Monaural perception under dichotic conditions

by

Daniel E. Shub
B.S.E., Bioengineering
University of Pennsylvania, 1997

SUBMITTED TO THE HARVARD-MIT DIVISION OF HEALTH SCIENCES AND TECHNOLOGY IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY IN HEALTH SCIENCES AND TECHNOLOGY
AT THE
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
[February 2007]
OCTOBER 2006

Signature of Author: ______________________
Harvard-MIT Division of Health Sciences and Technology
October 19, 2006

Certified by: ______________________
H. Steven Colburn, Ph.D.
Professor of Biomedical Engineering
Boston University

Accepted by: ______________________
Martha L. Gray, Ph.D.
Edward Hood Taplin Professor of Medical and Electrical Engineering
Co-Director, Harvard-MIT Division of Health Sciences and Technology

Monaural perception under dichotic conditions

by

Daniel E. Shub

SUBMITTED TO THE HARVARD-MIT DIVISION OF HEALTH SCIENCES AND TECHNOLOGY IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY IN HEALTH SCIENCES AND TECHNOLOGY

Abstract

Most people have two ears, but we can hear with only one ear. The ability to use two ears can substantially improve performance in many circumstances. There are times, however, when the addition of a second ear results in poorer performance (i.e., contra-aural interference). Contra-aural interference is of interest because it is not explained by current auditory models, it has theoretical ramifications, and its understanding could lead to improvements in the quality of life of the hearing-impaired. More generally, the techniques and results can be applied to fields in which information is combined across an array of sensors (e.g., vision with two eyes and radar arrays). This thesis includes both psychophysical measurements and black-box modeling of level discrimination. Level discrimination was chosen to study contra-aural interference since it has traditionally been considered a monaural task (dependent on only a single ear) even though the loudness of a sound depends on both ears (i.e., binaural).

This thesis demonstrates that the ability to discriminate small changes in the level of a low-frequency target stimulus presented at one ear can be adversely affected by a distractor stimulus presented simultaneously and contra-aurally to the target. The thesis focuses on conditions in which the target and distractor perceptually fuse; the dominant perception of the stimulus is a compact auditory image with a salient loudness and position and a secondary image referred to as the "time-image". Contra-aural interference was greatest when the introduction of the distractor decreased the reliability of both the perceived loudness and position of the dominant-image.

Although the tasks used in this thesis are artificial, their simplicity allows for detailed computational modeling. The results are consistent with a model based on non-optimal integration of the information carried by the dominant-image and the time-image. The modeling separates the effects of internal coding noise and decision noise (criterion jitter). The techniques used to separate the internal coding noise from the criterion jitter can be applied to a broad range of psychology experiments.

Acknowledgements

In order for me to have completed this thesis, I required a tremendous amount of support. Some of this support came in terms of academic contributions while other portions were social. The academic environment of the Speech and Hearing Bioscience and Technology program and the Hearing Research Center was supportive, nurturing, and above all rigorous. I am grateful to all the faculty, staff, and students who have helped me become a scientist.

I would especially like to thank my thesis committee for their help. Bertrand Delgutte not only served as the chair of committee he was also my academic advisor and chair of my concentration area. He always kept me on track even when I wanted to veer. Andrew Oxenham helped in the conceptualization of the psychophysical tasks and the interpretations of their results. He stuck by me even as much of what he was most interested in got cut from the thesis document. Nat Durlach and his ability to question the most basic aspects of my research shaped the later portions of the thesis. Steve Colburn was not only my thesis supervisor, but also a friend. This thesis represents an infinitesimal portion of what I learned from Steve. I can only hope that Steve enjoyed working with me a fraction of how much I enjoyed working with him.

Apart from my committee, I would also like to thank the students, post docs, and faculty members who comprise the binaural gang at Boston University. In my years as a graduate student, there were too many members of the gang to name, and all have shaped me. My time interacting with Barbara Shinn-Cunningham at our weekly lab meetings has shown me that research can be intellectually stimulating and a great deal of fun.

With the quality of my mentors it would have been hard for this thesis to fail. Without my friends and family, however, the thesis would not have succeeded. I would like to thank my friends who have kept me climbing and reminded me to get out occasionally. Mot importantly I want to thank Kristine Juncker for making home such a nice place to go.

Whether your contributions were academic or social, this thesis could not have been completed without you. Thank you.

Table of Contents

CHAPTER I. INTRODUCTION

I. Thesis Overview

Most people have two ears, but we can hear with only one ear. The ability to use two ears can substantially improve performance in many circumstances, although for many auditory tasks, the addition of a second ear results in only a marginal improvement. Most important for this thesis, however, is that there are times when the addition of a second ear results in poorer performance. These cases are said to exhibit "contra-aural interference". This thesis includes both empirical measurements and theoretical analyses which seek to identify when and how a second ear results in a decrease in performance. This chapter begins with an introduction of the aspects of auditory psychophysics and modeling which motivated this thesis.

II. Background

From a broad perspective, this thesis examines, through psychophysical measurements and black-box modeling, the interweaving of binaural and monaural processing. This interweaving is studied by considering the disadvantages of binaural listening over monaural listening in a level discrimination task. Binaural hearing research has traditionally focused on advantages in localization, discrimination, detection, and intelligibility (Durlach and Colburn 1978; Blauert 1997) that arise from listening with two ears. Based on psychophysical research and auditory modeling one might erroneously believe that binaural and monaural processing are independent; psychophysical research and auditory modeling have traditionally been classified as either binaural or monaural (dependent on only a single ear). Further, monaural models do not consider the "other" ear and binaural models typically include separate monaural and binaural pathways. This monaural/binaural division suggests that listeners should be quite adept at attending to a single ear when it is desirable. The psychophysical measurements in this thesis evaluate some limitations on the ability to extract information presented to a single ear under conditions in which it is subjectively difficult to attend to a single ear.

Systematic studies of contra-aural interference are needed not only to increase our understanding of the fundamentals of auditory processing, but also to help the hearing impaired. The hearing impaired often use bilateral aids to improve their ability to localize sounds and understand speech in noise. Speech intelligibility with two hearing aids, however, can be worse than intelligibility with one aid (Walden and Walden 2005). Localization with no aids can be better than with two aids (Van den Bogaert et al. 2006). These decreases in performance may be a result of mismatches in the aided ears due to both the pathology of the hearing loss and the independent processing in the two hearing aids. Typically, each hearing aid has its own automatic gain control, level setting, and phase delays that can lead to the corruption of the binaural information. The possibility for corruption of the binaural information is even greater in bilateral cochlear implants where the pulse rate and pulse timing also can vary across the ears.

How this corruption affects the ability of a listener to access monaural information is not understood.

There have been no systematic studies of contra-aural interference; instead contra-aural interference is generally treated as an anomalous finding. Treating contra-aural interference as an anomaly is unjustified given the range of experimental conditions for which contra-aural interference has been reported, including "central masking" (Zwislocki 1972; Mills et al. 1996), binaural masking (Taylor and Clarke 1971; Taylor et al. 1971; Yost et al. 1972; Koehnke and Besing 1992), and speech-on-speech masking (Brungart and Simpson 2002; Kidd et al. 2003) studies. Contra-aural interference has also been measured in studies of the precedence effect (Zurek 1979) and pitch perception (Bernstein and Oxenham 2003). Heller and Trahiotis (1995) reported that the ability of subjects to discriminate monotic noise targets can be decreased by presenting a distractor in the ear contralateral to the target. Most important for this thesis are the reports of contra-aural interference in level discrimination tasks (Rowland and Tobias 1967; Yost 1972; Bernstein 2004; Stellmack et al. 2004).

A series of psychophysical experiments on contra-aural interference in a level discrimination task were conducted. All of the psychophysical work involved normal-hearing subjects making judgments about the level of a target in the presence of a distractor. The target (a 600-Hz tone) and the distractor (also a 600-Hz tone) were presented simultaneously, but in opposite ears. The dominant perception of the combined target and distractor stimulus was a single compact auditory image with a salient loudness and position. Therefore, changes to the target level affected both the perceived loudness and position.

Traditional binaural models have strong predictive power for classical psychoacoustic experiments involving changes in the perceive loudness and position (Colburn and Durlach 1978; Colburn 1996; Blauert 1997; Stern and Trahiotis 1997), but they fail to predict any contra-aural interference. In traditional models the decision devices typically have access to both purely monaural and binaural representations of the signals under all stimulus conditions (monotic, diotic and dichotic). Modeling of contra-aural interference has been limited to *ad hoc* models included in some of the studies in which the interference was reported (Zurek 1979; Heller and Trahiotis 1995). These models, however, are not able to predict the same breadth of psychophysical results as traditional models.

This thesis is a systematic study of contra-aural interference in a level discrimination task. The aim of this thesis is to demonstrate that contra-aural interference is not an anomalous result, but rather can be interpreted by considering the perceptual attributes of the stimuli. Interpreting the results based on perceptual attributes is substantially different than the traditional interpretation based on independent monaural and binaural processing. To achieve this aim, the stimuli in the psychophysical experiments were manipulated to give variable amounts of contra-aural interference. In the modeling, predictions based on non-optimal use of the obvious perceptual dimensions were calculated.

III. Overview of the Thesis Document

The thesis work can be divided into psychophysical and modeling work, which are nevertheless highly intertwined. The psychophysical tasks were designed to test the model predictions and the model formulation stemmed from the psychophysical tasks. For this reason both the psychophysical results and the modeling predictions are presented in close association throughout the thesis.

This thesis is divided into five chapters. Chapters II, III, and IV present the findings of both the psychophysical and the modeling work. These chapters are structured such that they can be read individually. Chapter II presents the findings of a level-discrimination experiment which used a multi-interval adaptive paradigm to estimate the just-noticeable difference in level and compares the results with the predictions of a model based on the perceived loudness and position. Chapter III presents the findings of a level-discrimination experiment which used a single-interval, constant-increment paradigm; the predictions of the model based on two decision variables, perceived loudness and position, are evaluated. Chapter IV focuses on predicting the psychophysical results of Chapter III with a model based on a non-optimal combination of three decision variables, namely the perceived loudness, position, and time-image. Chapter V presents an integrated picture of the psychophysical results and the modeling as well as some directions for future work.

References

Bernstein, J. G., and Oxenham, A. J. (**2003**). "Pitch discrimination of diotic and dichotic tone complexes: Harmonic resolvability or harmonic number?," J. Acoust. Soc. Am. **113,** 3323-34.

Bernstein, L. R. (**2004**). "Sensitivity to interaural intensitive disparities: Listeners' use of potential cues," J. Acoust. Soc. Am. **115,** 3156-60.

Blauert, J. (**1997**). "Spatial hearing: The psychophysics of human sound localization," (MIT Press, Cambridge).

Brungart, D. S., and Simpson, B. D. (**2002**). "Within-ear and across-ear interference in a cocktail-party listening task," J. Acoust. Soc. Am. **112,** 2985-95.

Colburn, H. S., and Durlach, N. I. (**1978**). "Models of binaural interaction" in *Handbook of perception,* Vol. IV edited by E. C. Carterette and M. P. Friedman.(Academic Press, New York)

Colburn, H. S. (**1996**). "Computational models of binaural processing" in *Auditory computation,* Vol. 6 edited by H. Hawkins and T. McMullen.(Springer Verlag, New York)

Durlach, N. I., and Colburn, H. S. (**1978**). "Binaural phenomena" in *Handbook of perception,* Vol. IV edited by E. C. Carterette and M. P. Friedman.(Academic Press, New York)

Heller, L. M., and Trahiotis, C. (**1995**). "The discrimination of samples of noise in monotic, diotic, and dichotic conditions," J. Acoust. Soc. Am. **97,** 3775-81.

Kidd, G. J., Mason, C. R., Arbogast, T. L., Brungart, D. S., and Simpson, B. D. (**2003**). "Informational masking caused by contralateral stimulation," J. Acoust. Soc. Am. **113,** 1594-603.

Koehnke, J., and Besing, J. M. (**1992**). "Effects of roving level variation on monaural detection with a contralateral cue," J. Acoust. Soc. Am. **92,** 2625-9.

Mills, J. H., Dubno, J. R., and He, N.-J. (**1996**). "Masking by ipsilateral and contralateral maskers," J. Acoust. Soc. Am. **100,** 3336-44.

Rowland, R. C. J., and Tobias, J. V. (**1967**). "Interaural intensity difference limen," Journal of Speech and Hearing Research **10,** 745-56.

Stellmack, M. A., Viemeister, N. F., and Byrne, A. J. (**2004**). "Monaural and interaural intensity discrimination: Level effects and the "binaural advantage"," J. Acoust. Soc. Am. **116,** 1149-59.

Stern, R. M., and Trahiotis, C. (**1997**). "Models of binaural perception" in *Binaural and spatial hearing in real and virtual environments,* Vol. edited by R. H. Gilkey and T. R. Anderson.(Lawrence Erlbaum Associates, Mahwah)

Taylor, M. M., and Clarke, D. P. J. (**1971**). "Monaural detection with contralateral cue (MDCC). II. Interaural delay of cue and signal," J. Acoust. Soc. Am. **49,** 1243-53.

Taylor, M. M., Clarke, D. P. J., and Smith, S. M. (**1971**). "Monaural detection with contralateral cue (MDCC). III. Sinusoidal signals at a constant performance level," J. Acoust. Soc. Am. **49,** 1795-804.

Van Den Bogaert, T., Klasen, T. J., Moonen, M., Van Deun, L., and Wouters, J. (**2006**). "Horizontal localization with bilateral hearing aids: Without is better than with," J Acoust Soc Am **119,** 515-26.

Walden, T. C., and Walden, B. E. (**2005**). "Unilateral versus bilateral amplification for adults with impaired hearing," J Am Acad Audiol **16,** 574-84.

Yost, W. A. (**1972**). "Tone-on-tone masking for three binaural listening conditions," J. Acoust. Soc. Am. **52,** 1234-7.

Yost, W. A., Penner, M. J., and Feth, L. L. (**1972**). "Signal detection as a function of contralateral sinusoid-to-noise ratio," J. Acoust. Soc. Am. **51,** 1966-70.

Zurek, P. M. (**1979**). "Measurements of binaural echo suppression," J. Acoust. Soc. Am. **66,** 1750-7.

Zwislocki, J. J. (**1972**). "A theory of central auditory masking and its partial validation," J. Acoust. Soc. Am. **52,** 644-59.

CHAPTER II. MONAURAL LEVEL DISCRIMINATION UNDER DICHOTIC CONDITIONS

Abstract

A psychophysical experiment using a monaural level discrimination task was carried out in order to assess the influence of the contralateral ear (the ear that was not involved in the monaural task) on performance. Subjects were asked to discriminate the level of a 600-Hz target tone at one ear in the presence of a 600-Hz distractor tone presented simultaneously and contra-aurally to the target. Subjects were unable to ignore the distractor, even though the distractor was presented contra-aurally to the target. In general, subjects perceived a single, compact, auditory image whose perceived loudness and lateral position were influenced by both the target and the distractor. The results of this experiment are compared to predictions from a model that utilizes two binaural dimensions that loosely correspond to the perceived loudness and lateral position. The information conveyed by the loudness and position dimensions of the model is severely degraded by roving the level and phase of the distractor tone. In the data, when the level and phase of the distractor were both roved, the measured level discrimination thresholds increased by an order of magnitude (average thresholds increased from 0.7 dB to over 7.0 dB). When only the distractor phase was roved, the information in the lateral position dimension was disrupted, but the information in the loudness dimension remained intact. Under this condition, the average threshold was 1.6 dB. The presence of a distractor with a fixed level

and phase had a small, but significant, detrimental effect; the threshold was 1.1 dB with a fixed distractor.

I. Introduction

Auditory research has traditionally been classified as either monaural (dependent on only a single ear) or binaural (dependent on interaural differences). Overall-level discrimination of monotic, diotic, and dichotic stimuli has traditionally been considered a monaural task. Classic models predict that performance on a monotic level discrimination task cannot be degraded by the addition of a stimulus at the other ear: Monaural models do not consider the "other" ear and, therefore, predictions are independent of the stimuli at that ear. Binaural models typically include both monaural and binaural pathways (cf., Colburn and Durlach 1978 for a review of binaural modeling) so that monotic performance provides a lower bound on performance. This monaural/binaural division suggests that listeners should be quite adept at attending to a single ear when it is desirable. In this work, we evaluate the ability (or inability) of normal-hearing listeners to attend to a target sound at one ear in the presence of a distractor sound at the other ear. This work tests the hypothesis that monotic level discrimination uses information carried by binaural channels.

In disagreement with the monaural/binaural division and traditional models, studies of informational masking (Brungart and Simpson 2002; Kidd et al. 2003), the precedence effect (Zurek 1979), pitch perception (Bernstein and

Oxenham 2003), and noise-token discrimination (Heller and Trahiotis 1995) have provided evidence that, under complex listening conditions, normal-hearing subjects are, in fact, unable to attend to a single ear. In studies of informational masking (e.g., Brungart and Simpson 2002; Kidd et al. 2003), monaural speech intelligibility was reduced by adding a masker to the other ear. Zurek (1979) demonstrated that detection of an acoustic reflection can be better under diotic conditions than under dichotic conditions. In Bernstein and Oxenham (2003), the ability to discriminate changes in the fundamental frequency of a harmonic tone complex was reduced under dichotic conditions. Heller and Trahiotis (1995) reported that subjects could discriminate different noise tokens under monotic conditions, but that the subjects were unable to discriminate the tokens under some dichotic conditions. These studies all reported substantial contra-aural interference (i.e., a decrease in performance when two ears are used relative to performance when one ear is used)[1]. In the studies with complex listening conditions, the contra-aural interference was interpreted using the concepts of the central spectrum (Bilsen 1977; Bilsen and Raatgever 2000) and non-optimal across-frequency processing.

Substantial contra-aural interference has also been reported in both masked (Taylor and Clarke 1971; Taylor et al. 1971a; Taylor et al. 1971b; Yost et al. 1972; Koehnke and Besing 1992) and absolute (Zwislocki 1972; Mills et al. 1996) detection experiments. Due to the nature of the stimuli used in these experiments, the central spectrum and non-optimal across-frequency processing are insufficient to explain the contra-aural interference. The reliability of some of these results, however, has been questioned. Durlach and Colburn (1978), for example, stated that the results of Zwislocki were highly dependent on the paradigm and training and were not easily replicated. Further, in the studies by Taylor and colleagues, the psychometric functions (percent correct versus signal to noise ratio) were non-monotonic, but performance was measured adaptively. Finally, in the masked detection studies, the relevant perceptual attributes were unstable. For example, while for one signal-to-noise ratio (SNR) the presence of the signal was detected as an increase in the tonality of the sound, for another SNR, differing by just a few decibels, the target was detected as a decrease in the tonality. The effects of these drastic changes in the relevant perceptual attributes on contra-aural interference have not been explored.

Small amounts of contra-aural interference have also been demonstrated in level discrimination experiments. Rowland and Tobias (1967) and Yost (1972) measured small, but significant, amounts of contra-aural interference in a level discrimination task. In a similar task, however, Stellmack et al. (2004) did not find significant contra-aural interference. Bernstein (2004) used a more complex distractor and reported slightly larger amounts of contra-aural interference.

Although contra-aural interference has been reported with a wide range of stimuli and paradigms, modeling of contra-aural interference is limited. Monaural models are incapable of predicting contra-aural interference since they do not include the required contra-aural inputs. In some cases, binaural models have been able to predict certain aspects of contra-aural interference (Zurek 1979; Heller and Trahiotis 1995). These models of contra-aural interference, however, have not been applied to a wide range of experiments. For example, the model used by Zurek (1979) bases its decisions on the depths of notches in the central spectrum. To our knowledge, this model has not been applied to the results of any other experiments.

Traditional binaural models based on equalization and cancellation (Durlach 1963), cross-correlation (Colburn 1973) and position variables (Hafter 1971; Yost 1972; Stern and Colburn 1978) require monaural processors to accurately predict performance on monotic and diotic tasks. In these models, the decision devices usually (1) use an optimal combination of the information from monaural and binaural processors, (2) optimally switch between using the information from the monaural processor and from the binaural processor, or (3) have unspecified mechanisms for using the monaural and binaural information. The monaural processor places a lower limit on performance in these models and prevents the models from predicting contra-aural interference. In general, it appears that the monaural/binaural division of research and the traditional framework of auditory models persist due to the paucity of the psychophysical data convincingly demonstrating contra-aural interference in simple tasks which can be thoroughly analyzed.

In the work described in this paper, a psychophysical experiment was conducted in order to assess the ability of listeners to attend to a single ear with tonal stimuli. This study extends the work of Rowland and Tobias (1967) and Bernstein (2004) and measures contra-aural interference in a level discrimination task with pure-tone stimuli. In the psychophysical experiment, subjects were asked to discriminate the level of a low-frequency target tone at one ear, in the presence of an identical frequency distractor tone, presented simultaneously and contra-aurally to the target. The psychophysical data are compared to predictions of a two-dimensional detection theoretic model.

Throughout this work, the level and phase of the target and the distractor are discussed individually. The experimental task is nominally monaural level discrimination, and traditional models of level discrimination process the monaural target and contra-aural distractor separately. Treating the target and distractor separately, however, does not agree with the perception of the stimuli; the target and distractor perceptually fuse into a single dominant image with a salient loudness and lateral position. Unlike monaural models, which predict no influence of the distractor on performance, a substantial influence of the distractor is predicted by a binaural model that relates the binaural level, the interaural difference in level (ILD), and the interaural difference in time delay (ITD)[2] to the loudness and lateral position.

Across the experimental conditions, the reliability of the loudness and the lateral position are systematically varied. The loudness and lateral position are both reliable for discriminating the level of the target when the distractor level and phase are not roved (i.e., are fixed). When the phase of the distractor is roved, the reliability of the lateral position for discriminating the target is greatly reduced. Roving the level of the distractor reduces the reliability of both the loudness and the lateral position. In order to reduce the reliability of the loudness and lateral position, the ranges over which the phase and level of the distractor are roved are large (30 dB range in level and π in phase). Discrimination performance with a fixed distractor, however, has been shown to be nearly independent of the level and the phase of the distractor over the entire range of distractor levels and phases used in this work (Hershkowitz and Durlach 1969; Domnitz 1973; Domnitz and Colburn 1977). Thus, any changes in discrimination performance are due to the roving of the distractor, rather than the specific level and phase of the distractor.

