mason, C. R., Brantley, M. A., and Owen, G. A. (1989). Roving-level tone-in-noise detection. *J. Acoust. Soc. Am.*, 86(4):1310–1317.

Green, D. ((1988)). *Profile Analysis: Auditory Intensity Discrimination.* New York: Oxford Univ. Press.

Green, D. M., Berg, B. G., Dai, H., Eddins, D. A., Onsan, Z., and Nguyen, Q. (1992). Spectral shape discrimination of narrow-band sounds. *J. Acoust. Soc. Am.*, 92(5):2586–2597.

Hugdahl, K. (2011). Fifty years of dichotic listening research – still going and going and. . . *Brain and Cognition*, 76:211–213.

Kass, R. E. and Raftery, A. E. (1995). Bayes factors. *Journal of the American Statistical Association*, 90:773–795.

Kimura, D. (1961). Cerebral dominance and the perception of verbal stimuli. *Canadian Journal of Psychology*, 15:166–171.

Leshowitz, B. and Wightman, F. L. (1971). On-frequency masking with continuous sinusoids. *J. Acoust. Soc. Am.*, 49(4):1180–1190.

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

Love, J., Selker, R., Marsman, M., Jamil, T., Dropmann, D., Verhagen, A. J., Ly, A., Gronau, Q. F., Smira, M., Epskamp, S., Matzke, D., Wild, A., Rouder, J. N., Morey, R. D., and Wagenmakers, E.-J. (2015). Jasp (version 0.7)[computer software].

Lutfi, R. A. and Jesteadt, W. (2006). Molecular analysis of the effect of relative tone level on multitone pattern discrimination. *J. Acoust. Soc. Am.*, 120(6):3853–3860.

Lutfi, R. A., Liu, C.-J., and Stoelinga, C. (2008). Level dominance in sound source identification. *J. Acoust. Soc. Am.*, 124(6):3784–3792.

Morey, R. D., Hoekstra, R., Rouder, J. N., Lee, M. D., and Wagenmakers, E.-J. (2015). The fallacy of placing confidence in confidence intervals. *Psychonomic Bulletin and Review*, pages 1–21. in-press.

Oberfeld, D. and Plank, T. (2011). The temporal weighting of loudness: Effects of the level profile. *Atten. Percep. Psychophys.*, 73:189–208.

Ofek, E. and Pratt, H. (2004). Ear advantage and attention: an erp study of auditory cued attention. *Hearing Research*, 189:107–118.

Reiss, L. J., Turner, C., Karsten, S., and Gantz, B. (2014). Plasticity in human pitch perception induced by tonotopically mismatched electro-acoustic stimulation. *Neuroscience*, 256(2014):43–52.

Richards, V. M. (1992). The detectability of a tone added to narrow bands of equal-energy noise. *J. Acoust. Soc. Am.*, 91(6):2424–2435.

Richards, V. M., Shen, Y., and Chubb, C. (2013). Level dominance for the detection of changes in level distribution in sound streams. *J. Acoust. Soc. Am.*, 134(2):EL237–EL243.

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

Studdert-Kennedy, M. and Shankweiler, D. (1970). Hemispheric specialization for speech perception. *J. Acoust. Soc. Am.*, 48(2):579–594.

Turner, M. and Berg, B. (2007). Temporal limits of level dominance in a sample discrimination task. *J. Acoust. Soc. Am.*, 121(4):1848–1851.

Wagenmakers, E.-J. (2007). A practical solution to the pervasive problems of p values. *Psychonomic Bulletin and Review*, 14:779–804.

Zatorre, R. J. and Belin, P. (2001). Spectral and temporal processing in human auditory cortex. *Cerebral Cortex*, 11(10):946–953.

Zwicker, E., Flottorp, G., and Stevens, S. (1957). Critical band width in loudness summation. *J. Acoust. Soc. Am.*, 29(5):548–557.

Appendices

A Ideal Weights

Ideal weights are derived from

$$\hat{a}_i = \frac{1}{\sigma_i^2 \sum(\frac{1}{\sigma_i^2})}, \tag{A.1}$$

where $\hat{a}_i$ is the ideal weight for the ith observation.

B Tonal Interactions

For any trial t and each of the stimulus tones $k = 1, 2, 3$, let $\alpha_{1,k}(t)$ be the amplitude of tone k in the first of the two stimuli presented on trial t, and let $\alpha_{2,k}(t)$ be the amplitude of tone k in the second stimulus. Then let $D_k(t) = \alpha_{2,k}(t) - \alpha_{1,k}(t)$, and let $\Delta_k(t)$ be the deviation of $D_k(t)$ from its mean value across all trials.