In the modeling portion of this work, a binaural model based on the binaural level and a simple position variable (a weighted sum of the ILD and ITD) is evaluated. Models based on this position variable have predicted the results of many binaural experiments (Hafter 1971; Yost 1972). Although this position variable has strong predictive power, it is insufficient to predict all binaural effects. Particularly, the position variable predicts a complete time-intensity trade, while the incomplete nature of the trade has been widely demonstrated (Hershkowitz and Durlach 1969; Hafter and Carrier 1972; Ruotolo et al. 1979). Further, Bernstein (2004) reported that subjects can outperform an ideal observer of this simple position variable when discriminating the ILD of tones with a random ITD. The shortcomings of this simple position variable may be overcome by a variety of modifications (Stern and Trahiotis 1997). These modifications add complexity and may not be required to predict monaural level discrimination under dichotic conditions. The modeling portion of this work calculates the information, for monaural level discrimination, carried by the binaural level and the unmodified position variable.

II. Experimental Methods

Subjects were asked to discriminate the level of a 600-Hz target tone at one ear, in the presence of a 600-Hz distractor tone presented simultaneously and contra-aurally to the target. Five different conditions were explored. Across the conditions, the target was the same and the properties of the distractor varied. Table 1 lists the properties of the distractor used in each condition. In the psychophysical experiment, a 4-interval, 2-cue, 2-alternative, forced-choice, (4I-2AFC) adaptive paradigm was used. All roving was done in an interval-by-interval manner.

A. Subjects

Four subjects (S1, S2, S3, and S4) completed the tasks. Subject S1 is an author. Subjects, with the exception of the author, received an hourly wage for their participation. All subjects had pure tone thresholds below 20 dB HL at frequencies of 250, 500, 1000, 2000, 4000, and 8000 Hz in both ears. Subject ages were between 19 and 31 years old. Subjects S1 and S2 had prior listening experience in similar tasks, while subjects S3 and S4 had no prior listening experience.

B. Apparatus and Materials

During the experiment, subjects sat in a sound treated in front of a computer monitor. They responded through a graphical interface via a computer mouse. On each trial, "lights" displayed on an LCD monitor denoted the current interval number. The experiment was self-paced and listening sessions lasted no more than 2 hours with frequent rest breaks. A PC and Tucker-Davis-Technology System II hardware (AP2, PD1, PA4, and HB6) generated the experimental stimuli at a sampling rate of 50 kHz. Stimuli were presented over Sennheiser HD 265 headphones.

C. Stimuli

The stimuli in the experiment consisted of both a target and a distractor. The target and distractor were both 600-Hz tones with 300-ms durations and 25-ms rise/fall times. The left ear received the target, while the right ear received the distractor. The target and distractor were gated on and off simultaneously. The target was presented at either the reference level of 50 dB SPL or the reference level plus an increment of ΔL (in decibels); the target phase was always zero.

As indicated in Table 1, five different distractor conditions were explored: (1) no-distractor, (2) fixed, (3) roving-phase, (4) roving-level, and (5) double-rove (i.e., roving-phase and roving-level). In conditions in which the level was fixed

(fixed and roving-phase), the level of the distractor was fixed at 50 dB SPL (the reference level of the target). In the conditions in which the level was roved (roving-level and double-rove), the level of the distractor was chosen randomly, on an interval-by-interval basis, from a uniform distribution between 50 and 80 dB SPL. In the conditions in which the phase was roved (roving-phase and double-rove), the phase of the distractor roved uniformly between $\pm\pi/2$ on an interval-by-interval basis. In the fixed-phase conditions (fixed and roving-level), the phase of the distractor was fixed at zero.

D. Experimental Procedures

In the psychophysical experiment, a 4I-2AFC paradigm (*ABAA* or *AABA*) was used (Bernstein and Trahiotis 1982). There was 500 ms of quiet between each interval. In the reference intervals, the target level was held fixed at the reference level; in the test interval, the target level was incremented by ΔL. A 2-down 1-up adaptive procedure, modeled after Levitt (1971), estimated the minimum change in the target level required to achieve a probability of a correct response of 0.7. The adaptive runs consisted of 16 reversals, and each adaptive run began with a random initial value of ΔL (chosen uniformly between 15 and 25 dB). According to the 2-down 1-up adaptive rule, ΔL was initially adjusted by multiplying/dividing its current value by a scale factor of 1.8. After two reversals occurred, the increment was adjusted by multiplying/dividing by a scale factor of 1.4. The magnitude of the scale factor was further reduced when the fourth, sixth, and eighth reversals occurred to values of 1.2, 1.1, and 1.05, respectively. After the eighth reversal, the scale factor remained at 1.05 for eight additional reversals, at which point the adaptive run was concluded. The adaptive trials were self-paced and the subjects had an unlimited time to respond. The subjects received correct-answer feedback after every trial.

During each testing session, subjects completed four adaptive runs for each of the five different conditions. The ordering in which the conditions were presented was fixed as follows: no-distractor, fixed, roving-phase, roving-level, and double-rove. Since the perceptual cues for each condition were different, subjects were alerted to the condition using a unique letter for each condition. Subjects ran the four adaptive runs for each distractor condition in succession. The blocking of the adaptive runs in this manner was done to allow the subjects to re-familiarize themselves with the relevant perceptual cues. Subjects were given a minimum of 10 hours of training before the reported data were collected. For each condition, the reported results are based on 16 post-training adaptive runs.

E. Data Analysis

Level discrimination thresholds have been reported in many forms and the analysis reported here follows the recommendations of Buus and Florentine (1991). In accordance with their recommendations, the performance metric used is ΔL, the decibel change of the target level. The data from the experiment are reported in two ways. First, level discrimination thresholds were determined from the geometric mean of the values of ΔL that occurred on the last eight reversals of each adaptive run. Second, psychometric functions were fitted to the data from all the trials of the adaptive runs. Note that, due to the experimental paradigm, only a few trials were conducted with each value of ΔL. (The adaptive runs began with a random initial value of ΔL; on each trial ΔL was adjusted according to the adaptive rule.) Further, on any trial, the subject was either correct or incorrect. The fitting of the psychometric functions was done with these binary data.

In accordance with Buus and Florentine (1991), d' is assumed to be proportional to ΔL. Due to the binary nature of the data, converting the data into d' is problematic since trials in which the subjects were correct have a d' of infinity. By assuming (1) that d' is proportional to ΔL, (2) an unbiased observer (which is reasonable in a 4I-2AFC task), and (3) that performance is limited by Gaussian noise, one can relate d' to the probability of a correct response. Further, with these assumptions, the probability of a correct response depends on ΔL, and can be expressed as

$$P_{Correct}(\Delta L) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\Delta L} e^{-\frac{x^2}{2\sigma^2}} dx, \tag{1}$$

where σ is a fitting parameter and is related to the proportionality between d' and ΔL. When fitting psychometric functions to the subject data, the parameter σ was adjusted to minimize the root-mean-squared (RMS) error between the predicted percent correct and all of the data for a given subject and condition (collapsed over adaptive runs).

For each subject and condition, analysis of the fitted psychometric function is based on the parameter σ and the RMS error statistic. Confidence intervals for both σ and the RMS error were calculated by randomly drawing, with replacement, the results of N trials (where N is the total number of trials for a given subject and condition) and fitting a psychometric function to this sampled data. For each subject and condition a total of 1000 random drawing were made. An additional estimate of threshold is obtained by substituting the appropriate value of σ into Eq. 1 and then solving for the ΔL that yields a probability of a correct response of 0.7. These estimates of threshold based on the psychometric functions agree with the measurements of threshold based on the reversals of the adaptive runs and are not explicitly presented. The binary nature of the data complicates the interpretation of the RMS error statistic. The RMS error statistic is not a direct measure of the difference between the data and the fitted function; even if the data were drawn from a random process defined by the fitted function, the expected RMS error would not be zero.

Also in accordance with Buus and Florentine (1991) the value ΔL is displayed on a logarithmic scale. The analysis and statistics are therefore based on the logarithm of ΔL. For example, threshold was estimated from the geometric mean of the reversals of the adaptive runs and when averaging threshold across conditions, the geometric mean was also used. The results were also analyzed by considering ΔL on a linear display (i.e. using the arithmetic mean). Both analyses lead to similar conclusions.

III. Modeling

This work considers a detection theoretic model based on a 2-dimensional decision space in a 4I-2AFC paradigm. The two dimensions, which loosely correspond to the perceived loudness and lateral position, are denoted Λ and Θ. These two dimensions are functions of the level of the left ear L_{Left} (in decibels), the level of the right ear L_{Right} (in decibels), the interaural time difference T (in microseconds), and two internal noises N_Λ and N_Θ (both in decibels). The dimensions are defined as

$$\Lambda = 10\log_{10}\left(10^{\frac{L_{Left}}{10}} + 10^{\frac{L_{Right}}{10}}\right) + N_\Lambda \tag{2}$$

and

$$\Theta = L_{Left} - L_{Right} + k\mathrm{T} + N_\Theta . \tag{3}$$

where k is the intensity-time trading ratio in dB per μs which is related to the time-intensity trading ratio used by Hafter (1971) and Yost (1972).

As mentioned in the Introduction, models based on the Θ dimension (often referred to as a position variable) have been shown to have strong predictive power. Models based on the Λ dimension (the binaural level or energy) have been evaluated as an aspect of central spectrum models (Bilsen 1977; Bilsen and Raatgever 2000). These studies of the central spectrum have focused on issues concerning across-frequency integration. To our knowledge, the information carried by the Λ dimension with narrowband stimuli has not been explicitly explored. Under monotic conditions, Λ is the monaural energy and models based on a monaural energy detector have been explored for many monotic and diotic tasks. The analysis is limited to the Λ and Θ dimensions, since they correlate well with the dominant perceptions of the stimuli. The analysis does not include dimensions that correlate with any additional percepts or images (e.g., time-image, spatial width, and spatial shape).

The model dimensions are further specified by using internal noises N_Λ and N_Θ that are zero-mean Gaussian random variables that are independent across the dimensions, the observation intervals, and the trials. With these assumptions, the model has only three parameters k, σ_Λ, and σ_Θ which correspond to the intensity-time trading ratio and the standard deviations of the two internal noises. These three parameters are chosen and held fixed throughout all the modeling such that the model predicts the results from studies of time-intensity trading and lateral position discrimination (cf. Blauert 1997) and studies of overall-level discrimination (cf. Viemeister 1988). The intensity-time trading ratio k is fixed at 1 dB per 20 μs, and the standard deviations of the internal noises (σ_Λ and σ_Θ) are fixed at 0.5 dB.

To complete the model specification, three maximum likelihood observers are considered as the decision device. These maximum likelihood observers are the ideal observer of (1) Λ alone, (2) Θ alone, and (3) Λ and Θ together. Note that none of these observers has direct access to the level of the monaural target; the two dimensions on which the model is based are both binaural dimensions. Performance of the three ideal observers, in a 4I-2AFC task, is derived in Appendix A. Model predictions are then compared to subject performance.

IV. Results

Figure 1 contains the geometric mean of the thresholds (calculated from the reversals of the adaptive runs) of the four subjects for each of the five different conditions. A global analysis of variance (ANOVA) test found statistically significant effects of distractor condition and subject and a statistically significant interaction between distractor condition and subject ($p<<0.02$). The measured thresholds vary with the type of distractor and subject. Multiple post hoc ANOVA tests were used to test for statistically significance differences across conditions (all combinations) and subjects (all combinations). There are statistically significant differences ($p<0.0002$) in performance across all conditions and subjects except between the roving-level and double-rove conditions ($p>0.05$). The post hoc ANOVA tests were not repeated measures, and therefore differences which were not found to be statistically significant with the non-repeated measures ANOVA, may in fact be statistically significant with a repeated measures ANOVA. Performance in the no-distractor condition was the best with an average (across-subject geometric-mean) threshold value of ΔL of 0.7

dB. Performances with the roving-level and double-rove distractors were the worst with average thresholds of 5.8 and 7.3 dB, respectively. The fixed and roving-phase distractors had small detrimental affects on performance, with average thresholds of 1.1 dB and 1.6 dB, respectively.

In addition to the thresholds of the subjects, Fig. 1 also shows thresholds for the ideal observer of either Λ or Θ individually, as well as the thresholds for an ideal observer of Λ and Θ jointly. The lowest threshold the ideal observer of Λ obtains is in the no-distractor condition. In this condition, the predicted threshold is 0.5 dB. In both the fixed and roving-phase conditions, the threshold is 0.9 dB, while in both the roving-level and double-rove conditions, the thresholds are 11.9 dB. The ideal observer of Θ performs best in the fixed condition with a threshold of 0.6 dB. The ideal observer of Θ never obtains threshold performance in the no-distractor condition. In the roving-phase, roving-level, and double-rove conditions, thresholds are 15.8, 11.9, and 15.8 dB, respectively using Θ alone. The ideal observer of Λ and Θ together performs best in the fixed condition with a threshold of 0.3 dB, while in the no-distractor condition the threshold is 0.4 dB. In the roving-phase, roving-level, and double-rove conditions, thresholds were, respectively, 0.8, 0.5, and 4.5 dB.

Figure 2 shows an example of psychometric functions for a representative subject (S2) for all five conditions. The psychometric functions take into account all trials, whereas the threshold measurements take into account only trials at which reversals in performance occurred. Visually, the data appear to be well fitted by psychometric functions defined by Eq. 1. Since the data were collected using an adaptive paradigm, most trials had values of ΔL near threshold. Traditionally psychometric functions are not constructed from data collected with adaptive paradigms. In this work, however, this analysis was required to determine the extent to which performance was monotonic in ΔL. The values of σ (the single fit parameter for the psychometric function) that best fit the data are presented in the top panel of Fig. 3. Consistent with the threshold data presented in Fig. 1, the value of σ that best fits the data varies systematically across the conditions. Paired t-tests show no statistically significant differences ($p>0.05$) in the values of σ between the no-distractor, fixed, and roving-phase conditions, nor are there statistically significant differences between the roving-level and double-rove conditions ($p>0.05$). The differences in the values of σ between the no-distractor, fixed, and roving-phase conditions and the roving-level and double-rove conditions, however, are statistically significant ($p<0.025$).

In addition to the changes in σ, there are also changes in the RMS error between the fitted psychometric function and the data. The bottom panel of Fig. 3 shows the RMS error. Paired t-tests show no statistically significant differences ($p>0.05$) in the RMS error between the no-distractor, fixed, and roving-phase conditions, nor between the roving-level and double-rove conditions. The differences in the RMS error between the no-distractor, fixed, and roving-phase conditions and the roving-level and double-rove conditions, however, are statistically significant ($p<0.025$). These changes in the RMS error are indicative of a change in how well the fitted psychometric function fits the data. These changes could be a result of either a change in the underlying psychometric function, or an artifact of the adaptive paradigm, which arises from the change in σ. Independent of the source, these changes in the quality of the fit are relatively small and do not indicate a drastic departure from the assumed shape of the psychometric function.

In addition to the thresholds reported in Fig. 1, the model also predicts psychometric functions for the five different distractor conditions. Figure 4 shows predicted psychometric functions of the ideal observer of Λ alone (dashed), Θ alone (dotted), and Λ and Θ together (solid). The predicted psychometric functions of the ideal observer of Λ alone visually agree with the reconstructed psychometric functions of the subjects in the no-distractor, fixed, and roving-phase conditions (cf. Figs. 2 and 4). The predicted psychometric functions of the ideal observer of Θ alone agree only with the psychometric functions of the subjects in the fixed condition. The subject psychometric functions for the roving-level and double-rove conditions are neither predicted by the ideal observer of Λ alone, nor the ideal observer of Θ alone. For all conditions except the roving-level condition, the predicted psychometric functions of the ideal observer of Λ and Θ together are consistent with the psychometric functions of the subjects. The predicted psychometric functions of the ideal observer of Λ and Θ together for the roving-level condition has an appropriate shape, but shows substantially better performance than that obtained by the subjects.

V. Explanation of the Model Predictions

Insight into the model predictions (cf. Figs. 1 and 4) is gained by considering the three ideal observers (Λ alone, Θ alone, and Λ and Θ together) that are derived in Appendix A. All three ideal observers are completely characterized by the conditional joint probability density function of observing Λ and Θ, given the target level L_{Target}. A closed form expression of this conditional joint probability density function, denoted $f_{\Lambda,\Theta|L_{Target}}$, could not be derived analytically. Rather, $f_{\Lambda,\Theta|L_{Target}}$ was approximated with a mixture of analytical and numerical techniques; a detailed derivation of this calculation is presented in Appendix B. This section considers how the joint and marginal probability density functions of Λ and Θ, given L_{Target}, are affected by changes to ΔL and the stimulus condition.

Figure 5 shows the conditional probability density for Λ given L_{Target}, denoted as $f_{\Lambda|L_{Target}}$, for the five conditions and a range of ΔL sizes. In the no-distractor, fixed, and roving-phase conditions, $f_{\Lambda|L_{Target}}$ has a Gaussian shape. Changing the size of ΔL shifts the mean (Fig. 5 panels A-C). The width (variance) of $f_{\Lambda|L_{Target}}$ is independent of the size of ΔL. For a given L_{Target} (i.e. ΔL), the shape is Gaussian because variation in the random variable Λ depends only on the internal Gaussian noise. In the fixed and roving-phase conditions, the distributions (and the effects of the increment size on the distributions) are identical. For a given ΔL, the mean in the no-distractor condition is slightly less than the mean in the fixed and roving-phase conditions. These across-condition differences in the mean are greatest for small ΔL. The difference is due to the presence of the distractor, and as ΔL increases, the distractor influence decreases.[3] The shifts in the distributions associated with a change in the increment size are slightly smaller in the fixed and roving-phase conditions than the shifts in the no-distractor condition and, therefore, an ideal observer of only Λ performs slightly worse in the fixed and roving-phase conditions.

In the roving-level and double-rove conditions, the distributions of Λ (Fig. 5 panels D and E) are more complicated; the distributions are much broader than the distributions of Λ in the no-distractor condition. The increased complexity of $f_{\Lambda|L_{Target}}$ in the roving-level and double-rove conditions arises because the random variable Λ depends not only on L_{Target} and the internal noise (as is the case in the other conditions) but also on the random level of the distractor. The level of the distractor, in decibels, was chosen from a uniform distribution. A uniform distribution of level is neither a uniform distribution of pressure (units of force per area) nor intensity (units of power per area). The random variable Λ adds the intensities, not the levels, of the target and distractor. In the roving-level and double-rove conditions, the distributions of Λ and the effect of changes to the increment size on the distributions are identical. When ΔL is changed, the mean of the distribution shifts and the shape of the distribution is also affected. In these two conditions, changes in ΔL have a much smaller influence than in the no-distractor condition. The performance of an ideal observer of Λ is considerably worse in these two conditions.

A detection theoretic analysis reveals that performance is limited by the width and how the mean depends on ΔL; further, the analysis demonstrates that performance is neither affected by the changes in shape of $f_{\Lambda|L_{Target}}$ associated with changes in ΔL nor the complexity of the shape of $f_{\Lambda|L_{Target}}$. Many $f_{\Lambda|L_{Target}}$ shapes (e.g., rectangular) will lead to the same predictions. [For example, the psychometric functions for the ideal observers of Λ alone and Θ alone in the roving-level condition (refer to Fig. 3) are nearly identical even though $f_{\Lambda|L_{Target}}$ and $f_{\Theta|L_{Target}}$ are strikingly different in this condition (see Figs. 5 and 6).]

Figure 6 shows $f_{\Theta|L_{Target}}$, the conditional probability density function for the observed Θ given L_{Target}, for four of the five different conditions and a range of increment sizes. In the no-distractor condition, the decision variable Θ is undefined and carries no information, and this case is not included in Fig. 6. In the fixed condition (Fig. 6 panel A), $f_{\Theta|L_{Target}}$ is Gaussian; changing the increment size shifts the mean and has no effect on its variance.[4] For a given L_{Target} the shape is Gaussian because the randomness of the variable Θ comes only from the Gaussian internal noise. In the roving-phase condition (Fig. 6 panel B), $f_{\Theta|L_{Target}}$ is nearly uniform over a large range. For a given L_{Target}, the variable Θ depends on both internal noise and the phase of the distractor. The predictions are robust to the method used to empirically extract the phase since the stimuli are tones with relatively long durations. These predictions were derived by using the actual phase of the target and distractor stimuli. Note that in the roving-phase condition, Θ carries substantially less information about ΔL than Λ carries.

In the roving-level condition (Fig. 6 panel C), $f_{\Theta|L_{Target}}$ is also nearly uniform over a large range (although the range is smaller than in the roving-phase condition). In this case, for a given L_{Target}, the variable Θ depends on both internal noise and the level of the distractor. In the roving-level condition, $f_{\Theta|L_{Target}}$ is nearly uniform, as opposed to $f_{\Lambda|L_{Target}}$, since the random variable Θ is calculated by subtracting the levels in decibels, not the intensities (units of power per area), of the target and distractor. Note that changes in ΔL result in only small changes in the marginal distributions of both Λ and Θ. Finally, in the double-rove condition (Fig. 6 panel D), the marginal distributions of Θ are nearly trapezoidal in shape since they arises from the sum of two uniformly distributed random variables. Changes to ΔL have only small effects on this distribution.

In the roving-level condition, neither Λ nor Θ carry useful information individually. The variables Λ and Θ, however, carry information jointly and therefore, there is substantial information for discriminating changes in the target level. The joint probability distributions of Λ and Θ for a range of increment sizes for the roving-level condition are shown in Fig. 7. Although both Λ and Θ take on a large range of values, for each increment size, the probability of Θ conditioned on Λ is narrow (and vice-versa) and changes in ΔL substantially shift the distribution. If the effects of the internal noise are disregarded, for each increment size, every value of Λ corresponds to a single value of Θ (performance is limited only by the internal noise). Thus, for each increment size, there is a unique $\{\Lambda,\Theta\}$, making monaural level discrimination possible with these two dimensions.

In the double-rove condition, the variables Λ and Θ do not carry useful information (neither individually nor jointly). The joint probability distributions of Λ and Θ for a range of increment sizes for the double-rove condition are shown in Fig 8. The variables Λ and Θ take on large ranges of values, as in the roving-level condition. Unlike the roving-level condition, however, in the double-rove condition, the probability of Θ conditioned on Λ is broad (each value of Λ now corresponds to a range of Θ values and vice-versa) and changes in the increment size have only small effects (cf. Fig. 8). Therefore, performance on monaural level discrimination based on Λ and Θ is greatly reduced. Note that adding almost any third dimension would result in monaural level discrimination performance in the double-rove condition being limited only by internal noise.

In summary, the Λ and Θ dimensions carry information both individually and jointly about the target level (cf. Figs. 5-8). The roving-phase condition decreases the information in Θ, without compromising the information in Λ. The roving-level condition reduces the information in Λ and Θ individually, but does not reduce the joint information. The double-rove condition reduces the information in Λ and Θ individually and also reduces the joint information. Note that in all the conditions, however, the signal at the target ear is unchanged; therefore, any decreases in performance are a result of an inability of the subjects to access the information available at the target ear.

VI. Discussion

The results are consistent with the hypothesis that monotic level discrimination uses information carried by binaural channels. The experimental stimuli were chosen such the monaural level of the target ear was always reliable. Thus, if a subject were able to focus exclusively on the ipsilateral ear (the ear in which the target is presented), performance would not vary across conditions. Alternatively, attending to only the contralateral ear would result in chance performance in all conditions because the distractor never carried information useful for level discrimination. The results presented in Figs. 1 and 3 clearly demonstrate that it is possible to measure large amounts of contra-aural interference with a tonal stimulus. Both the analysis of the measured thresholds and the analysis of the reconstructed psychometric functions show a significant effect of the contralateral ear on monaural level discrimination (cf. Figs. 1 and 2).

The contra-aural interference is greatest when the reliability of the perceived loudness and lateral position is reduced. In the no-distractor and fixed conditions, subjects could have relied upon any perception that varied with the level at the target ear. For these conditions, subjects typically reported listening for a change in the loudness (no-distractor condition) or changes in both the loudness and lateral position (fixed condition). These well-studied conditions were repeated (1) to provide a baseline for comparison, (2) to decrease the effects of inter-subject variability, and (3) as a training condition. The measured thresholds in the no-distractor condition agree with previous measures of monotic (and diotic) level discrimination (Viemeister 1988). The measured thresholds in the fixed condition also agree with previous measures of interaural level discrimination (Blauert 1997). The small, but significant, differences between the no-distractor and fixed conditions agree with previous reports of small increases in threshold due to the addition of a distractor (Rowland and Tobias 1967; Yost 1972; Bernstein 2004; Stellmack et al. 2004).

The measured thresholds in the roving-phase condition are approximately twice as large as the thresholds in the no-distractor condition (1.6 dB versus 0.7 dB). Although this increase in threshold is contra-aural interference, there may be other interpretations. For example, this relatively small increase in threshold could be due to the perceptual complexity of the task when the distractor has a random phase. In the no-distractor condition, three of the four intervals are identical; whereas in the roving-phase condition, all four intervals are different and subjects must focus on the correct perceptual attribute. The further increases in the thresholds in the roving-level and double-rove conditions, however, challenge this perceptual complexity argument. If subjects are basing their decisions on the monaural level, the roving-level and double-rove conditions have the same perceptual complexity as the roving-phase condition and, therefore, the thresholds should be the same in these three conditions.

Another competing hypothesis is that the decrease in performance in the roving-level and double-rove conditions is due to crosstalk. A theoretical analysis of crosstalk, however, suggests that its contributions are not sufficient to explain the results. The predictions of the ideal observer of Λ are equal to the predictions of the monaural energy detector when there is no attenuation of the crosstalk signal. Since the predictions of the ideal observer of Λ are so close to the measured thresholds, the amount of attenuation of the crosstalk signal would need to be near zero, which is unreasonable.

One hypothesis consistent with all observations is that monotic level discrimination uses information carried by binaural channels. Predictions of an ideal observer of Λ and Θ agree with the measured thresholds (cf. Fig. 1) except for three discrepancies. The first is that subject S3 outperformed the ideal observer in the no-distractor condition. The second is that the model predicts a slight improvement (lower threshold) due to the addition of a fixed distractor and the data indicates a slight decrease (higher threshold) in performance. The third is that the predicted performance in the roving-level condition is much better than the actual performance. These three discrepancies are considered in the following two paragraphs.

The inconsistency between the measured threshold of subject S3 and the threshold predicted by the ideal observer of Λ and Θ is a result of how the parameters of the model were selected. The values of the parameters were fit to the average performance across many studies and subjects. In the no-distractor condition, the performance of the ideal observer is limited only by internal noise in the Λ dimension and therefore highly influenced by the σ_Λ parameter. A small decrease in the amount of internal noise would result in the ideal observer having a lower threshold than subject S3. In fact, the predictions of the no-distractor, fixed, and roving-phase conditions would be improved by tuning the model parameters for each subject.

The discrepancies in the fixed condition and roving-level condition are related to the assumption that the joint information in Λ and Θ is used optimally. The amount of joint information[5] is minimal in the no-distractor and roving-phase conditions and maximal in the fixed, roving-level, and double-rove conditions. Further, the predictions in the double-rove condition are less sensitive to perturbations of the ideal decision rule than the predictions in the fixed and roving-level conditions.

This invariance to non-ideal perturbations is best highlighted by a one-interval analysis, even though the experiment used a four-interval paradigm. Figure 8 shows the joint probability density function of Λ and Θ in the double-rove condition for the un-incremented target and targets with level increments of 4 and 16 dB. The probability density functions are nearly identical for many values of Λ and Θ. For example, comparing the un-incremented target distribution and the distribution of a target with a level increment of 4 dB, the probability density functions are nearly identical for Λ above 65 dB and Θ between -20 and 20 dB. Making non-ideal decisions in the regions where the probability of the reference target is similar to the probability of the incremented target does not drastically decrease performance. In the roving-level condition, however, perturbations of the ideal decision rule drastically affect performance. There are no regions where the probability of the un-incremented target is similar to the probability of the incremented target. (See Fig. 7.)

The departure from ideal processing may also be related to the ability to reliably observe Λ and Θ on the same interval. In the studies of overall-level discrimination and lateral position discrimination used to select the parameters of the model, observations of only a single dimension were required to obtain ideal performance (e.g., the Λ dimension carries all the relevant information in overall-level discrimination studies). It is not known whether subjects can reliably observe Λ and Θ in our paradigm. Clearly, subjects can rate the loudness or the position of a tonal stimulus. In the paradigm used in this work, in order to make optimal decisions, subjects would be required to accurately rate both the loudness and position of a single observation of a short tonal stimulus on each interval. This joint rating would most likely need to be done rapidly (during the brief quiet between each interval) or slowly, but delayed (i.e., after the trial and therefore requiring more memory). Studies on divided attention (e.g., Bonnel and Hafter 1998) report that performance decreases when subjects are required to judge multiple dimensions.

Most of the model analysis has focused on comparing the predicted thresholds to the measured thresholds. The model also predicts psychometric functions (cf. Fig. 4), and it is possible to compare the predicted and measured psychometric functions. Due to the experimental design, the scope of this comparison is limited since comprehensive measurements of the psychometric functions were not obtained with the adaptive paradigm. One can make, however, some general statements about the psychometric functions. The first is that the predicted psychometric functions of the ideal observer of Λ and Θ have the general shape of the psychometric functions of the subjects (compare Figs. 2 and 4) for all conditions. Further, the shapes of the predicted psychometric functions of the ideal observer of only Λ or Θ are incorrect in many conditions. Second, there are slight variations across conditions in the shapes of the psychometric functions predicted by the ideal observer of Λ and Θ. This is consistent with the change, across conditions, in the quality of fit of the psychometric function defined by Eq. 1 and the data (refer to Fig. 3). As mentioned previously, the change in the quality of fit across conditions could also be an artifact of the adaptive paradigm. Most importantly, it appears reasonable to assume that the psychometric functions in all the conditions are monotonic in ΔL and smooth. Therefore, the measured increases in threshold due to the addition of the distractor are neither an artifact of adaptive procedure nor an artifact of a major perceptual changes as a function of ΔL. The influence of the distractor on performance cannot be explained by monaural processing. The psychophysical results are consistent with a binaural model based on correlates to the perceived loudness and lateral position.

VII. Summary

This work measured large amounts of contra-aural interference in a monaural level discrimination task. Subjects were unable to ignore the ear contralateral to the target. As a result of this failure, thresholds were increased by nearly an order of magnitude in some conditions. In general, subjects perceived a single compact auditory image. The reliability of the perceived loudness and lateral position of this image were systematically varied. The increases in the measured thresholds were greatest when both the perceived loudness and lateral position were unreliable. These results are predicted by a detection theoretic model based on only binaural information. Further, the results are consistent with the hypothesis that monotic level discrimination uses information carried by binaural channels.

Acknowledgements

This research was supported by NIH/NIDCD Grants DC00100, DC004663, and DC6769. The authors would also like to thank Dr. Frederick Gallun for reading a previous version of this manuscript.

Appendix A

In this appendix, performance in a 4I-2AFC task is calculated for three observers in the $\{\Lambda, \Theta\}$ space, as defined by Eqs. 2 and 3 in the text. Three models are considered: the ideal observer of Λ and Θ together, the ideal observer of Λ alone, and the ideal observer of Θ alone. The task requires the observer to discriminate the level of the target L_{Target}. On a given interval the target is either un-incremented, such that L_{Target} is equal to L_0 (the reference level) or the target is incremented, such that L_{Target} is equal to the sum of L_0 and ΔL. To calculate performance we note: (1) on each interval there is a single observation of both Λ and Θ and (2) on a single trial there are eight total observations (four of Λ and four of Θ). Due to the experimental paradigm, the observations in the first and last intervals carry no information for the ideal observers[6] and, therefore, only four observed values (two pairs) are relevant. The observation of Λ on the second interval is λ_2; the observation of Λ on the third interval is λ_3. Similarly, the observation of Θ on the second interval is θ_2 and on the third interval is θ_3. The ideal observer of Λ and Θ together is considered first, since the performance of the ideal observers of Λ alone and Θ alone follow from the ideal observer of Λ and Θ.

The ideal observer of both Λ and Θ depends on two 4-dimensional joint probability functions. The first is the probability densities of λ_2, λ_3, θ_2, and θ_3 given that an increment of size ΔL occurred on the second interval; the second is the probability density of λ_2, λ_3, θ_2, and θ_3 given that an increment of size ΔL occurred on the third interval. These 4-dimensional joint probability functions can be written as the product of two 2-dimensional joint probability functions by noting that when the interval in which the increment occurred is given, the observation of λ_2 is independent of λ_3 and the observation θ_2 is independent of θ_3. Since there are two intervals in which the target level can be incremented, there are four relevant 2-dimensional probability density functions.

The relevant 2-dimensional conditional joint probabilities are the probability of observed values of Λ and Θ on a particular interval given a target level. The log-likelihood ratio $\eta_{\Lambda,\Theta}$ is defined in terms of these probabilities as

$$\eta_{\Lambda,\Theta}(\lambda_2,\theta_2,\lambda_3,\theta_3,\Delta L)=10\log_{10}\left(\frac{f_{\Lambda,\Theta|L_{Target}}(\lambda_2,\theta_2|L_0)\,f_{\Lambda,\Theta|L_{Target}}(\lambda_3,\theta_3|L_0+\Delta L)}{f_{\Lambda,\Theta|L_{Target}}(\lambda_2,\theta_2|L_0+\Delta L)\,f_{\Lambda,\Theta|L_{Target}}(\lambda_3,\theta_3|L_0)}\right).$$

The ideal observer is defined by a binary indicator function $\psi_{\Lambda,\Theta}$, which depends on the likelihood ratio. Specifically, when the second interval is most likely to have the incremented target $\psi_{\Lambda,\Theta}$ is equal to zero. Similarly, when the third interval is most likely to have the incremented target $\psi_{\Lambda,\Theta}$ is equal to one. Mathematically the indicator function is

$$\psi_{\Lambda,\Theta}(\lambda_2,\theta_2,\lambda_3,\theta_3,\Delta L)=\begin{Bmatrix}1 & \text{when} & \eta_{\Lambda,\Theta}(\lambda_2,\theta_2,\lambda_3,\theta_3,\Delta L)\geq 0\\ 0 & \text{when} & \eta_{\Lambda,\Theta}(\lambda_2,\theta_2,\lambda_3,\theta_3,\Delta L)<0\end{Bmatrix}.$$

The probability that the ideal observer of both Λ and Θ results in the correct answer is a function of ΔL and can be written as

$$\begin{aligned}P_{Correct}(\Delta L)=&\frac{1}{2}\iiiint\begin{bmatrix}\psi_{\Lambda,\Theta}(\lambda_2,\theta_2,\lambda_3,\theta_3,\Delta L)\\ \times f_{\Lambda,\Theta|L_{Target}}(\lambda_2,\theta_2|L_0)\,f_{\Lambda,\Theta|L_{Target}}(\lambda_3,\theta_3|L_0+\Delta L)\end{bmatrix}d\theta_2 d\lambda_2 d\theta_3 d\lambda_3\\ &+\frac{1}{2}\iiiint\begin{bmatrix}(1-\psi_{\Lambda,\Theta}(\lambda_2,\theta_2,\lambda_3,\theta_3,\Delta L))\\ \times f_{\Lambda,\Theta|L_{Target}}(\lambda_2,\theta_2|L_0+\Delta L)\,f_{\Lambda,\Theta|L_{Target}}(\lambda_3,\theta_3|L_0)\end{bmatrix}d\theta_2 d\lambda_2 d\theta_3 d\lambda_3\end{aligned}.$$

The probability of a correct response depends only on the joint probability density function of Λ and Θ given L_{Target}. This conditional joint probability density function $f_{\Lambda,\Theta|L_{Target}}$ is derived in Appendix B. The probabilities of correct for the ideal observers of Λ alone and Θ alone are calculated in an analogous manner and the derivation is not presented. The probability of a correct response of the ideal observer of Λ alone depends only on the probability of observing Λ given L_{Target}. This conditional probability $f_{\Lambda|L_{Target}}$ is a marginal of $f_{\Lambda,\Theta|L_{Target}}$. Similarly, the probability of a correct response of the ideal observer of Θ alone depends only on the probability of observing Θ given L_{Target}. This conditional probability $f_{\Theta|L_{Target}}$ is also a marginal of $f_{\Lambda,\Theta|L_{Target}}$.

Appendix B

In this appendix, analytical and numerical techniques are used to approximate the joint density function of Λ and Θ, as defined in Eqs. 2 and 3, for a target level L_{Target} equal to the sum of L_0 (the reference level) and an increment Δ*L*. (In this notation, Δ*L* is zero when the target is un-incremented.) Before the details of the derivation of $f_{\Lambda,\Theta|L_{Target}}$ are outlined, the model variables are related to the experimental variables. Specifically, the values that are appropriate for the psychophysical experiment are substituted into Eqs. 2 and 3. In the experiment, L_{Left} is the level of the target L_{Target} and the level of the distractor is L_{Right}. The level of the target is the sum of a reference level L_0 and an increment Δ*L*, such that $L_{Left} = L_{Target} = L_0 + \Delta L$. The distractor has a level that is equal to the reference level plus a random variable *A*; the level of the right ear can, therefore, be written as $L_{Right} = L_0 + A$. The interaural time difference T is the negative of the distractor phase -Φ divided by the radian frequency ω (i.e., $\mathrm{T} = {-\Phi}/{\omega}$). The psychophysical experiment specifies that *A* has a uniform probability density function between a_{min} and a_{max}. The experiment also specifies that Φ has a uniform probability density function between ϕ_{min} and ϕ_{max}. Using this notation, when the distractor level is roved $a_{min} = 0$ and $a_{max} = 30$ and when the distractor level is fixed,

$a_{min} = a_{max} = 0$. Similarly, when the distractor phase is roved $\phi_{min} = -\pi/2$ and $\phi_{max} = \pi/2$ and when the distractor phase is fixed, $\phi_{min} = \phi_{max} = 0$. Making these substitutions into Eqs. 2 and 3 results in

$$\Lambda = 10\log_{10}\left(10^{\frac{L_{Target}}{10}} + 10^{\frac{L_0+A}{10}}\right) + N_\Lambda \tag{B1}$$

and

$$\Theta = L_{Target} - (L_0 + A) - \frac{k}{\omega}\Phi + N_\Theta . \tag{B2}$$

This derivation of $f_{\Lambda,\Theta|L_{Target}}$ begins by using the definition of conditional probability to expand the joint density function to

$$f_{\Lambda,\Theta|L_{Target}} = f_{\Lambda|L_{Target}}\left(\lambda|L_0 + \Delta L\right) f_{\Theta|\Lambda,L_{Target}}\left(\theta|\lambda, L_0 + \Delta L\right).$$

Then expanding $f_{\Lambda,\Theta|L_{Target}}$ into its marginal distribution with respect to A and Φ results in

$$f_{\Lambda,\Theta|L_{Target}} = f_{\Lambda|L_{Target}}\left(\lambda|L_0 + \Delta L\right)\iint\begin{bmatrix} f_{\Theta|A,\Phi,\Lambda,L_{Target}}\left(\theta|a,\phi,\lambda,L_0 + \Delta L\right) \\ \times f_{A,\Phi|\Lambda,L_{Target}}\left(a,\phi|\lambda,L_0 + \Delta L\right) \end{bmatrix} d\phi da$$

Making a substitution of Θ based on Eq. B2 and using the definition of conditional probability yields

$$f_{\Lambda,\Theta|L_{Target}} = f_{\Lambda|L_{Target}}\left(\lambda|L_0 + \Delta L\right)\iint f_{N_\Theta}\left(\frac{k}{\omega}\phi - \mu_\Theta(a)\right) f_{A,\Phi|\Lambda,L_{Target}}\left(a,\phi|\lambda,L_0 + \Delta L\right) d\phi da,$$

where $\mu_\Theta(a)$ is equal to $\Delta L - a - \theta$. Using the definition of conditional probability for $f_{A,\Phi|\Lambda,L_{Target}}$, and then noting the independence of Φ on A, Λ, and L_{Target} gives

$$f_{\Lambda,\Theta|L_{Target}} = f_{\Lambda|L_{Target}}\left(\lambda|L_0+\Delta L\right)\int\int f_{N_\Theta}\left(\frac{k}{\omega}\phi-\mu_\Theta(a)\right)f_{A|\Lambda,L_{Target}}\left(a|\lambda,L_0+\Delta L\right)f_\Phi(\phi)\,d\phi da.$$

Using the definition of conditional probability on $f_{A|\Lambda,L_{Target}}$, noting the independence of $f_{A|L_{Target}}$ on L_{Target} and simplifying yields

$$f_{\Lambda,\Theta|L_{Target}} = \int f_A(a)\,f_{\Lambda|A,L_{Target}}\left(\lambda|a,L_0+\Delta L\right)\int f_\Phi(\phi)\,f_{N_\Theta}\left(\frac{k}{\omega}\phi-\mu_\Theta(a)\right)d\phi da.$$

Making use of the uniform probability density functions of the random variables A and Φ, $f_{\Lambda,\Theta|L_{Target}}$ can be rewritten as

$$f_{\Lambda,\Theta|L_{Target}} = c\int_{a_{min}}^{a_{max}} f_{\Lambda|A,L_{Target}}\left(\lambda|a,L_0+\Delta L\right)\int_{\phi_{min}}^{\phi_{max}} f_{N_\Theta}\left(\frac{k}{\omega}\phi-\mu_\Theta(a)\right)d\phi da,$$

where c is equal to $1/(a_{max}-a_{min})(\phi_{max}-\phi_{min})$. Substituting the density function of N_Θ and simplifying gives

$$f_{\Lambda,\Theta|L_{Target}} = \frac{c}{\sqrt{2\pi\sigma_\Theta^{\,2}}}\int_{a_{min}}^{a_{max}} f_{\Lambda|A,L_{Target}}\left(\lambda|a,L_0+\Delta L\right)\int_{\phi_{min}}^{\phi_{max}} e^{-\frac{\left(\phi-\frac{\omega}{k}\mu_\Theta(a)\right)^2}{2\left(\frac{\omega}{k}\sigma_\Theta\right)^2}}\,d\phi da.$$

Defining the integral of an exponential squared function as

$$\mathrm{G}(\alpha,\beta,\mu,\sigma) = \int_\alpha^\beta e^{-\frac{(x-\mu)^2}{2\sigma^2}}\,dx,$$

we rewrite $f_{\Lambda,\Theta|L_{Target}}$ as

$$f_{\Lambda,\Theta|L_{Target}} = \frac{c}{\sqrt{2\pi\sigma_\Theta^{\ 2}}} \int_{a_{min}}^{a_{max}} f_{\Lambda|A,L_{Target}}\left(\lambda|a, L_0 + \Delta L\right) \mathrm{G}\left(\phi_{min}, \phi_{max}, \frac{\omega}{k}\mu_\Theta(a), \frac{\omega}{k}\sigma_\Theta\right) da .$$

Noting that $f_{\Lambda|A,L_{Target}}$ is equal to $f_{N_\Lambda}\left(\lambda - 10\log_{10}\left(10^{\frac{L_0+\Delta L}{10}} + 10^{\frac{L_0+a}{10}}\right)\right)$ (cf. Eq. B1) and making a substitution of the density function of N_Λ yields

$$f_{\Lambda,\Theta|L_{Target}} = \frac{c}{2\pi\sigma_\Theta\sigma_\Lambda} \int_{a_{min}}^{a_{max}} e^{-\left(\frac{(\lambda-\mu_\Lambda(a))^2}{2\sigma_\Lambda^{\ 2}}\right)} \mathrm{G}\left(\phi_{min}, \phi_{max}, \frac{\omega}{k}\mu_\Theta(a), \frac{\omega}{k}\sigma_\Theta\right) da$$

where $\mu_\Lambda(a)$ is equal to $10\log_{10}\left(10^{\frac{L_0+\Delta L}{10}} + 10^{\frac{L_0+a}{10}}\right)$.