In each of the different conditions in which the listener is tested, the listener is asked to judge which of stimulus 1 or stimulus 2 on a given trial is higher in some target property T. To describe the results from a given condition, we use a general linear model in which we assume that the participant judges stimulus 2 to be higher in T than stimulus 1 if

$$M(t) + X_t > 0 \tag{B.2}$$

where the X_t's across trials t are jointly independent, standard norm random variables and

$$\begin{aligned} M(t) = {} & w_1\Delta_1(t) + w_2\Delta_2(t) + w_3\Delta_3(t) + w_{12}\Delta_1(t)\Delta_2(t) + w_{13}\Delta_1(t)\Delta_3(t) \\ & + w_{23}\Delta_2(t)\Delta_3(t) + w_{123}\Delta_1(t)\Delta_2(t)\Delta_3(t) + Bias \end{aligned} \tag{B.3}$$

for model parameters w_1, w_2, w_3, w_{12}, w_{13}, w_{23}, w_{123} and $bias$.

A Bayesian method was used to estimate the joint posterior density of the model parameters from the data. Figure B.1 shows that the first 3 tones are significant with patterns that match the aforementioned three models, which is observed in listener data. However, Figure B.2 shows that none of the calculated weights for the interaction terms differed significantly from 0 for simulated data. This is further corroborated by analysis of data collected from listeners which also showed interaction terms that did not differ from 0.

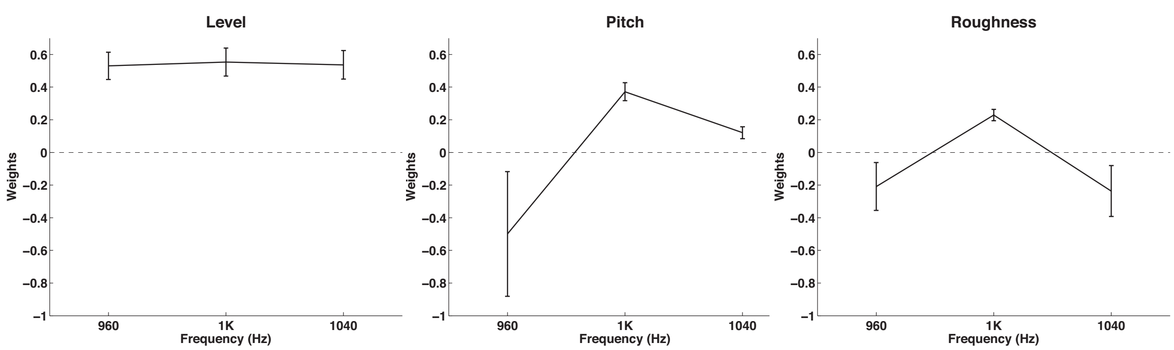


Figure B.1: Data from model simulations for level, envelope, and pitch and thresholds show significance for the first order statistics with tones 1, 2, and 3.

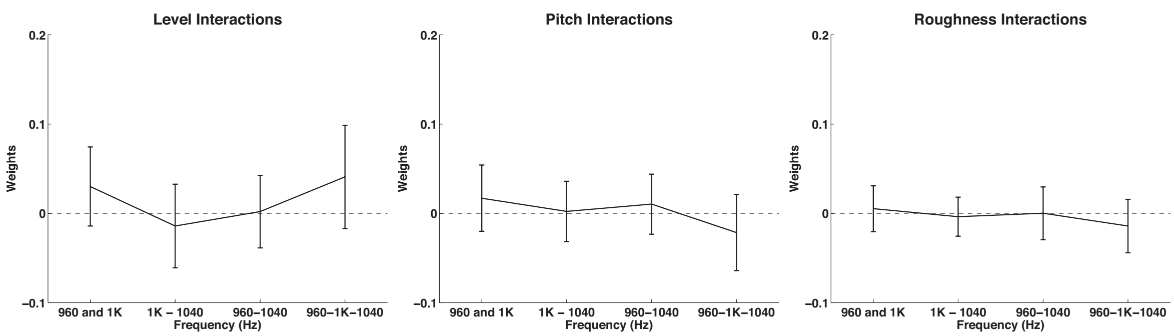


Figure B.2: Data from computer simulations for the interaction terms for level, envelope and pitch indicate no significant difference from 0.