Further analytical manipulations of $f_{\Lambda,\Theta|L_{Target}}$ do not appear to reduce the complexity of the solution, but, using the above equation $f_{\Lambda,\Theta|L_{Target}}$ can be approximated numerically. The first step of the numerical implementation is to approximate the definite integral over *A* through finite summation. Let us denote $a[n]$ as a sampled version of the continuous random variable *A*. Further let $a[1]$ equal a_{min} and $a[N]$ equal a_{max}. The probability density function $f_{\Lambda,\Theta|L_{Target}}$ can then be numerically approximated as

$$f_{\Lambda,\Theta|L_{Target}} \approx \frac{1}{N(\phi_{max} - \phi_{min})2\pi\sigma_\Theta\sigma_\Lambda} \sum_{n=1}^{N} e^{-\left(\frac{(\lambda-\mu_\Lambda(a[n]))^2}{2\sigma_\Lambda^{\ 2}}\right)} \mathrm{G}\left(\phi_{min}, \phi_{max}, \frac{\omega}{k}\mu_\Theta(a[n]), \frac{\omega}{k}\sigma_\Theta\right).$$

One should note that in the limit when $a_{max} - a_{min} = 0$ or $\phi_{max} - \phi_{min} = 0$ one cannot simply evaluate this expression at zero, but must rather evaluate the expression in the limit as the difference approaches zero.

---

[1] In Zurek (1979) the reference condition was diotic and not monotic.

[2] Note that the binaural level and the interaural differences in level and time delay fully define the target and distractor stimuli.

[3] The random variable Λ is calculated by adding the intensities (units of power per area), not the decibel levels, of the target and distractor. When the target is 50 dB and the distractor is absent, the expected value of Λ is 50. When the target and distractor both have a level of 50 dB, the expected value of Λ is 53. Thus, adding the 50-dB distractor to the 50-dB target increases Λ by 3. Adding a 50 dB distractor to a 66 dB target increases Λ by only 0.1.

[4] The ILD component of the random variable Θ is calculated by subtracting the decibel levels, not the intensities (units of power per area), of the target and distractor. Therefore the change in the mean of $f_{\Theta|L_{Target}}$ associated with a change in the level of the target from 50 dB to 51 dB is identical to the change in the mean of $f_{\Theta|L_{Target}}$ associated with a change in the level of the target from 70 dB to 71 dB. The effects of changing the level of the target on $f_{\Theta|L_{Target}}$ are different than the effects on $f_{\Lambda|L_{Target}}$.

[5] The joint information can be quantified by considering the joint probability density functions. The joint probability density functions for the roving-level and double-rove conditions are shown in Figs. 7 and 8, respectively. Joint probability density functions for the no-distractor, fixed, and roving-phase conditions are not explicitly derived or presented in this work.

[6] Although the first and last intervals convey no information for the ideal observer, these intervals may aid the non-ideal subjects.

References

Bernstein, J. G., and Oxenham, A. J. (**2003**). "Pitch discrimination of diotic and dichotic tone complexes: Harmonic resolvability or harmonic number?," J. Acoust. Soc. Am. **113,** 3323-34.

Bernstein, L. R., and Trahiotis, C. (**1982**). "Detection of interaural delay in high frequency noise," J. Acoust. Soc. Am. **71,** 147-52.

Bernstein, L. R. (**2004**). "Sensitivity to interaural intensitive disparities: Listeners' use of potential cues," J. Acoust. Soc. Am. **115,** 3156-60.

Bilsen, F. A. (**1977**). "Pitch of noise signals: Evidence for a "Central spectrum"," J. Acoust. Soc. Am. **61,** 150-61.

Bilsen, F. A., and Raatgever, J. (**2000**). "On the dichotic pitch of simultaneously presented interaurally delayed white noises. Implications for binaural theory," J. Acoust. Soc. Am. **108,** 272-84.

Blauert, J. (**1997**). "Spatial hearing: The psychophysics of human sound localization," (MIT Press, Cambridge).

Bonnel, A. M., and Hafter, E. R. (**1998**). "Divided attention between simultaneous auditory and visual signals," Percept Psychophys **60,** 179-90.

Brungart, D. S., and Simpson, B. D. (**2002**). "Within-ear and across-ear interference in a cocktail-party listening task," J. Acoust. Soc. Am. **112,** 2985-95.

Buus, S., and Florentine, M. (**1991**). "Psychometric functions for level discrimination," J. Acoust. Soc. Am. **90,** 1371-80.

Colburn, H. S. (**1973**). "Theory of binaural interaction based on auditory-nerve data. I. General strategy and preliminary results on interaural discrimination," J. Acoust. Soc. Am. **54,** 1458-70.

Colburn, H. S., and Durlach, N. I. (**1978**). "Models of binaural interaction" in *Handbook of perception,* Vol. IV edited by E. C. Carterette and M. P. Friedman.(Academic Press, New York)

Domnitz, R. (**1973**). "The interaural time jnd as a simultaneous function of interaural time and interaural amplitude," J. Acoust. Soc. Am. **53,** 1549-52.

Domnitz, R. H., and Colburn, H. S. (**1977**). "Lateral position and interaural discrimination," J. Acoust. Soc. Am. **61,** 1586-98.

Durlach, N. I. (**1963**). "Equalization and cancellation theory of binaural masking-level differences," J. Acoust. Soc. Am. **35,** 1206-18.

Durlach, N. I., and Colburn, H. S. (**1978**). "Binaural phenomena" in *Handbook of perception,* Vol. IV edited by E. C. Carterette and M. P. Friedman.(Academic Press, New York)

Hafter, E. R. (**1971**). "Quantitative evaluation of a lateralization model of masking-level differences," J. Acoust. Soc. Am. **50,** 1116-22.

Hafter, E. R., and Carrier, S. C. (**1972**). "Binaural interaction in low-frequency stimuli: The inability to trade time and intensity completely," J. Acoust. Soc. Am. **51,** 1852-62.

Heller, L. M., and Trahiotis, C. (**1995**). "The discrimination of samples of noise in monotic, diotic, and dichotic conditions," J. Acoust. Soc. Am. **97,** 3775-81.

Hershkowitz, R. M., and Durlach, N. I. (**1969**). "Interaural time and amplitude jnds for a 500-hz tone," J. Acoust. Soc. Am. **46,** 1464-7.

Kidd, G. J., Mason, C. R., Arbogast, T. L., Brungart, D. S., and Simpson, B. D. (**2003**). "Informational masking caused by contralateral stimulation," J. Acoust. Soc. Am. **113,** 1594-603.

Koehnke, J., and Besing, J. M. (**1992**). "Effects of roving level variation on monaural detection with a contralateral cue," J. Acoust. Soc. Am. **92,** 2625-9.

Levitt, H. (**1971**). "Transformed up-down methods in psychoacoustics," J. Acoust. Soc. Am. **49,** 467-77.

Mills, J. H., Dubno, J. R., and He, N.-J. (**1996**). "Masking by ipsilateral and contralateral maskers," J. Acoust. Soc. Am. **100,** 3336-44.

Rowland, R. C. J., and Tobias, J. V. (**1967**). "Interaural intensity difference limen," Journal of Speech and Hearing Research **10,** 745-56.

Ruotolo, B. R., Stern, R. M. J., and Colburn, H. S. (**1979**). "Discrimination of symmetric time-intensity traded binaural stimuli," J. Acoust. Soc. Am. **66,** 1733-7.

Stellmack, M. A., Viemeister, N. F., and Byrne, A. J. (**2004**). "Monaural and interaural intensity discrimination: Level effects and the "binaural advantage"," J. Acoust. Soc. Am. **116,** 1149-59.

Stern, R. M., and Trahiotis, C. (**1997**). "Models of binaural perception" in *Binaural and spatial hearing in real and virtual environments*, Vol. edited by R. H. Gilkey and T. R. Anderson.(Lawrence Erlbaum Associates, Mahwah)

Stern, R. M., Jr., and Colburn, H. S. (**1978**). "Theory of binaural interaction based in auditory-nerve data. IV. A model for subjective lateral position," J. Acoust. Soc. Am. **64,** 127-40.

Taylor, M. M., and Clarke, D. P. J. (**1971**). "Monaural detection with contralateral cue (MDCC). II. Interaural delay of cue and signal," J. Acoust. Soc. Am. **49,** 1243-53.

Taylor, M. M., Clarke, D. P. J., and Smith, S. M. (**1971a**). "Monaural detection with contralateral cue (MDCC). III. Sinusoidal signals at a constant performance level," J. Acoust. Soc. Am. **49,** 1795-804.

Taylor, M. M., Smith, S. M., and Clarke, D. P. (**1971b**). "Monaural detection with contralateral cue (MDCC). IV. Psychometric functions with sinusoidal signals," J. Acoust. Soc. Am. **50,** 1151-61.

Viemeister, N. F. (**1988**). "Psychophysical aspects of auditory intensity coding" in *Auditory function: Neurobiological bases of hearing*, Vol. edited by G. M. Edelman, W. E. Gall and W. M. Cowan.(John Wiley & Sons, NYC)

Yost, W. A. (**1972**). "Tone-on-tone masking for three binaural listening conditions," J. Acoust. Soc. Am. **52,** 1234-7.

Yost, W. A., Penner, M. J., and Feth, L. L. (**1972**). "Signal detection as a function of contralateral sinusoid-to-noise ratio," J. Acoust. Soc. Am. **51,** 1966-70.

Zurek, P. M. (**1979**). "Measurements of binaural echo suppression," J. Acoust. Soc. Am. **66,** 1750-7.

Zwislocki, J. J. (**1972**). "A theory of central auditory masking and its partial validation," J. Acoust. Soc. Am. **52,** 644-59.

Tables and Figures

	Frequency (Hz)	Duration (ms)	Phase (radians)	Level (dB SPL)
No-Distractor	-	-	-	-
Fixed	600	300	0	50
Roving-Phase	600	300	$Uniform\left(-\frac{\pi}{2},\frac{\pi}{2}\right)$	50
Roving-Level	600	300	0	$Uniform(50,80)$
Double-Rove	600	300	$Uniform\left(-\frac{\pi}{2},\frac{\pi}{2}\right)$	$Uniform(50,80)$

Table 1. Distractor properties in the five conditions. In all conditions, the target has a frequency of 600 Hz, duration of 300 ms, a phase of zero, and a reference level of 50 dB SPL. The distractor was presented simultaneously but contra-aurally to the target. Roving of the level and phase of the distractor was done on an interval-by-interval basis, with values chosen from uniform distributions.

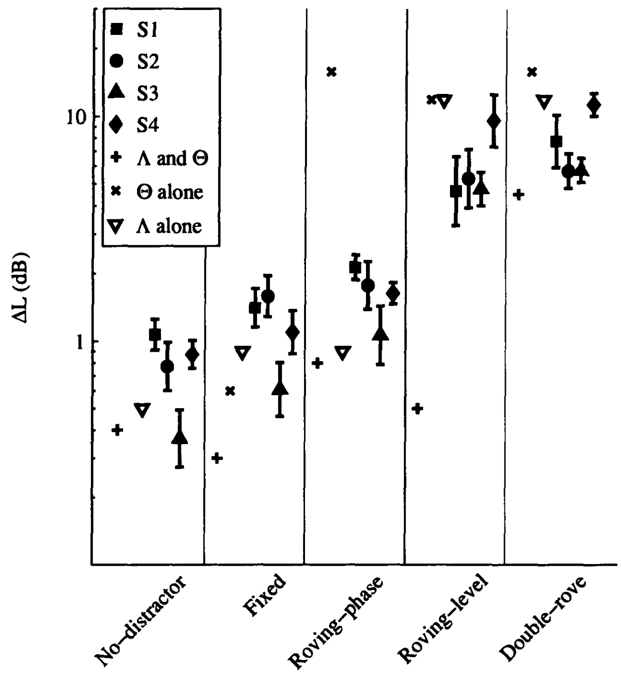

Figure 1. Mean thresholds for the four subjects under the five different conditions. Error bars are two times the standard error of the mean. Thresholds of the ideal observer of Λ and Θ, both alone and together, are also shown. In the no-distractor condition, the ideal observer of Θ never obtains threshold performance.

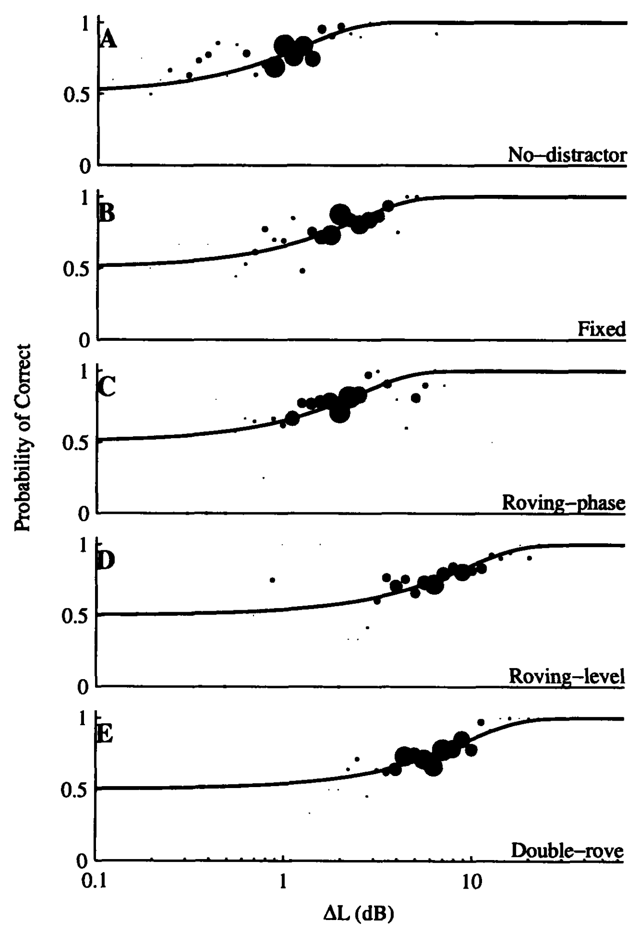

Figure 2. Examples of the dependence of the probability of a correct response as a function of the target increment for subject S2 for the five different conditions. Panels A-E correspond to the no-distractor, fixed, roving-phase, roving-level, and double-rove conditions, respectively. The data has been binned according to ΔL, and the size of the symbol is proportional to the number of trials that occurred within the bin. The best fitting psychometric functions, based on unbinned data, are also shown for each condition.

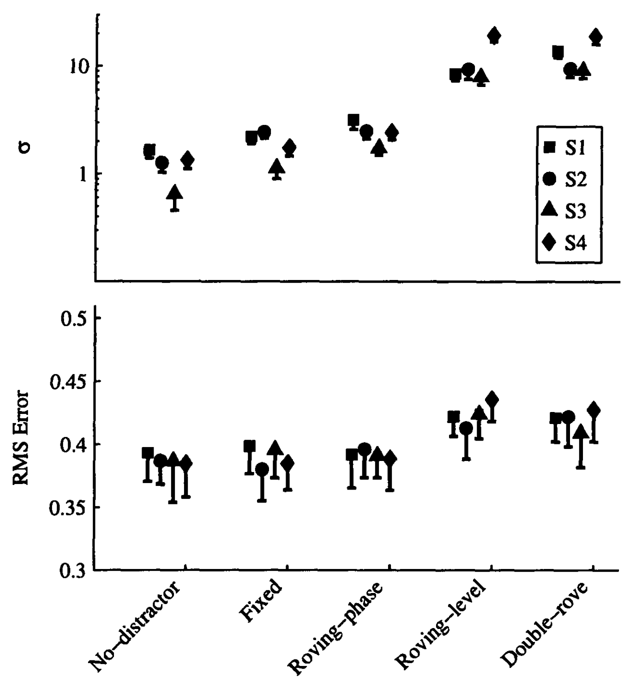

Figure 3. Fit parameter (top panel) and the root-mean-square (RMS) error (bottom panel) for the psychometric functions that were fitted to the data of the four subjects under the five different conditions. Error bars are the 95 percent confidence intervals derived by re-sampling the data. When the error bars are absent, the confidence interval is on the order of the size of the symbol.

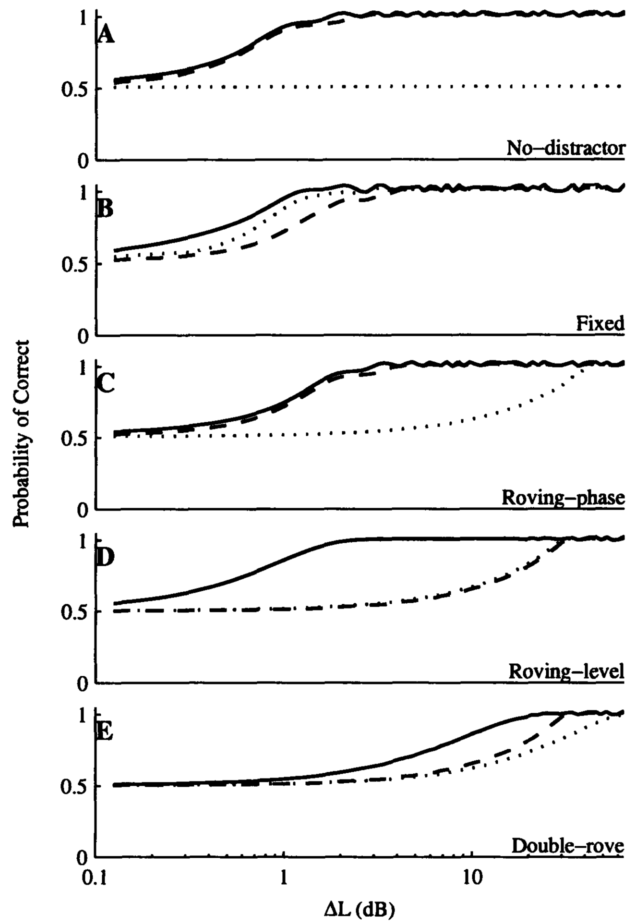

Figure 4. Predicted psychometric functions for the five conditions for the ideal observer of both Λ and Θ (solid), Λ alone (dashed) and Θ alone (dotted). Panels A-E correspond to the no-distractor, fixed, roving-phase, roving-level, and double-rove conditions, respectively.

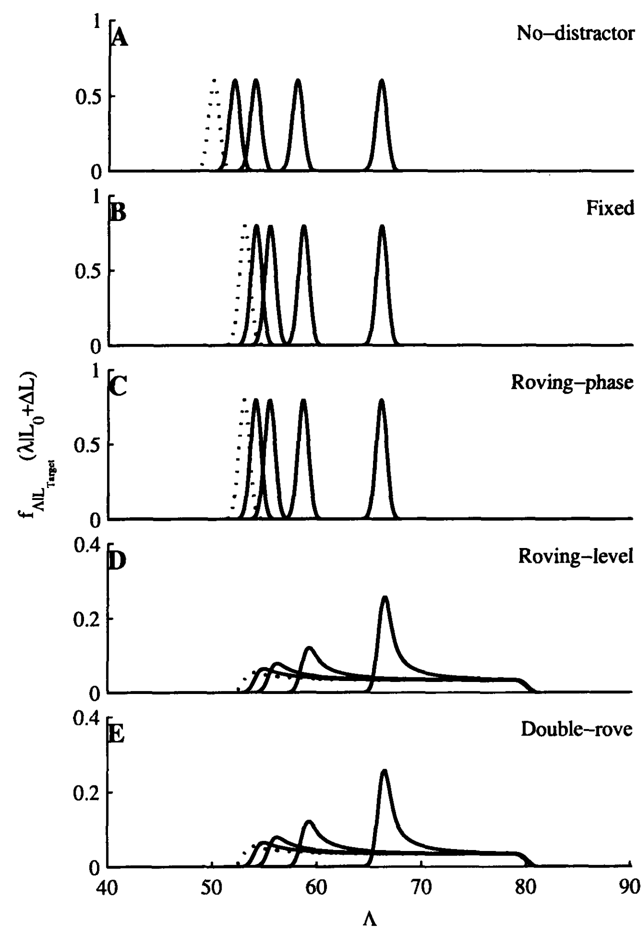

Figure 5. Marginal probability distributions of Λ for the five different conditions. Panels A-E correspond to the no-distractor, fixed, roving-phase, roving-level, and double-rove conditions, respectively. Each panel plots $f_{\Lambda|L_{Target}}\left(\Lambda|L_0+\Delta L\right)$ for five different values of ΔL (0, 2, 4, 8, and 16 dB). The marginal probability distribution corresponding to a ΔL of 0 dB (the reference in the experiment) is shown as a dotted line in each panel. One should note the change in scale between panels A-C and panels D and E.

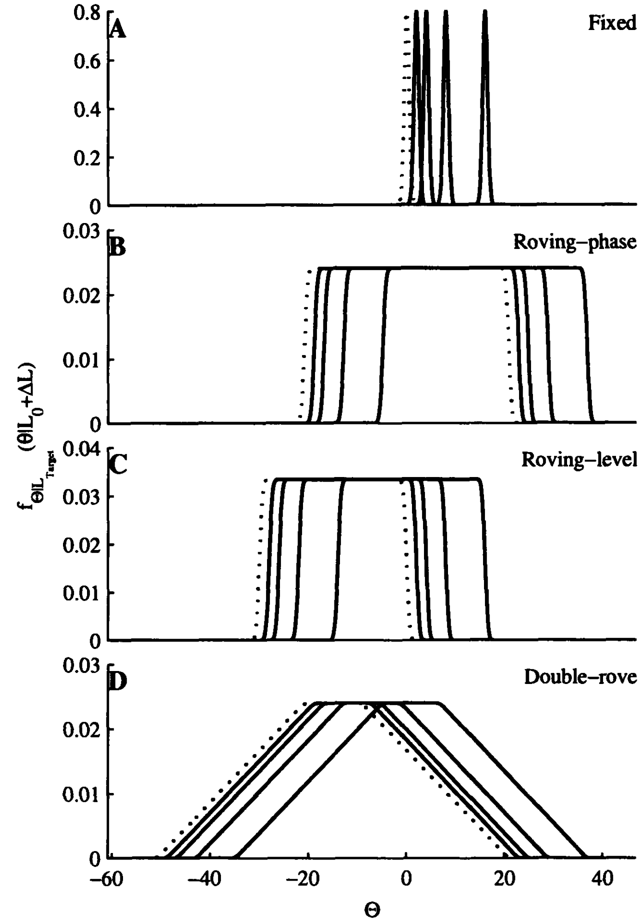

Figure 6. Marginal probability distributions of Θ for the five different conditions. Panels A-D correspond to the no-distractor, fixed, roving-phase, roving-level, and double-rove conditions, respectively. Each panel plots $f_{\Theta|L_{Target}}\left(\Theta|L_0+\Delta L\right)$ for five different values of ΔL (0, 2, 4, 8, and 16 dB). The marginal probability distribution corresponding to a ΔL of 0 dB (the reference in the experiment) is shown as a dotted line in each panel. In the no-distractor condition Θ is infinite.

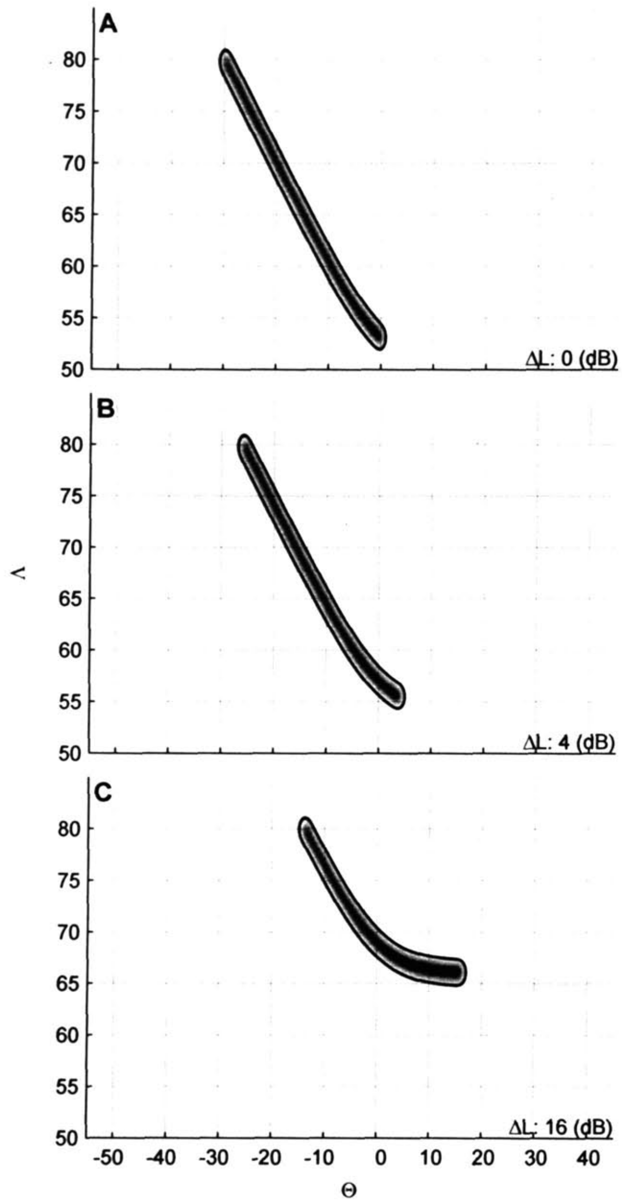

Figure 7. Joint probability density functions of Λ and Θ for the roving-level condition with three different values of ΔL. Panels A-C, respectively, correspond to values of ΔL of 0, 4, and 16 dB. Areas of high probability are dark and areas of low probability are light. In addition to the color coding, contours of equal probability are also shown.

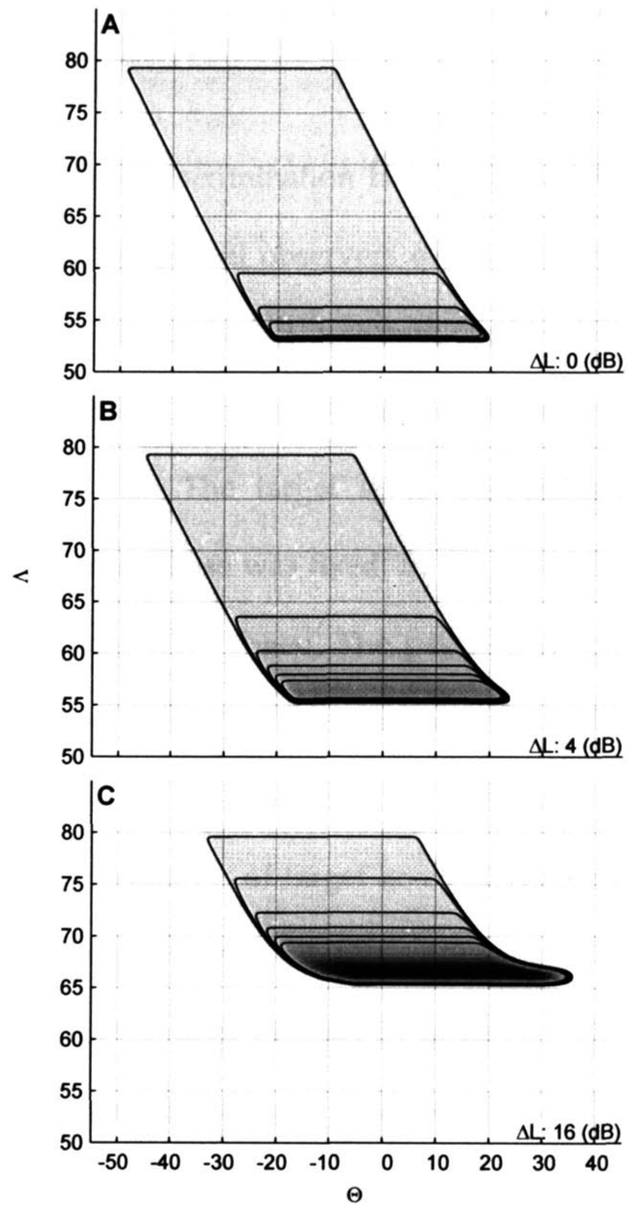

Figure 8. Joint probability density functions of Λ and Θ for the double-rove condition with three different values of ΔL. Panels A-C, respectively, correspond to values of ΔL of 0, 4, and 16 dB. Areas of high probability are dark and areas of low probability are light. Note that the color scale used in this figure is different than in Fig. 7. In addition to the color coding, contours of equal probability are also shown.

CHAPTER III. ONE INTERVAL LEVEL DISCRIMINATION UNDER DICHOTIC CONDITIONS

Abstract

A one-interval level-discrimination task was used to test three binaural models (corresponding to optimal observers of loudness alone, position alone, and loudness and position together). The stimuli used in the psychophysical task consisted of a 600-Hz target tone presented at the left ear and a 600-Hz distractor presented at the right ear. The target level was either un-incremented or incremented and the target phase was fixed. In the most complex condition, the distractor level and phase were roved. The psychophysical results show little inter-subject variability of the average percent correct and bias, but considerable inter-subject variability in the conditional probabilities of responding *"Incremented"* given combinations of target level, distractor level, and distractor phase. These results suggest that subjects can adopt substantially different strategies and still obtain the same probability of correct. No single model could predict the average percent correct and bias for all conditions. Further, all the models considered failed to predict the conditional probabilities as a function of the distractor level and phase.

I. Introduction

The monaural level discrimination experiment presented in this chapter has many similarities to the experiment in Chapter II. Both experiments involved discrimination of the level of a target at one ear in the presence of a distractor presented simultaneously, but contra-aurally to the target. One notable exception between the experiments is the paradigm. In the previous experiment, the paradigm was multi-interval and adaptive; in this experiment, a one-interval, two-alternative-forced-choice (1I, 2AFC) constant-increment paradigm is used. The 1I, 2AFC constant-increment paradigm allows for analyses which are not possible with the multi-interval, adaptive paradigm. In particular, the conditional probability of a given response (either *"Un-Incremented"* or *"Incremented"*) is estimated as a function of the distractor attributes (e.g., the distractor level and phase). These conditional probabilities are then used to compare performance across subjects as well as to test several models.

In the psychophysical portion of the work described here, subjects are asked to judge the level of a target tone at one ear in the presence of a distractor tone with the same frequency at the other ear. The levels and phases of the target and distractor were manipulated independently. The dominant perception of the stimulus is a single, compact, auditory image with a salient loudness and position. It is hypothesized that subjects may judge the level of the target tone in one of three ways. The subjects may attend to (1) the perceived loudness, (2) the perceived lateral position, or (3) a combination of the perceived loudness and lateral position. The modeling includes predictions of three models which correspond to the three hypothesized listening strategies. Based on a growing body of evidence we assume that subjects are unable to attend to a single ear.

In some early studies, an inability to attend to a single ear was reported by Rowland and Tobias (1967) [monaural level discrimination] and Zwislocki (1972) [central masking]. Work by Taylor and colleagues (Taylor and Clarke 1971; Taylor et al. 1971a; Taylor et al. 1971b) and Yost et al. (1972) clearly demonstrated that, in masked detection, subjects are not able to attend to a single ear. Studies of the precedence effect also have reported contra-aural interference; Zurek (1979) reported that the detection of an acoustic reflection was better under diotic conditions than under dichotic conditions. There is also a growing body of literature on informational masking in which listening with one ear is better than listening with two ears (Brungart and Simpson 2002; Kidd et al. 2003). More recently Bernstein (2004) demonstrated contra-aural interference in a level discrimination task. Finally, in the Chapter II, it was shown that, with identical stimuli in a similar paradigm, subjects are unable to attend to a single ear.

Given that subjects are unable to attend to a single ear in a level discrimination task when tones of the same frequency are presented simultaneously, it is hypothesized that subjects are basing their decisions on binaural percepts. It is clear from studies of binaural phenomenon [refer to Durlach and Colburn (1978) for a review of binaural psychophysics] that normal hearing listeners are sensitive to changes in interaural differences in level (ILD) and time delay (ITD). In general, it is believed that discrimination of changes to the ILD or ITD is based on changes in the perceived lateral position. This belief stems from the strong predictive power of models based on correlates of the lateral position (e.g., Hafter and Carrier 1972; Stern and Colburn 1978). Since changing the level at one ear changes the perceived lateral position, subjects could discriminate changes in level at one ear by attending to the position. Changing the level at one ear may also change the loudness. The extent to which the loudness varies with the change in level at one ear depends strongly on the interaural difference in level. Under some circumstances, attending to the loudness is a reliable way to discriminate changes in the level at one ear.

While subjects could choose to attend exclusively to the loudness or exclusively to the position, it is possible that some subjects would attend to a combination of the loudness and position. Studies of divided attention suggest that there is a limited cost of dividing attention across two independent dimensions of the same modality (Bonnel and Hafter 1998). But, these studies have not examined dependent dimensions or conditions in which there is considerable joint information.

Although the modeling focuses on the position and loudness, there are many other perceptual attributes of tonal stimuli. For example, a tonal stimulus has a strong pitch percept, but the pitch of a tone is generally not considered to be strongly level dependent. Two less salient aspects of the tonal stimulus are the "time-image" (Hafter and Jeffress 1968a; Hafter and Carrier 1972) and the spatial width (Smith 1973; Ruotolo et al. 1979). The role of these less salient dimensions in the discrimination of the level at one ear is unknown, and only the two salient and obvious dimensions are incorporated into the models analyzed here. The increased complexity required to include the less salient dimensions in a model is substantial and it is the intent to present the limitations of models based on these obvious dimensions.

II. Methods

Subjects were asked to discriminate the level of a target at the left ear, in the presence of a distractor presented simultaneously, but contra-aurally, to the target. Performance was measured under three experimental conditions which were differentiated by the properties of the distractor as listed in Table 1. Across conditions, the properties of the target were identical and the properties of the distractor varied. The experimental paradigm in all three conditions was one-interval, two-alternative-forced-choice (1I, 2AFC) with a constant-increment that was held fixed for each subject.

A. Stimuli and Paradigm

Each stimulus consisted of a target and a distractor. The target and distractor were both 600-Hz tones with rise and fall times of 25 ms and a total duration of 300 ms. The left ear always received the target, while the right ear received the distractor. The target and distractor were gated on and off simultaneously. The phase of the target was fixed at zero and the level of the target was either un-incremented or incremented. The level of the un-incremented target was always 50 dB SPL. Performance was measured with fixed, roving-level, and double-rove (i.e., roving-phase and roving-level) distractors. In the fixed condition, the level of the distractor was fixed at

50 dB SPL (the same level as the un-incremented target). In the other two conditions (roving-level and double-rove) the decibel level of the distractor was chosen randomly from a uniform distribution between 50 and 80 dB SPL. In the fixed and roving-level conditions, the phase of the distractor was fixed at zero. In the double-rove condition, the phase of the distractor was also roved, uniformly between $-\pi/2$ and $+\pi/2$.

Noticeably absent from the conditions tested is a roving-phase condition in which the level of the distractor is fixed and the phase of the distractor is roved. In Chapter II fixed, roving-level, and double-rove as well as target-only and roving-phase conditions were tested with a multi-interval adaptive paradigm. It was demonstrated that in the roving-phase condition subjects attend to the loudness and perform essentially the same as they perform in the target-only and fixed conditions. Additionally the predictions of a model based on the perceived loudness and position are also essentially the same for the target-only, fixed, and roving-phase conditions. Therefore only the fixed, roving-level, and double-rove conditions were tested.

In all three conditions, the size of the target level increment ΔL was fixed and the size of ΔL was adjusted individually for each subject. Specifically, for each subject, ΔL was adjusted such that the subject obtained a probability of correct of approximately 0.75 in the double-rove condition. For subjects S1 and S2, ΔL was fixed at 8 dB; the target was presented at either 50 or 58 dB SPL with equal a priori probability. For subject S4, ΔL was fixed at 14 dB (the target was either 50 or 64 dB SPL). These individualized values of ΔL were used for all trials of all runs of all three conditions. For all subjects, data were collected for the double-rove condition first, then for the roving-level condition second, and finally for the fixed condition last. Data were collected in sessions which lasted between one and a half and two hours with frequent rest breaks and approximately 2000 trials were collected per session. The trials were self paced and subject received correct-answer feedback after every trial. Subjects completed two or three sessions per week.

B. Subjects

Three of the four subjects (S1, S2, and S4) who participated in the psychophysical experiment of Chapter II also completed this study. Subject S3 from that study did not obtain satisfactory performance in the initial stages of this experiment and therefore data collection for this subject was stopped. More specifically, data collection was stopped for subject S3 since in the double-rove condition with a ΔL of 20 dB subject S3 reported a feeling of "total confusion" and was performing at near chance levels. Note, however, that subject S3 was among the best performers with the multi-interval, adaptive paradigm of the Chapter II. An additional subject was recruited and given a moderate amount of training (approximately 10 hours) and was also not able to perform satisfactorily with a 1I, 2AFC constant-increment paradigm[1] and therefore additional data were not collected for this subject either. Only the results from the three subjects who completed the study are reported. Subject S1 is an author. Subjects, with the exception of the author, received an hourly wage for their participation. All subjects had pure tone thresholds below 20 dB HL at frequencies of 250, 500, 1000, 2000, 4000, and 8000 Hz in both ears. Subject ages were between 19 and 31 years old.

Subjects S1 and S2 completed 30,000 trials in the double-rove condition, then they completed 12,000 trials in the roving-level condition and finally 2,000 trials in the fixed condition. Subject S4 completed 24,000 trials in the double-rove condition, 6,000 trials in the roving-level condition, and 2,000 trials in the fixed condition. Fewer trials were conducted with subject S4 due to the limited availability of the subject for participation in the experiment. The number of trials varied across the conditions since sufficient data were needed across the entire range of distractor properties to estimate the conditional probabilities of each response. In the double-rove condition, the distractor varied in both level and phase (two dimensions). In the roving-level condition, the distractor only varied in one dimension. In the fixed condition, the distractor did not vary.

C. Apparatus

The experiments were conducted in a double walled sound treated booth. Stimulus generation was implemented with Tucker-Davis-Technology System II hardware (AP2, PD1, PA4, and HB6) with a sampling rate of 50 kHz. The stimuli were presented over Sennheiser HD265 headphones. The subjects responded through a graphical interface displayed on a computer monitor. The graphical response buttons were labeled "*A*" and "*B*". When the target level was un-incremented, the correct response was *A*. Conversely, when the target level was incremented the correct response was *B*. The subjects were given instructions (and correct-answer feedback) intended to aid in the use of these un-informative labels.

The instructions were based on the previous experience of the subjects with a multi-interval version of the experiment in which the target level was un-incremented in all but one of the intervals (cf. Chapter II). The instructions referred to "*same*" and "*different*" stimuli. Based on the previous paradigm, a "*same*" stimulus was one in which the target level was un-incremented and a "*different*" stimulus was one in which the target level was incremented. The terms "*Un-Incremented*" and "*Incremented*" were not used in the instructions given to the subjects. Rather, the subjects were only instructed that *A* corresponded to a "*same*" stimulus and that *B* corresponded to a "*different*" stimulus. The subjects who completed this experiment reported no difficulties assigning these labels and felt comfortable with their prior experience of the "*same*" and "*different*" stimuli.[2] For ease of discussion, throughout this chapter, the responses *A* and *B*, respectively, will be referred to by the more informative labels "*Un-Incremented*" and "*Incremented*". These more informative labels were not given to the subjects due to concerns about biasing the subjects towards using a particular response strategy.

D. Data Analysis

The responses are analyzed by calculating the probability of the subject responding *Incremented,* conditioned on the binary target level (the target level was either un-incremented or incremented). These conditional probabilities are referred to using the nomenclature of signal detection theory where the signal to be detected is the target with a level that is incremented. In particular, the probability of the subject responding *Incremented* given that the target level is un-incremented is P_F (the probability of a false alarm) and the probability of the subject responding *Incremented* given that the target level is incremented is P_D (the probability of a detection).

In the fixed condition, P_F and P_D are calculated for each subject by averaging across all trials. In the roving-level condition, P_F and P_D are calculated as a function of the distractor level every 0.25 dB; distractors which have levels within ±1 dB of the nominal level are included in the calculations. Overlap between calculations both increases the number of trials included in each average and smoothes the data. For subjects S1 and S2 approximately 400 trials are used in each calculation and for subject S4 approximately 200 trials are used.

In the double-rove condition, P_F and P_D are calculated as a function of both distractor level and phase. These probabilities are calculated for 20,000 cases (every 0.25 dB in the level dimension and every 5 μs in the phase/time dimension). Distractors which have levels and phases/times within ±1 dB and

±20 μs of the specified values are included in each calculation. In the double-rove condition, for subjects S1 and S2, each calculation is based on approximately 50 trials; calculations for subject S4 are based on a slightly smaller number of trials (approximately 40 in the double-rove condition).

In addition to P_F and P_D, the across distractor level and phase average probability of a correct response $\overline{P_C}$ and bias towards responding incremented $\overline{\beta}$ are also calculated. These values depend on P_F and P_D as well as the probability of the target level being incremented. Although the *a priori* probability of the target level being incremented is equal to 0.5 (the target level is un-incremented and incremented with equal probability), the actual fraction of trials incremented is different and this actual fraction is used to estimate $\overline{P_C}$ and $\overline{\beta}$. Note that an observer who performs perfectly has a $\overline{P_C}$ of unity, an observer who performs at chance has a $\overline{P_C}$ of 0.5, and an observer who is always wrong has a $\overline{P_C}$ of zero. The expected bias towards responding *Incremented* $\overline{\beta}$ is given by $\overline{\beta} = P_{Incremented}\overline{P_D} + (1 - P_{Incremented})\overline{P_F} - P_{Incremented}$. An observer who responds *Un-Incremented* and *Incremented* with the same frequency has a $\overline{\beta}$ of zero. While an observer that always responds *Un-Incremented* has a $\overline{\beta}$ of -0.5 and one that always responds *Incremented* has a $\overline{\beta}$ of 0.5. Note that an ideal (maximum likelihood) observer can have a non-zero $\overline{\beta}$ even when the probability of the target level being incremented is 0.5 and $\overline{P_C}$ is substantially greater than chance levels.[3]

III. Modeling

The modeling considers the responses of an ideal (maximum likelihood) observer of a two-dimensional decision space as well as ideal observers of each dimension individually. The two dimensions loosely correspond to the perceived loudness and lateral position and are denoted Λ and Θ, respectively. The two dimensions are functions of the level (in decibels) of the left ear (L_{Left}), the level of the right ear (L_{Right}), the interaural difference in time delay (T), and two hypothesized internal noises (N_Λ and N_Θ). They are defined as

$$\Lambda = 10\log_{10}\left(10^{\frac{L_{Left}}{10}} + 10^{\frac{L_{Right}}{10}}\right) + N_\Lambda \qquad (1)$$

and

$$\Theta = L_{Left} - L_{Right} + k\mathrm{T} + N_\Theta, \qquad (2)$$

where k corresponds to the intensity-time trading ratio. The additive internal noises are Gaussian random variables which are statistically independent of each other and across trials. The zero-mean internal noises N_Λ and N_Θ are characterized by their respective standard deviations σ_Λ and σ_Θ. The free parameters of the model (σ_Λ, σ_Θ, and k) are fixed for all conditions: both σ_Λ and σ_Θ are fixed at 0.5 dB and k is fixed at 0.05 dB per μs.[4] The Θ dimension is a noisy version of the weighted the sum of the interaural differences in level and time delay. The Λ dimension is a noisy representation of the sum of the intensities (units of watts per area) at the two ears expressed in decibels.

The model predictions are based on calculations of P_F and P_D. Specifically, P_F and P_D are calculated as a function of the distractor level and phase for a 1I, 2AFC constant-increment paradigm. A mixture of analytic and numerical techniques is used to approximate P_F and P_D. A detailed derivation is presented in the Appendix. The derivation defines P_F and P_D in terms of a maximum likelihood indicator function. The indicator functions divides the $\{\Lambda, \Theta\}$ space into two regions; one in which the target level is more likely to be un-incremented and the other in which the target level is more likely to be incremented. In Chapter II an approximation of the maximum likelihood indicator function for a four-interval paradigm was derived. The derivation with a one-interval paradigm is similar and therefore is not provided. Note that the approximations of the indicator function, P_F, and P_D require no Monte-Carlo type simulations. The approximations arise from the use of numerical techniques to calculate the integrals of continuous functions.

IV. Psychophysical Results

The top panels of Fig. 1 show the average probability of correct $\overline{P_C}$ for each of the three subjects in each of the three conditions. Recall that the target level increment ΔL is constant across the three conditions, but varies across the subjects. For subjects S1 and S2, ΔL is 8 dB; for subject S4, ΔL is 14 dB. Due to the large number of trials, a formal statistical analysis is not presented since many small differences are statistically reliable. For example, the 95 percent confidence intervals for $\overline{P_C}$ based on binomial variability are approximately ±0.01. More conservative estimates of the reliability based on day-to-day and block-to-block repeatability also reveal a high degree (although less than with the binomial estimation) of reliability.

All three subjects perform best in the fixed condition. On average, $\overline{P_C}$ in the fixed condition was 0.14 higher than $\overline{P_C}$ in the roving-level condition and 0.23 higher than $\overline{P_C}$ in the double-rove condition. Although there is inter-subject variability (e.g., S4 performed much worse than S1 and S2 in the fixed condition and S1 was better in the roving-level condition), overall $\overline{P_C}$ amongst the subjects was similar both in magnitude and the trends across conditions. The small across-subject differences of $\overline{P_C}$ in the double-rove condition are not surprising since ΔL was adjust for each subject to yield a $\overline{P_C}$ of approximately 0.75 in the double-rove condition. Subject S4 required a 6 dB increase in ΔL to obtain approximately the same $\overline{P_C}$ as the other subjects in all the conditions.

The bottom panel of Fig. 1 shows the bias $\overline{\beta}$ for the three subjects in the three conditions. As was the case with $\overline{P_C}$, no formal statistical analysis on $\overline{\beta}$ is presented due to the large number of trials. Essentially all visible differences of

$\overline{\beta}$ are statistically significant. In all three conditions, all the subjects show little bias with values of $\overline{\beta}$ near zero (note the expanded scale). The mean value of $\overline{\beta}$ was -0.03; the probability of the subjects responding *Un-Incremented* (response button "*A*") was 0.03 greater than the probability of responding *Incremented* (response button "*B*"). The largest magnitude of $\overline{\beta}$ was 0.07 (S1 in the double-rove condition), in which a response of *Un-Incremented* was favored. In summary, a strong bias was not exhibited by any of the subjects in any of the conditions.

Although the across-subject differences in $\overline{P_C}$ and $\overline{\beta}$ are relatively small, there are large across-subject differences when the effects of the distractor level and phases are separated. Across-subject variability can be further assessed by considering performance as a function of both (1) the distractor level in the roving-level condition and (2) the distractor level and phase in the double-rove condition. (This is possible because such a large number of trials were collected and the paradigm was 1I, 2AFC constant-increment.) Performance as a function of the distractor properties is reported in terms of P_F and P_D (conditional probabilities of responding *Incremented* given that the target level was un-incremented and incremented, respectively). The results of the roving-level condition are reported first, followed by the results of the double-rove condition.

For the roving-level condition, P_F and P_D as a function of the distractor level are shown as thick lines (solid and dashed, respectively) in Fig. 2. The dependency of P_F and P_D on the distractor level (e.g., increasing, decreasing or flat) varies from subject to subject even though the average $\overline{P_C}$ and $\overline{\beta}$ are similar (cf. Fig. 1). Additionally, the range which the P_F and P_D functions span varies from subject to subject. The general trends of the results for the three subjects are as follows. For subject S1, P_F and P_D are nearly constant at 0.17 and 0.88, respectively. For subject S2, P_F increases with distractor level from near zero with the lowest distractor levels to near 0.5 with the highest distractor levels and P_D is nearly constant at 0.78. For subject S4, as the distractor level increases, both P_F and P_D decrease. For this subject, when the distractor level is low, P_F is approximately 0.5 and P_D is nearly unity and at the highest distractor levels, P_F is near zero and P_D is about 0.25. The differences in the dependencies of P_F and P_D on the distractor level suggests that the three subjects have chosen three different strategies; these different strategies, however, all yield similar $\overline{P_C}$ and $\overline{\beta}$.

For the double-rove condition, P_F and P_D are two-dimensional functions which depend on the distractor level and phase. Surface plot representations of these two-dimensional functions are shown in Fig. 3. Each panel of Fig. 3 corresponds to either P_F (left column) or P_D (right column) for a single subject. Note that a positive phase distractor leads the target (i.e., T in Eq. 2 is negative for positive phase distractors). For each subject, the dependencies of P_F and P_D on the distractor level and phase are complex. Furthermore, similar to the roving-level condition, each subject performs differently, even though all subjects obtain similar values of $\overline{P_C}$ and $\overline{\beta}$ (cf. Fig. 1).

Although the response patterns are complex, some generalities can be made about P_F and P_D in the double-rove condition. First, as expected, for a given distractor level and phase, P_D is generally greater than P_F. Further, P_D is less dependent than P_F on the distractor level and phase. Finally, areas of low and high P_F appear to be separated by a simple boundary that varies across the subjects. For subject S1, P_F depends most strongly on the distractor phase; P_F is near unity for distractors with negative phases (distractor lags the target) and near zero for distractors with positive phases (distractor leads the target). Subjectively, this dependence is reasonable; distractors with negative phases should be perceived more towards the ear at which the target is presented which is similar to the effects of incrementing the target level. For subject S2, P_F depends predominately on the distractor level. For distractors with a level lower than 65 dB SPL, P_F is near zero and for higher level distractors P_F increases to nearly 0.8. For subject S4, the boundary depends on both the distractor level and phase. For distractors with levels in excess of approximately 75 dB SPL or phases greater than $\pi/4$, P_F is near zero while for lower level distractors with less positive phases, P_F increases to 0.7. This dependence is also subjectively reasonable; a distractor with a positive phase and low level should have a perceived position and loudness more similar to that of the un-incremented target.

In the double-rove condition the distractor phase is different on every trial. On some trials the distractor phase is near zero making the distractor similar to the distractor in the roving-level condition. The probabilities P_F and P_D, for trials on which the distractor phase is near zero in the double-rove condition, are shown in Fig. 2 as thin lines along with the corresponding results from the roving-level condition (thick lines). For each distractor level, the number of trials in the double-rove condition for which the distractor phase is near zero was small (40 on average, but as few as 10). Therefore the statistical reliability of the data from the double-rove condition is much less than for the data from the roving-level condition.

Although the dependence of P_F and P_D on the distractor level varies across the subjects, for each subject the dependence of P_F and P_D on the distractor level is the same in the roving-level and double-rove conditions when the distractor phase is near zero. Statistically, the correlation coefficient between the P_F and P_D functions from the roving-level condition and the P_F and P_D functions from the double-rove condition when the distractor phase is near zero are 0.94, 0.96 and 0.88 for subjects S1, S2, and S4, respectively. Visually, as suggested by the correlation coefficients, the gross trends of P_F and P_D are captured, while the fine details are often different. These high correlation coefficients suggest that subjects are using the same decision rule in both the double-rove and roving-level conditions.

V. Model Results

The ability of an ideal observer to discriminate changes in the target level is limited by either (1) internal noise alone (i.e., a combination of N_{Λ} and N_{Θ}) or (2) a combination of internal noise and distractor variability. Whether an ideal observer is limited by internal noise alone or by a combination of internal noise and distractor variability depends on both (1) the number of dimensions observed (one for the Λ-alone and Θ-alone observers and two for the joint, Λ and Θ together, observer) and (2) the number of dimensions in which the distractor properties vary (none in the fixed condition, one in the roving-level condition and two in the double-rove condition). Specifically, when the number of dimensions on which decisions are based is greater than the number of dimensions in which the distractor properties vary, performance is limited by the internal noise. When the number of dimensions on which decisions are based is less than or equal to the number of dimensions in which the distractor properties vary, performance is limited by a combination of internal noise and distractor variability.

In the model, the internal noise is much smaller than the distractor variability. The standard deviations of the internal noises on the Λ and Θ dimensions are 0.5 dB. In the double-rove condition, the distractor level varies uniformly over a 30 dB range (standard deviation of 8.7 dB) and the intensity-time traded distractor phase varies over a range equivalent to 40 dB (standard deviation of 11.5 dB). Further, the size of ΔL (either 8 or 14 dB) is large compared to the internal noise and of the same order of magnitude as the distractor variability. Therefore, when performance is limited by the internal noise, the ideal observer should perform nearly perfectly (P_F of zero, P_D of unity, $\overline{P_C}$ of unity, and $\overline{\beta}$ of zero). When performance is limited by a combination of internal noise and distractor variability, optimum performance decreases substantially. In these cases of decreased performance, the influence of the internal noise is small compared to the influence of the distractor variability. The exact performance of the model depends on the dimensions observed (Λ alone, Θ alone, or Λ and Θ together), ΔL, and the experimental condition (fixed, roving-level, or double-rove).

The model results are presented in two parts. A graphical analysis of the model is presented first, followed by predictions for the psychophysical results. The graphical analysis focuses on combinations of observers and conditions in which the performance is limited by a combination of internal noise and distractor variability; the graphical analysis does not consider the combinations of observers and conditions in which the performance is limited by only internal noise (i.e., when the distractor variability has no effect). Further, the graphical analysis is limited to the values of L_{Target} used in the psychophysical experiment. Specifically, three values of L_{Target} are considered: 50 dB (the reference level for all three subjects), 58 dB (the incremented level for S1 and S2), and 64 dB (the incremented level for S4). The psychophysical predictions include predictions of P_F and P_D as a function of the distractor level (roving-level and double-rove conditions) and phase (double-rove condition) for values of ΔL of 8 and 14 dB. Additionally, predictions of $\overline{P_C}$ and $\overline{\beta}$ (as well as $\overline{P_F}$ and $\overline{P_D}$) are presented for all three conditions and both values of ΔL.

A. Graphical Analysis of the Model

The graphical analysis is intended to give insight without requiring a formal mathematical derivation of the model. (The formal derivation of the model is presented in the Appendix.) The graphical analysis, however, does require some mathematical notation to be introduced. First, note that when the level of the target L_{Target} is un-incremented, L_{Target} is equal to the reference level L_0 (in decibels) and when the level of the target is incremented, L_{Target} is equal to the sum of the reference level and ΔL (in decibels). On each trial, the modeled observer (Λ-alone observer, Θ-alone observer, Λ-and-Θ-together observer) must decide whether the target level was incremented or not based on an observation of Λ, Θ or, both. The observation of Λ is denoted λ and the observation of Θ is denoted θ.

Decisions are based on the conditional probability of the observation (Λ alone, Θ alone, or Λ and Θ together) occurring, given L_{Target}. These conditional probabilities are denoted $f_{\Lambda|L_{Target}}$, $f_{\Theta|L_{Target}}$, and $f_{\Lambda,\Theta|L_{Target}}$, respectively.[5] The observers are defined in terms of a log-likelihood ratio of these density functions when L_{Target} is equal to the un-incremented (L_0) or the incremented levels ($L_0 + \Delta L$). Specifically, the log-likelihood ratio for the Λ-alone, Θ-alone, and Λ-and-Θ-together observers are denoted η_Λ, η_Θ and $\eta_{\Lambda,\Theta}$, respectively and are defined as

$$\eta_\Lambda(\lambda, \Delta L) = 10\log_{10}\left(\frac{f_{\Lambda|L_{Target}}\left(\lambda \middle| L_{Target} = L_0\right)}{f_{\Lambda|L_{Target}}\left(\lambda \middle| L_{Target} = L_0 + \Delta L\right)}\right),$$

$$\eta_\Theta(\theta, \Delta L) = 10\log_{10}\left(\frac{f_{\Theta|L_{Target}}\left(\theta \middle| L_{Target} = L_0\right)}{f_{\Theta|L_{Target}}\left(\theta \middle| L_{Target} = L_0 + \Delta L\right)}\right),$$

and

$$\eta_{\Lambda,\Theta}(\lambda, \theta, \Delta L) = 10\log_{10}\left(\frac{f_{\Lambda,\Theta|L_{Target}}\left(\lambda, \theta \middle| L_{Target} = L_0\right)}{f_{\Lambda,\Theta|L_{Target}}\left(\lambda, \theta \middle| L_{Target} = L_0 + \Delta L\right)}\right). \tag{3}$$

When the log-likelihood ratio is equal to zero, the conditional probability of the observed values given that the target level is un-incremented is equal to the conditional probability of the observed values given that the target level is incremented. Positive values for the log-likelihood ratio imply that the observed values are more likely to occur with a target level that is un-incremented while negative values imply that the target level is more likely to be incremented.

According to our assumptions, the maximum likelihood indicator functions defined by comparing η_{Λ}, η_{Θ} and $\eta_{\Lambda,\Theta}$ to zero are relatively simple. For example, the indicator function for the ideal observer of Λ alone is to respond *Incremented* when Λ is above a criterion value. For the observers of Λ alone and Θ alone, these indicator functions can be inferred from η_{Λ} and η_{Θ}, respectively. Therefore the graphical analysis begins with $f_{\Lambda|L_{Target}}$ and η_{Λ}, followed by an analysis of $f_{\Theta|L_{Target}}$ and η_{Θ}. For the observer of Λ and Θ together, the increased dimensionality of $f_{\Lambda,\Theta|L_{Target}}$ and $\eta_{\Lambda,\Theta}$ make a graphical analysis of limited use and therefore it is not presented. Although a graphical analysis of $f_{\Lambda,\Theta|L_{Target}}$ and $\eta_{\Lambda,\Theta}$ is not practical, a graphical analysis of the indicator function for the observer of Λ and Θ together is presented. In the graphical analysis, the three conditions (fixed, roving-level, and double-rove) are treated separately.

In the fixed condition, the performance of all three observers (Λ alone, Θ alone, and Λ and Θ together) is limited by only internal noise and therefore a graphical analysis is not presented. In the roving-level condition, the observer of Λ and Θ together is only limited by internal noise and a graphical analysis is not presented; the observers of Λ alone and Θ alone, however, are limited by a combination of internal noise and distractor variability, and therefore graphical analyses for these two observers in the roving-level condition are presented. Figure 4 graphically displays $f_{\Lambda|L_{Target}}$ (top) and η_{Λ} (bottom) in the roving-level condition, while the left column of Fig. 5 displays $f_{\Theta|L_{Target}}$ (top) and η_Θ (bottom) in the roving-level condition.

In the roving-level condition, the density function $f_{\Lambda|L_{Target}}$ (cf. Fig. 4) is heavily skewed towards smaller values of Λ. Further, the minimum value of Λ for which $f_{\Lambda|L_{Target}}$ is substantially greater than zero shifts systematically with ΔL. In contrast, the maximum value of Λ for which $f_{\Lambda|L_{Target}}$ is substantially greater than zero is nearly independent of ΔL. The properties of $f_{\Lambda|L_{Target}}$ are a direct consequence of the definition of Λ (cf. Eq. 1) which sums the target and distractor intensities. That is units of power (per area), not the decibel levels, are summed to be more consistent with measurements of binaural loudness. The lack of a change in the maximum value of Λ associated with changes in ΔL is expected since at high distractor levels the distractor dominates the loudness. The change in the minimum value of Λ is also expected since at low distractor levels the target influences Λ. For the un-incremented target level (ΔL of 0 dB) the low level distractors influence Λ to nearly the same degree as the high level distractors. For the incremented target (ΔL of 8 and 14 dB), the low level distractors do not influence Λ while the high level distractors do influence Λ. This variable influence of the distractor results in the skew of $f_{\Lambda|L_{Target}}$.

Given the dependence of $f_{\Lambda|L_{Target}}$ on Λ and ΔL, the dependence of η_Λ on Λ and ΔL is not surprising. For values of Λ less than some criterion (approximately 58 dB and 64 dB for values of ΔL of 8 and 14 dB, respectively), η_Λ is positive with a large magnitude. For values of Λ above this criterion, η_Λ behaves in a complex manner. Near the criterion η_Λ rapidly becomes negative. Although the dependence of the magnitude of η_Λ on Λ is complicated, for all values of Λ greater than the criterion, η_Λ is negative (i.e. the target level is more likely to be incremented). Comparing the top and bottom panels of Fig. 4 reveals that generally if the probability of Λ given an incremented target is substantially greater than zero, the value of η_Λ is negative. This suggests that the ideal observer of Λ in the roving-level condition is biased towards responding incremented.

The left column of Fig 5 displays $f_{\Theta|L_{Target}}$ (top) and η_Θ (bottom) in the roving-level condition. Note that For each L_{Target}, $f_{\Theta|L_{Target}}$ is nearly uniform over a 30 dB range (recall that in Eq. 2 Θ is defined as having units of decibels) and the limits of the range are shifted by ΔL. The distractor level is uniformly distributed, but the internal noise results in slight deviations. Since the internal noise has a standard deviation of 0.5 dB and the distractor level is roved over a 30 dB range, these deviations are small. For each ΔL the span is approximately 30 dB, the upper (right) limit is approximately equal to ΔL, and the lower (left) limit is approximately 30–ΔL. The dramatic difference between $f_{\Theta|L_{Target}}$ and $f_{\Lambda|L_{Target}}$ is that Θ is defined by summing the levels (in decibels) of the target and distractor and an intensity-time traded phase component (in the roving-level condition the phase component is always equal to zero).

The log-likelihood ratio η_Θ follows from $f_{\Theta|L_{Target}}$. Changes to ΔL horizontally shift the η_Θ function substantially. There is a wide range of values of Θ such that η_Λ is nearly zero. However, there is only a single value of Θ for which η_Θ is equal to zero. There is a single criterion value of Θ such that for values of Θ less than the criterion, η_Θ is positive; for values greater than the criterion, η_Θ is negative. Both $f_{\Theta|L_{Target}}$ and η_Θ are symmetric and therefore the ideal observer of Θ alone will be unbiased. The wide range of values of Θ such that η_Λ is nearly zero implies that numerous non-ideal indicator functions can lead to nearly the same probability of correct as achieved by the ideal indicator function, but these non-ideal indicator functions will have different biases.

In the double-rove condition the ideal observers of Λ alone, Θ alone, and Λ and Θ together are limited by a combination of internal noise and distractor variability. Since Λ is independent of the distractor phase, $f_{\Lambda|L_{Target}}$ and η_Λ in the double-rove condition is identical to $f_{\Lambda|L_{Target}}$ and η_Λ in the roving-level condition and the functions are not discussed again. The Θ dimension depends on both the distractor level and phase and therefore $f_{\Theta|L_{Target}}$ and η_Θ are different in the double-rove condition. The right column of Fig 5 displays $f_{\Theta|L_{Target}}$ (top) and η_Θ (bottom) in the double-rove condition. The $f_{\Theta|L_{Target}}$ function is nearly trapezoidal in the double-rove condition. Mathematically, $f_{\Theta|L_{Target}}$ in the double-rove condition is the convolution of $f_{\Theta|L_{Target}}$ in the roving-level condition with a rectangle corresponding to the variability of the intensity-time-traded distractor phase. The log-likelihood ratio η_Θ in the double-rove condition is visually different than η_Θ in the roving-level condition, however, two important properties of η_Θ are maintained: η_Θ is again symmetric and there is only one value of Θ for which η_Θ is equal to zero. Unlike in the roving-level condition, η_Θ in the double-rove condition does not have a wide range of values near zero; there is a more limited set of non-ideal indicator functions which can lead to nearly the same probability of correct.

The dimensionality of $f_{\Lambda,\Theta|L_{Target}}$ and $\eta_{\Lambda,\Theta}$ makes a graphical analysis of the functions of limited use. However, the maximum likelihood indicator function based on comparing $\eta_{\Lambda,\Theta}$ to zero results in a boundary which divides the $\{\Lambda,\Theta\}$ plane into regions of indicate-*Un-Incremented* and indicate-*Incremented* for all three conditions. A graphical analysis of this boundary is insightful. The top panel of Fig. 6 shows the decision boundary for the roving-level condition with values of ΔL of 8 and 14 dB. The boundary is approximately a diagonal line with a negative slope. The indicate-*Incremented* regions are above the negative sloped diagonal line (larger Λ and Θ) while indicate-*Un-Incremented* regions are below the line (smaller Λ and Θ). Increasing ΔL increases the Λ intercept and does not affect the slope. The bottom panel of Fig. 6 shows the decision boundary for the ideal observer of Λ and Θ together in the double-rove condition for values of ΔL of 8 and 14 dB. The boundary is a continuous function such that larger Λ and Θ fall into the *Incremented* region. Increasing ΔL shifts the boundary function towards higher values of Λ and Θ.

Without $f_{\Lambda,\Theta|L_{Target}}$, it is not possible to infer the performance of the ideal observer of Λ and Θ together; however, two important aspects of the ideal observer of Λ and Θ together are gained from an analysis of the decision boundary. The first aspect is that the decision boundary in the roving-level condition is different than in the double-rove condition. The second is that the decision boundary in all three conditions is relatively simple and well behaved. If the decision boundary was highly complex, it would be unlikely that subjects would be able to respond accurately based on the boundary. Surprisingly, as discussed in the next section, none of the subjects behave similar to any of the ideal observers considered (Λ alone, Θ alone, and Λ and Θ together).

B. Model Predictions of the Psychophysical Results

Predicting P_F and P_D from the graphical analysis of the probability density functions, the likelihood ratios and the indicator functions is not trivial. The graphical analysis is performed in the $\{\Lambda, \Theta\}$ space while the predictions are made relative to the distractor level and phase. To aid the transition $\overline{\Lambda}$ and $\overline{\Theta}$ (the expectation of Λ and Θ, respectively) are considered. Figure 7 displays contours of equal $\overline{\Lambda}$ (top row) and equal $\overline{\Theta}$ (bottom row) as a function of the distractor level and phase for when the target level is un-incremented (left column) and incremented (right column). The contours of equal $\overline{\Lambda}$ are horizontal lines. The contours are closer for higher distractor levels than for lower distractor levels since the influence of the target on $\overline{\Lambda}$ is greater at lower distractor levels. To achieve an equal $\overline{\Lambda}$, the distractor level needs to be higher with the un-incremented target than with an incremented target; the difference is greatest for lower distractor levels. The contours of equal $\overline{\Theta}$ are diagonal lines (they depend on both the distractor level and phase). For a given distractor level and phase, $\overline{\Theta}$ is larger for the incremented target. Unlike the contours of $\overline{\Lambda}$, the spacing between the contours of $\overline{\Theta}$ is constant.

The modeling described here explores the hypothesis that subjects base their decisions on Λ and Θ as defined in Eqs. 1 and 2, respectively. Specifically, P_F and P_D are predicted under the three experimental conditions for (1) the ideal observer of Λ alone, (2) the ideal observer of Θ alone, and (3) the ideal observer of both Λ and Θ together. Each model can be used to predict $\overline{P_C}$ and $\overline{\beta}$ (as well as $\overline{P_F}$ and $\overline{P_D}$) in all three conditions with any value of ΔL. In addition to these average values, the models can also be used to predict the dependence of P_F and P_D on the distractor level in the roving-level condition and on the distractor level and phase in the double-rove condition. The average values of the predicted $\overline{P_C}$, $\overline{\beta}$, $\overline{P_F}$, and $\overline{P_D}$ are presented first. This is followed by the predictions of P_F and P_D as a function of the distractor level in the roving-level condition. Finally, the dependence of P_F and P_D on the distractor level and phase in the double-rove condition is reported.

Table 2 contains the predictions ($\overline{P_C}$, $\overline{\beta}$, $\overline{P_F}$, and $\overline{P_D}$) of the ideal observers of Λ alone, Θ alone, and both Λ and Θ together in the fixed, roving-level, and double-rove conditions for values of ΔL of 8 and 14 dB. In the fixed condition, the performances of all three models are limited by internal noise and the predictions are determined by the standard deviations chosen for the model. Theoretically, performance is slightly better with a ΔL of 14 dB as opposed to a ΔL of 8 dB, but this difference is small. Further, the ideal observer of both Λ and Θ together has a small advantage over either the ideal observer of Λ alone or the ideal observer of Θ alone. Due to the large values of ΔL used relative to the internal noise, however, the performances of all the models considered with either value of ΔL are nearly perfect.

In the roving-level condition, the ideal observer of both Λ and Θ together is again limited by internal noise and also performs nearly perfectly for both values of ΔL. In contrast, the ideal observers of Λ alone and Θ alone in the roving-level condition are limited by a combination of internal noise and distractor variability. The ideal observer of Θ alone is theoretically unbiased (due to computational rounding errors, the prediction are slightly biased) as expected, $\overline{P_C}$ increases with increasing ΔL. The ideal observer of Λ alone obtains values of $\overline{P_C}$ similar to the ideal observer of Θ alone, but is biased (the ideal observer of Λ alone responds *"Incremented"* with higher probability). Increasing ΔL, increases $\overline{P_C}$ and decreases $\overline{\beta}$; in other words, increasing ΔL only decreases $\overline{P_F}$ and does not affect $\overline{P_D}$. Note that the ideal observer of Λ alone responds *"Incremented"* more frequently than *"Un-Incremented"*, which is the opposite of the slight bias of the subjects.

In the double-rove condition, all the ideal observers considered are limited by a combination of internal noise and distractor variability and therefore none achieves near perfect performance. The predictions of the three observers are different. The observer of Λ and Θ together has the highest $\overline{P_C}$. The observer of Λ alone has a value of $\overline{P_C}$ similar to the observer of Θ alone. For all the observers, as expected, $\overline{P_C}$ increases with increasing ΔL. The observer of Θ alone is unbiased for both values of ΔL. The observer of Λ alone and the observer of Λ and Θ

together, however, have high values of $\overline{\beta}$ for both values of ΔL. For these observers, increasing ΔL, decreases $\overline{P_F}$ and does not affect $\overline{P_D}$ and therefore $\overline{\beta}$ also decreases.

Figure 8 presents the average differences in $\overline{P_C}$ (top panel) and $\overline{\beta}$ (bottom panel) between the ideal observers and the subjects in the three different conditions. In the fixed condition, there are only small differences between the subjects and all three modeled observers; the model observers have a $\overline{P_C}$ less than 0.1 better than the subjects. In the roving-level condition, there are substantial differences between the models. The subjects outperform the ideal observers of Λ alone and Θ alone; on average, the subjects have values of $\overline{P_C}$ 0.2 greater than the ideal observers of Λ alone and Θ alone. This indicates that the subjects must be basing their decisions on more information than is contained in either Λ alone and Θ alone. The ideal observer of Λ and Θ together, has a $\overline{P_C}$ 0.2 better than the subjects indicating that there is sufficient information in Λ and Θ together to predict the average subject performance. In the double-rove condition, the values of $\overline{P_C}$ for the ideal observers of Λ alone and Θ alone are again worse than the subjects (by at least 0.1), while the ideal observer of Λ and Θ together has a value of $\overline{P_C}$ nearly identical to that of the subjects.

By considering the differences in $\overline{P_C}$ between the subjects and modeled observer, one can dismiss certain observer models since they do not have access to sufficient information. Overall, the models do a reasonable job of predicting $\overline{P_C}$ in both the fixed and double-rove conditions. All the models fail to predict $\overline{P_C}$ in the roving-level condition. Since the subjects outperform the ideal observers of Λ alone and Θ alone, these cannot be the correct model for the roving-level condition; the subjects must have access to additional information. The subjects never outperform the ideal observer of Λ and Θ together and therefore the predictions of this model are considered in detail. Although the ideal observers of Λ alone and Θ alone do not have access to sufficient information, considering the predictions of these observers is insightful to understanding the ideal observer of Λ and Θ together.

By considering differences in $\overline{\beta}$, one can gain insight into the discrepancies in the $\overline{P_C}$ between the model observers and the subjects. In the fixed condition, the magnitude of the difference in $\overline{\beta}$ between the modeled observers and the subjects is less than 0.05; the modeled observers respond *Incremented* only slightly more often than the subjects. In the roving-level condition, the differences in $\overline{\beta}$ between the subjects and the modeled observers varies. The ideal observers of Θ alone and Λ and Θ together have nearly the same values of $\overline{\beta}$ as the subjects, while the ideal observer of Λ alone has a value of $\overline{\beta}$ that is 0.4 smaller than the subjects (i.e., the ideal observer of Λ alone responds *Incremented* more often than the subjects). In the double-rove condition, the ideal observers of Λ alone and Λ and Θ together have substantially different values of $\overline{\beta}$ than the subjects (-0.3 and -0.4, respectively). Again, these observers respond *Incremented* more often than the subjects. The ideal observer of Θ alone performs much more similarly to the subjects with a difference in $\overline{\beta}$ of -0.05.

As was the case with the psychophysical results in the roving-level and double-rove conditions, the average values of $\overline{P_C}$ and $\overline{\beta}$ do not completely describe the data since there is an interaction between performance and the distractor level (roving-level and double-rove conditions) and phase (double-rove condition). The predicted P_F and P_D, for the roving-level and double-rove conditions, are shown in Figs. 9, 10, and 11. Figure 9 shows the predicted P_F and P_D as a function of the distractor level in the roving-level condition. Only the predictions for the ideal observers of Λ alone and Θ alone, with values of ΔL of both 8 and 14 dB are shown since the ideal observer of Λ and Θ together performs nearly perfectly. Figures 10 and 11 show the predicted dependence of P_F and P_D on the distractor level and phase for the ideal observers of Λ and Θ together, Λ alone, and Θ alone in the double-rove condition for values of ΔL of 8 and 14 dB, respectively.

In the roving-level condition (cf. Fig. 9), the ideal observer of Λ alone predicts a P_F near zero for low distractor levels and a P_F near unity for high distractor levels. The transition in P_F is extremely rapid and its exact location depends on ΔL. The ideal observer of Λ alone predicts a P_D of nearly unity for all but the lowest distractor levels. The ideal observer of Θ alone predicts a substantially different dependence of P_F and P_D on the distractor level. Both P_F and P_D are near unity for low level distractors and near zero for high level distractors. The P_F and P_D function have rapid transitions between these two probabilities. The distractor level at which the transitions occur depend on ΔL, but the steepness does not. The distractor level at which the transition occurs is smaller for P_F than for P_D.

Figures 10 and 11 show the predicted dependence of P_F and P_D on the distractor level and phase for the ideal observers of Λ and Θ together, Λ alone, and Θ alone in the double-rove condition. Figure 10 contains the predictions with a ΔL of 8 dB while Fig. 11 is for a ΔL of 14 dB. The three models predict substantially different dependencies of P_F and P_D on the distractor level and phase. A general trend is that when performance depends on the distractor level and phase, there is a rapid transition from areas of low probability to areas of high probability. The shapes of these transition regions depend on the model and the statistics (P_F and P_D) but not on ΔL. The locations of the transitions, however, depend on ΔL. The presentation of the predictions in the double-rove condition will begin with the ideal observer of both Λ and Θ together, followed by the ideal observer of Λ alone and then finally the ideal observer of Θ alone.

The ideal observer of both Λ and Θ together predicts a P_F which depends on the distractor level and phase. The predicted P_F is near zero for distractors with low levels and large phases and near unity elsewhere. The predictions of P_D are nearly independent of the distractor level and phase and are near unity for all distractors. Consistent with the results presented in Table 2, the ideal observer of both Λ and Θ together is biased. The ideal observer of Λ alone predicts a P_F that depends predominantly on the distractor level. Low-level distractors yield a P_F near zero and high-level distractors yield a P_F near unity. The distractor level at which the transition occurs is approximately the same for the ideal observer of both Λ and Θ together and the ideal observer of Λ alone. Similar to the ideal observer of both Λ and Θ together, the ideal observer of Λ alone, predicts a P_D which is nearly independent of the distractor level and phase, and near unity for all distractors. Consistent with the change in the predicted P_F, the ideal observer of Λ alone is even more biased than the ideal observer of both Λ and Θ together.

The ideal observer of Θ alone is substantially different than either the ideal observer of both Λ and Θ together or the ideal observer of Λ alone. The ideal observer of Θ alone predicts that both P_F and P_D depend on the distractor level and phase. The transition region from high to low P_F and P_D is nearly linear. Distractors whose level and phase fall in the lower triangle (low level and small phase) have values of P_F and P_D near unity, while distractors whose level and phase fall in the upper triangle (low level and small phase) have values of P_F and P_D near zero. The transition regions of P_F and P_D are not identical resulting in a narrow region of extremely high probability of correct. Due to the similarities of the predictions of P_F and P_D, the ideal observer of Θ alone is the least biased of all the models.

VI. Discussion

This study is a follow up to the work in Chapter II in which it was reported that in a monaural level discrimination task subjects could not focus exclusively on a single ear. In that work, the measured level discrimination thresholds were well predicted by a model based upon an ideal observer of the Λ and Θ dimensions. It was suggested that the decisions of the subjects may be based upon observations of the loudness and position. Here, psychophysical data were collected with a different paradigm, but with nearly identical stimuli and subjects, to further test the model based upon the ideal observer of Λ and Θ. The current results confirm the psychophysical findings that roving the level (and phase) of a contra-aural distractor decreases performance on a monaural level discrimination task. The psychophysical results, however, refute models based upon ideal observers of Λ and Θ (either alone or together).

In the original study which used a multi-interval adaptive paradigm, in order to obtain a probability of correct of 0.7 in the double-rove condition, a target level increment ΔL of 7.8, 5.7 and 11.3 dB was needed for S1, S2, and S4, respectively. In this study which used a one-interval constant increment paradigm, subjects S1 and S2 obtained a $\overline{P_C}$ of 0.69 and 0.72 with a ΔL of 8 dB and S4 obtained a $\overline{P_C}$ of 0.72 with a ΔL of 14 dB. This study does not directly assess the degree to which performance was affected by the change in paradigm; however, the overall effect appears to be small. For a ΔL near the threshold measured in the previous work, performance was approximately the same as the threshold criterion of a probability of correct of 0.7.

Unlike the original study, which included a no-distractor control condition, this work did not investigate the no-distractor condition. Stellmack et al. (2004) argue that two interval monotic (i.e., no distractor) level discrimination is the same as one-interval interaural (i.e., fixed distractor) level discrimination. Jesteadt and Bilger (1974) measured monotic level discrimination thresholds under 2 dB with both one-interval and two-interval paradigms. Based on these previous studies, it would not be expected to measure values of $\overline{P_C}$ near 0.7 (approximately the measured $\overline{P_C}$ of the subjects in the roving-level and double-rove conditions) in a one-interval monotic level discrimination task with values of ΔL of either 8 or 14 dB given the relatively high $\overline{P_C}$ in the fixed condition. Although not directly tested, performance is consistent with our previous conclusions that subjects cannot exclusively attend to a single ear.

In the fixed condition, all three subjects perform nearly perfectly with values of $\overline{P_C}$ near unity (cf. Fig. 1). Subject S4, had the largest ΔL of 14 dB, but the worst performance with a $\overline{P_C}$ of 0.88. Although this is substantially less than one might expect, the relatively small number of trials (2000) in the fixed condition and the one-interval paradigm without any explicit training might account for the relatively poor performance. All three models predict nearly perfect performance in the fixed condition. The lack of explicit training is less of a factor in the roving-level and double-rove conditions since the number of trials is much greater (at least 6,000 and as great as 30,000). In the roving-level condition (6,000 trials for S4 and 12,000 trials for S1 and S2), for all three subjects $\overline{P_C}$ decreases compared to the fixed condition. The ideal observer of both Λ and Θ together does not get worse, but the ideal observers of either Λ alone or Θ alone predict the values of $\overline{P_C}$ obtained by the subjects. In the double-rove condition (24,000 trials for S4 and 30,000 trials for S1 and S2) the performance of the ideal observer of Λ and Θ together is degraded appreciably. In this condition, all three models predict the values of $\overline{P_C}$ obtained by the subjects to within 0.2. These findings are similar to those reported in Chapter II. Specifically, all three models predicted the average performance in the fixed and double-rove conditions and the ideal observer of both Λ and Θ together substantially outperformed all the subjects in the roving-level condition.

Although the models predict $\overline{P_C}$ to some degree, an exploration of the predictions and data in greater depth reveals serious shortcomings of these reasonable models. In particular, the subjects outperform the ideal observers of either Λ alone or Θ alone in both the roving-level and double-rove conditions. In these conditions, the performance of the ideal observers of either Λ alone or Θ alone is limited by a combination of internal noise and distractor variability. Removing the internal noise does not sufficiently increase performance. Additionally, the ideal observers of either Λ alone or Θ alone demonstrate a bias that the subjects do not. Further evidence against the models based on ideal observers of either Λ alone or Θ alone can be seen in a comparison of Figs. 2 and 9. The dependence of P_F and P_D on the distractor level in the roving-level condition for the subjects differs from the predictions of both models. Finally, the predictions of P_F and P_D in the double-rove condition are also in disagreement with the subject data. These discrepancies imply that the subjects are basing their decisions on more than either Λ alone or Θ alone.

One could simply dismiss the model based on the ideal observer of Λ and Θ together due to the sizable discrepancy between the predictions and the values of $\overline{P_C}$ obtained by the subjects in the roving-level condition. We believe, however, that although large, the discrepancy is not serious since the ideal observer is performing better than the subjects. One possible remedy to this discrepancy is to use a non-ideal observer. Calculations of performance in the roving-level with a non-ideal indicator function are not presented, since, we believe that the model based on Λ and Θ together has other flaws. Specifically, these flaws are most evident in the double-rove condition.

In the double-rove condition, there is only sufficient information (not an excess of information) contained in Λ and Θ together to predict the values of $\overline{P_C}$ obtained by the subjects. In other words, the ideal observer of Λ and Θ together does not greatly outperform the subjects. The model, however, is much more biased than the subjects (cf. Fig. 8). Further, the predicted dependence of P_F and P_D on the distractor level and phase does not match the subject performance (refer to Figs. 3, 10, and 11). When evaluating the models with $\overline{P_C}$, all three models seem to have reasonable amounts of predictive power. By evaluating the models on the finer aspects of the data (P_F and P_D as a function of the distractor properties), one can determine that the models in fact have very little predictive power. Although these binaural models failed to predict the data, a monaural model, would predict no effect of the distractor since the target is the same in all three conditions and the distractor is presented contra-aurally to the target. By using models based on binaural cues (e.g., Λ and Θ) performance that varies across the different conditions can be predicted.

We conclude that models based on an ideal observer of decision variables related to the perceived loudness (Λ) and position (Θ) or their combination are inadequate to predict the psychophysical results. We attribute the failure to predict the data to three independent aspects of the models. These three aspects are the use of (1) only two perceptual dimensions, (2) decision variables which drastically simplify the perceptual dimensions, and (3) ideal observers.

Psychophysical studies of time-intensity trading have demonstrated that the perceived loudness and position does not completely describe the perception of a dichotic tone (Hafter and Jeffress 1968a; Hafter and Carrier 1972; Smith 1973; Ruotolo et al. 1979). Additional perceptual aspects such as the "time-image" (Hafter and Carrier 1972), "level-difference" (Hartmann and Constan 2002), and "spatial width" (Ruotolo et al. 1979) have been suggested. One could also include a monaural loudness dimension, but this percept would have to be extremely noisy to allow the model to predict the decreases in performance associated with the addition of the distractor. Adding dimensions, however, increases the theoretical complexity of the model.

In the modeling work presented here, Λ is the sum of the intensities at the left and right ears and Θ is a weighted sum of the ILD and ITD. Extensive studies on the perception of both overall level (e.g., Viemeister 1988) and interaural differences (e.g., Durlach and Colburn 1978) have been conducted in the past. At best, the definitions of Λ and Θ used are first-order approximations to the perceived loudness and position. The exact forms of the definitions of Λ and Θ were chosen to balance the following three competing factors (1) correlation to perception, (2) computational simplicity, and (3) theoretical simplicity.

Increasing the complexity of the definitions of Λ (e.g., a binaural version of the loudness incorporating the findings of Edmonds and Culling 2006) and Θ (e.g., Stern and Colburn 1978) might lead to better correlation with perception. Even with the simple definitions of Λ and Θ, however, a mixture of analytical and numerical techniques were required to compute the model predictions. Increases in the complexity of the definitions, could result in an intractable model. Further, increasing the complexity of the definitions would raise questions as to what aspects of the definitions are crucial to predict the psychophysical results.

The final aspect of the model which may give rise to the poor predictive power is the use of an ideal observer. Overall, all three subjects in this experiment performed similarly achieving an average $\overline{P_C}$ between 0.68 and 0.72 in the double-rove condition (cf. Fig. 1). Notably, subject S4 did require a larger ΔL than the other two subjects. The performance as a function of the distractor level in the roving-level condition and as a function of the distractor level and phase in the double-rove condition were extremely different for the three subjects. The inter-subject variability is likely to be even greater than that reported here since in addition to the three subjects who completed the study, two subjects were dismissed from the study due to extremely poor performance[6].

An ideal observer cannot predict inter-subject variability. Since there are an infinite number of non-ideal observers, each subject could be modeled as a different non-ideal observer. For example one non-ideal observer could weight Λ more than Θ and another could weight Λ and Θ equally. One could also "jitter" the decision boundaries (cf. Fig. 6). The properties of the jitter could be adjusted for each subject. One could also impose a cost for attending to multiple dimensions (i.e., divided attention). The imposed cost matrix could be different for each subject.

Predictions of models based on many of the possible non-ideal observers of the currently defined Λ and Θ can be calculated. The predictions of the ideal observer of Λ and Θ are so different from the psychophysical data that substantial modifications to the ideal observer would be required. One problem with using a non-ideal observer is that the $\overline{P_C}$ predicted by the ideal observer of Λ and Θ is only marginally better than the $\overline{P_C}$ obtained by the subjects. Since the $\overline{P_C}$ predicted by the non-ideal observer will be less than that predicted by the ideal observer[7], a non-ideal observer will most likely not have access to sufficient information and will be outperformed by the subjects.

After resolving the first two issues with the model (the simplicity of the dimensions and the number of dimensions), there may be more information available to a model decision device which would allow for the use of non-ideal observers. Designing and implementing a model which addresses these issues is beyond the scope of this chapter, but is included in Chapter IV. The simplicity of the dimensions seems secondary to the number of dimensions. The relatively simple Λ and Θ have strong predictive power for many psychophysical tasks, and studies of time-intensity trading (Hafter and Carrier 1972; Ruotolo et al.

1979) have demonstrated that two dimensions are insufficient to characterize the stimuli. The modeling framework presented here can be expanded to three or more dimensions.

VII. Summary

Monaural level discrimination was measured with a contra-aural distractor. The subjects could neither attend exclusively to the ear at which the target was presented, nor make optimal use of the information carried by Λ and Θ (correlates to the perceived loudness and position). The finding that subjects could not exclusively attend to the ear at which the target was presented is in agreement with the results of Chapter II, and demonstrates that those findings are not highly dependent on the experimental paradigm. Also in agreement with the findings of Chapter II, the subjects obtained a probability of correct similar to that of the ideal observer of Λ and Θ together. The subjects, however, did not make optimal use of the information carried by Λ and Θ. Performance as a function of the distractor parameters (level and phase) was substantially different between the subjects and the ideal observer, suggesting that the subjects used additional information than that which is carried by Λ and Θ. Finally, there was little inter-subject variability in the overall probabilities of correct and bias, but substantial inter-subject variability in performance (P_F and P_D) as a function of the distractor parameters, implying that the listeners used different listening strategies.

Acknowledgements

This research was supported by NIH/NIDCD Grant DC00100.

Appendix

This appendix presents a detailed derivation of the performance of the ideal observer of the two dimensions defined by Eqs. 1 and 2 in the text. Specially, the probability of the ideal observer responding *Incremented* conditioned on the target level, the distractor level, and the distractor phase is derived. Both the notation and the general framework of the derivation are similar to that in Chapter II. The derivation only considers the Λ and Θ together model and does not explicitly derive the Λ alone or Θ alone models. These derivations, however, follow trivially from the given derivation.

The derivation begins by substituting the experimental values into Eqs. 1 and 2. In the experiment the level of the left ear L_{Left} is the target level L_{Target}; the un-incremented target level is equal to the reference level L_0 (in decibels) and the incremented target level is equal to the sum of L_0 and the target increment ΔL (in decibels). The level of the right ear L_{Right} is the sum of L_0 and the distractor level increment A (in decibels). The interaural time difference T is the negative of the distractor phase Φ divided by the distractor frequency ω. Making these substitutions into Eqs. 1 and 2 results in

$$\Lambda = 10\log_{10}\left(10^{\frac{L_{Target}}{10}} + 10^{\frac{L_0+A}{10}}\right) + N_\Lambda, \tag{A1}$$

and

$$\Theta = L_{Target} - \left(L_0 + A\right) - \frac{k}{\omega}\Phi + N_\Theta . \tag{A2}$$

On each trial there is a single observation of both Λ and Θ. These observations are denoted Λ and Θ, respectively. The maximum likelihood observer is defined in terms of two joint probability density functions. The first is the probability of observing Λ and Θ given L_{Target} is equal to the reference level L_0, $f_{\Lambda,\Theta|L_{Target}}\left(\lambda,\theta|L_0\right)$. The second is the probability of Λ and Θ given the target level was incremented, $f_{\Lambda,\Theta|L_{Target}}\left(\lambda,\theta|L_0+\Delta L\right)$. The log-likelihood ratio $\eta_{\Lambda,\Theta}$ is defined in Eq. 3.

The ideal observer is defined by a binary indicator function $\psi_{\Lambda,\Theta}$ which depends on $\eta_{\Lambda,\Theta}$. Specifically, when $\psi_{\Lambda,\Theta}$ is equal to zero, the target level is most likely incremented. Similarly, when $\psi_{\Lambda,\Theta}$ is equal to one the target level is most likely to be equal to the reference level L_0. Mathematically the indicator function is

$$\psi_{\Lambda,\Theta}\left(\lambda,\theta,\Delta L\right) = \begin{Bmatrix} 1 & \text{when} & \eta_{\Lambda,\Theta}\left(\lambda,\theta,\Delta L\right) \geq 1 \\ 0 & \text{when} & \eta_{\Lambda,\Theta}\left(\lambda,\theta,\Delta L\right) < 1 \end{Bmatrix}.$$

In order to calculate $\psi_{\Lambda,\Theta}$, $f_{\Lambda,\Theta|L_{Target}}$ must be determined. In Chapter II it was found that a closed form solution to $f_{\Lambda,\Theta|L_{Target}}$ could not be found with analytic methods. An approximation of $f_{\Lambda,\Theta|L_{Target}}$ was calculated in the appendix to Chapter II with a mixture of analytic and numerical techniques. That derivation is not presented here. Recall that the approximations used in that derivation involved approximating integration with summation. No Monte-Carlo type simulations were required.

In Chapter II, the probability of correct as a function of ΔL was calculated. In this work, P_F and P_D are calculated as a function of the distractor level and phase. The probabilities P_F and P_D depend on the indicator function and the joint probability of Λ and Θ given (1) a, the particular value of the random variable A (the distractor level), (2) ϕ the particular value of the random variable Φ, and (3) L_{Target}. This conditional joint probability of Λ and Θ is denoted as $f_{\Lambda,\Theta|A,\Phi,L_{Target}}$. The probabilities P_F and P_D can be written as

$$P_F(a,\phi,\Delta L)=\int\int\psi_{\Lambda,\Theta}(\lambda,\theta,\Delta L)\,f_{\Lambda,\Theta|A,\Phi,L_{Target}}(\lambda,\theta|a,\phi,L_0)\,d\lambda d\theta$$

and

$$P_D(a,\phi,\Delta L)=\int\int\psi_{\Lambda,\Theta}(\lambda,\theta,\Delta L)\,f_{\Lambda,\Theta|A,\Phi,L_{Target}}(\lambda,\theta|a,\phi,L_0+\Delta L)\,d\lambda d\theta\,. \qquad \text{(A3)}$$

The derivation of the model proceeds by considering only P_D. The derivation of P_F follows trivially from the derivation of P_D by adjusting L_{Target}. To solve for P_D, the definition of conditional probability is used to expand the joint density function $f_{\Lambda,\Theta|A,\Phi,L_{Target}}$ to

$$f_{\Lambda,\Theta|A,\Phi,L_{Target}}=f_{\Lambda|A,\Phi,L_{Target}}(\lambda|a,\phi,L_0+\Delta L)\,f_{\Theta|\Lambda,A,\Phi,L_{Target}}(\theta|\lambda,a,\phi,L_0+\Delta L)\,.$$

Noting from Eq. A1 that Λ is independent of Φ and from Eq. A2 that Θ is conditionally independent of Λ when A is given, allows $f_{\Lambda,\Theta|A,\Phi,L_{Target}}$ to be simplified to

$$f_{\Lambda,\Theta|A,\Phi,L_{Target}} = f_{\Lambda|A,L_{Target}}\left(\lambda|a,L_0+\Delta L\right) f_{\Theta|A,\Phi,L_{Target}}\left(\theta|a,\phi,L_0+\Delta L\right)$$

Making a substitutions of Λ based on Eq. A1 and Θ based on Eq. A2 and using the definition of conditional probability yields

$$f_{\Lambda,\Theta|A,\Phi,L_{Target}} = f_{N_\Lambda}\left(\lambda-\mu_\Lambda(a)\right) f_{N_\Theta}\left(\theta-\mu_\Theta(a,\phi)\right)$$

where $\mu_\Lambda(a)$ is equal to $10\log_{10}\left(10^{\frac{L_0+\Delta L}{10}}+10^{\frac{L_0+a}{10}}\right)$ and $\mu_\Theta(a,\phi)$ is equal to $\Delta L - a - \frac{k}{\omega}\phi$. Then substituting the Gaussian density functions of N_Λ and N_Θ gives

$$f_{\Lambda,\Theta|A,\Phi,L_{Target}} = \frac{1}{2\pi\sigma_\Lambda\sigma_\Theta} e^{-\frac{(\lambda-\mu_\Lambda(a))^2}{2\sigma_\Lambda^2}} e^{-\frac{(\theta-\mu_\Theta(a,\phi))^2}{2\sigma_\Theta^2}}.$$

Substituting the expanded form of $f_{\Lambda,\Theta|A,\Phi,L_{Target}}$ into Eq. A3 yields

$$P_D(a,\phi,\Delta L) = \frac{1}{2\pi\sigma_\Lambda\sigma_\Theta}\int e^{-\frac{(\theta-\mu_\Theta(a,\phi))^2}{2\sigma_\Theta^2}} \int \psi_{\Lambda,\Theta}(\lambda,\theta,\Delta L) e^{-\frac{(\lambda-\mu_\Lambda(a))^2}{2\sigma_\Lambda^2}} d\lambda d\theta .$$

To solve for P_D, we note that the conditions of the experiment lead to an indicator function $\psi_{\Lambda,\Theta}$ which is relatively simple. For a fixed ΔL, $\psi_{\Lambda,\Theta}$ is characterized by a boundary function $\Psi_{\Lambda,\Theta}$ which divides the $\{\Lambda,\Theta\}$ plane into two regions (cf., Fig. 6). The probability of a detection can be expanded by splitting the integration over Λ into two segments, such that

$$P_D(a,\phi,\Delta L)=\frac{1}{2\pi\sigma_\Lambda\sigma_\Theta}\int e^{-\frac{(\theta-\mu_\Theta(a,\phi))^2}{2\sigma_\Theta^2}}\left[\begin{array}{l}\displaystyle\int_{-\infty}^{\Psi_{\Lambda,\Theta}(\theta,\Delta L)}\psi_{\Lambda,\Theta}(\lambda,\theta,\Delta L)e^{-\frac{(\lambda-\mu_\Lambda(a))^2}{2\sigma_\Lambda^2}}d\lambda\\ \displaystyle+\int_{\Psi_{\Lambda,\Theta}(\theta,\Delta L)}^{\infty}\psi_{\Lambda,\Theta}(\lambda,\theta,\Delta L)e^{-\frac{(\lambda-\mu_\Lambda(a))^2}{2\sigma_\Lambda^2}}d\lambda\end{array}\right]d\theta .$$

Noting that $\psi_{\Lambda,\Theta}$ is equal to one when Λ is greater than $\Psi_{\Lambda,\Theta}$ and zero otherwise, allows for P_D to be simplified to

$$P_D(a,\phi,\Delta L)=\frac{1}{2\pi\sigma_\Lambda\sigma_\Theta}\int e^{-\frac{(\theta-\mu_\Theta(a,\phi))^2}{2\sigma_\Theta^2}}\int_{\Psi_{\Lambda,\Theta}(\theta,\Delta L)}^{\infty}e^{-\frac{(\lambda-\mu_\Lambda(a))^2}{2\sigma_\Lambda^2}}d\lambda d\theta .$$

Defining the integral of an exponential squared function as

$$\mathrm{G}(\alpha,\beta,\mu,\sigma)=\int_\alpha^\beta e^{-\frac{(x-\mu)^2}{2\sigma^2}}dx,$$

allows for further simplification of P_D

$$P_D(a,\phi,\Delta L)=\frac{1}{2\pi\sigma_\Lambda\sigma_\Theta}\int e^{-\frac{(\theta-\mu_\Theta(a,\phi))^2}{2\sigma_\Theta^2}}\mathrm{G}\left(\Psi_{\Lambda,\Theta}(\theta,\Delta L),\infty,\mu_\Lambda(a),\sigma_\Lambda\right)d\theta .$$

In the experiment, the maximum likelihood boundary $\Psi_{\Lambda,\Theta}$ with a fixed ΔL is relatively smooth; the boundary $\Psi_{\Lambda,\Theta}$ can be approximated as a constant over a small range of Θ. We denote $\theta[m]$ as a sampled version of the continuous random variable Θ. Splitting the integration over Θ into M minus one segments (where M is the number of samples of Θ), P_D can be written as

$$P_D(a,\phi,\Delta L)=\frac{1}{2\pi\sigma_\Lambda\sigma_\Theta}\sum_{m=1}^{M-1}\int_{\theta[m]}^{\theta[m+1]} e^{-\frac{(\theta-\mu_\Theta(a,\phi))^2}{2\sigma_\Theta^2}} \mathrm{G}\left(\Psi_{\Lambda,\Theta}(\theta,\Delta L),\infty,\mu_\Lambda(a),\sigma_\Lambda\right)d\theta .$$

Noting that $\Psi_{\Lambda,\Theta}$ is nearly independent of Θ over the range $\{\Theta[m],\Theta[m+1]\}$, P_D is well approximated even when the G function is moved outside the integration (but still inside the summation) only, namely,

$$P_D(a,\phi,\Delta L)\approx\frac{1}{2\pi\sigma_\Lambda\sigma_\Theta}\sum_{m=1}^{M-1}\mathrm{G}\left(\Omega_{\Lambda,\Theta}(\theta[m],\Delta L),\infty,\mu_\Lambda(a),\sigma_\Lambda\right)\int_{\theta[m]}^{\theta[m+1]} e^{-\frac{(\theta-\mu_\Theta(a,\phi))^2}{2\sigma_\Theta^2}} d\theta .$$

Substituting in the $\mathrm{G}(\alpha,\beta,\mu,\sigma)$ notation gives

$$P_D(a,\phi,\Delta L)\approx\frac{1}{2\pi\sigma_\Lambda\sigma_\Theta}\sum_{m=1}^{M-1}\begin{bmatrix}\mathrm{G}\left(\Omega_{\Lambda,\Theta}(\theta[m],\Delta L),\infty,\mu_\Lambda(a),\sigma_\Lambda\right)\\ \times\mathrm{G}\left(\theta[m],\theta[m+1],\mu_\Theta(a,\phi),\sigma_\Theta\right)\end{bmatrix}.$$

This approximation of P_D can be implemented computationally once $\Psi_{\Lambda,\Theta}$ is calculated.

---

[1] For this subject, the first few hours of training were with a multi-interval paradigm in a target only condition. Following the satisfactory performance in this target only condition, a few hours of training in the double-rove condition with a multi-interval, adaptive paradigm was conducted. The subject also performed similarly to the other subjects in this condition and had threshold values of ΔL substantially less than 20 dB. Finally the subject was trained in the double-rove condition with the one-interval paradigm with a ΔL of 20 dB. With the one-interval

paradigm the subject reported total confusion and performed at near chance levels.

[2] The abilities of the subjects who were excluded from the study to use the labels were not assessed. None of the subjects, however, complained about the use of the labels

[3] A simple example of a biased ideal observer is the case of two six-sided dice (one fair and one "loaded" slightly towards six). If the ideal observer must decide which die was rolled (fair or loaded) based on the outcome of a single roll, the ideal observer will only respond "loaded" for a roll of six and will respond "fair" for all other outcomes (one through five). Since six is only slightly more probable than the other values, the ideal observer is biased.

[4] The chosen value of 0.05 dB per μs for the intensity-time trading ratio k matches the classic time-intensity trading ratio of 20 μs per dB (Whitworth and Jeffress 1961; Hafter and Jeffress 1968b; Hafter and Carrier 1972). By using an intensity-time trade instead of a time-intensity trade, both the Λ and Θ dimensions are expressed in decibels.

[5] Note that $f_{\Lambda|L_{Target}}$ and $f_{\Theta|L_{Target}}$ are the marginal probability density functions of the joint density function $f_{\Lambda,\Theta|L_{Target}}$. Since Λ and Θ are not independent

(both depend on the target level and the distractor level), $f_{\Lambda,\Theta|L_{Target}}$ cannot be reconstructed from $f_{\Lambda|L_{Target}}$ and $f_{\Theta|L_{Target}}$.

[6] The two subjects that did not obtain satisfactory performance performed similar to the other subjects in the fixed condition (high probability of correct), but performed near chance in the double-rove condition even with a ΔL of 20 dB.

[7] The maximum likelihood ideal observer maximizes P_C.

References

Bernstein, L. R. (**2004**). "Sensitivity to interaural intensitive disparities: Listeners' use of potential cues," J. Acoust. Soc. Am. **115,** 3156-60.

Bonnel, A. M., and Hafter, E. R. (**1998**). "Divided attention between simultaneous auditory and visual signals," Percept Psychophys **60,** 179-90.

Brungart, D. S., and Simpson, B. D. (**2002**). "Within-ear and across-ear interference in a cocktail-party listening task," J. Acoust. Soc. Am. **112,** 2985-95.

Durlach, N. I., and Colburn, H. S. (**1978**). "Binaural phenomena" in *Handbook of perception,* Vol. IV edited by E. C. Carterette and M. P. Friedman.(Academic Press, New York)

Edmonds, B., A., and Culling, J., F. (**2006**). "Interaural correlation and loudness," The Journal of the Acoustical Society of America **119,** 3235-6.

Hafter, E. R., and Jeffress, L. A. (**1968a**). "Two-image lateralization of tones and clicks," J. Acoust. Soc. Am. **44,** 563-9.

Hafter, E. R., and Jeffress, L. A. (**1968b**). "Two-image lateralization of tones and clicks," The Journal of the Acoustical Society of America **44,** 563-9.

Hafter, E. R., and Carrier, S. C. (**1972**). "Binaural interaction in low-frequency stimuli: The inability to trade time and intensity completely," J. Acoust. Soc. Am. **51,** 1852-62.

Hartmann, W. M., and Constan, Z. A. (**2002**). "Interaural level differences and the level-meter model," J. Acoust. Soc. Am. **112,** 1037-45.

Jesteadt, W., and Bilger, R. C. (**1974**). "Intensity and frequency discrimination in one- and two-interval paradigms," The Journal of the Acoustical Society of America **55,** 1266-76.

Kidd, G. J., Mason, C. R., Arbogast, T. L., Brungart, D. S., and Simpson, B. D. (**2003**). "Informational masking caused by contralateral stimulation," J. Acoust. Soc. Am. **113,** 1594-603.

Rowland, R. C. J., and Tobias, J. V. (**1967**). "Interaural intensity difference limen," Journal of Speech and Hearing Research **10,** 745-56.

Ruotolo, B. R., Stern, R. M. J., and Colburn, H. S. (**1979**). "Discrimination of symmetric time-intensity traded binaural stimuli," J. Acoust. Soc. Am. **66,** 1733-7.

Smith, L. E. (**1973**). "The effects of time and intensity on the lateralization of sounds," Psychology. Urbana, University of Illinois at Urbana-Champaign: 60.

Stellmack, M. A., Viemeister, N. F., and Byrne, A. J. (**2004**). "Monaural and interaural intensity discrimination: Level effects and the "binaural advantage"," J. Acoust. Soc. Am. **116,** 1149-59.

Stern, R. M., Jr., and Colburn, H. S. (**1978**). "Theory of binaural interaction based in auditory-nerve data. IV. A model for subjective lateral position," J. Acoust. Soc. Am. **64,** 127-40.

Taylor, M. M., and Clarke, D. P. J. (**1971**). "Monaural detection with contralateral cue (MDCC). II. Interaural delay of cue and signal," J. Acoust. Soc. Am. **49,** 1243-53.

Taylor, M. M., Clarke, D. P. J., and Smith, S. M. (**1971a**). "Monaural detection with contralateral cue (MDCC). III. Sinusoidal signals at a constant performance level," J. Acoust. Soc. Am. **49,** 1795-804.

Taylor, M. M., Smith, S. M., and Clarke, D. P. (**1971b**). "Monaural detection with contralateral cue (MDCC). IV. Psychometric functions with sinusoidal signals," J. Acoust. Soc. Am. **50,** 1151-61.

Viemeister, N. F. (**1988**). "Psychophysical aspects of auditory intensity coding" in *Auditory function: Neurobiological bases of hearing,* Vol. edited by G. M. Edelman, W. E. Gall and W. M. Cowan.(John Wiley & Sons, NYC)

Whitworth, R. H., and Jeffress, L. A. (**1961**). "Time vs intensity in the localization of tones," The Journal of the Acoustical Society of America **33,** 925-9.

Yost, W. A., Penner, M. J., and Feth, L. L. (**1972**). "Signal detection as a function of contralateral sinusoid-to-noise ratio," J. Acoust. Soc. Am. **51,** 1966-70.

Zurek, P. M. (**1979**). "Measurements of binaural echo suppression," J. Acoust. Soc. Am. **66,** 1750-7.

Zwislocki, J. J. (**1972**). "A theory of central auditory masking and its partial validation," J. Acoust. Soc. Am. **52,** 644-59.

Tables and Figures

	Frequency (Hz)	Duration (ms)	Phase (radians)	Level (dB SPL)
Fixed	600	300	0	50
Roving-Level	600	300	0	$Uniform(50,80)$
Double-Rove	600	300	$Uniform\left(-\frac{\pi}{2},\frac{\pi}{2}\right)$	$Uniform(50,80)$

Table 1. Distractor properties in the three conditions. In all conditions the target has a frequency of 600 Hz, a duration of 300 ms, a phase of zero and a reference level of 50 dB SPL. The distractor was presented simultaneously but contra-aurally to the target. When roved, the values of the level and phase were chosen from uniform distributions.

Model	Condition	ΔL	$\overline{P_C}$	$\overline{\beta}$	$\overline{P_D}$	$\overline{P_F}$
Λ and Θ	Fixed	8 dB	1.00	0.00	1.00	0.00
		14 dB	1.00	0.00	1.00	0.00
	Roving-Level	8 dB	1.00	0.00	1.00	0.00
		14 dB	1.00	0.00	1.00	0.00
	Double-Rove	8 dB	0.70	0.27	0.97	0.57
		14 dB	0.82	0.16	0.98	0.34
Λ only	Fixed	8 dB	1.00	0.00	1.00	0.00
		14 dB	1.00	0.00	1.00	0.00
	Roving-Level	8 dB	0.62	0.37	0.99	0.75
		14 dB	0.72	0.28	0.99	0.56
	Double-Rove	8 dB	0.62	0.37	0.99	0.75
		14 dB	0.72	0.28	0.99	0.56
Θ only	Fixed	8 dB	1.00	0.00	1.00	0.00
		14 dB	1.00	0.00	1.00	0.00
	Roving-Level	8 dB	0.63	0.07	0.70	0.43
		14 dB	0.73	0.03	0.76	0.30
	Double-Rove	8 dB	0.60	0.00	0.60	0.40
		14 dB	0.67	0.00	0.66	0.33

Table 2. Model Predictions of the ideal observer of Λ and Θ together, Λ alone and Θ alone for the three conditions for the two values of ΔL used.

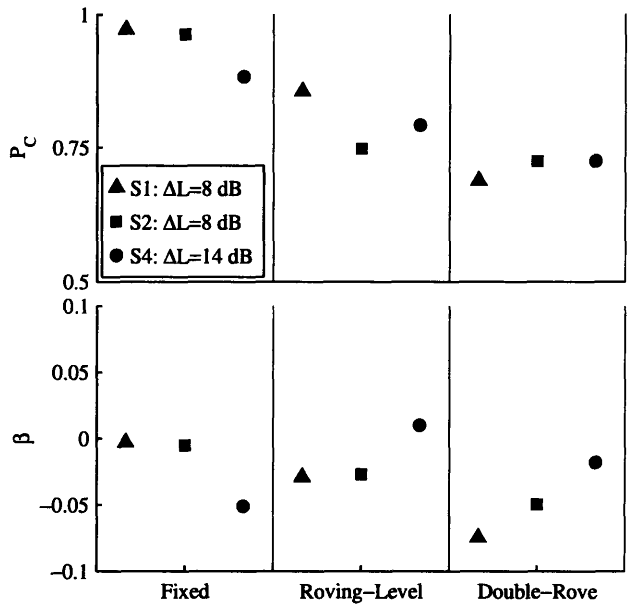

Figure 1. Probability of correct $\overline{P_C}$ (top panel) and the bias towards responding *"Incremented"* $\overline{\beta}$ (bottom panel) for each subject and conditions. The 95 percent confidence intervals, based on a binomial distribution with the appropriate number of trials for each condition, are on the order of the size of the symbols and are not shown.

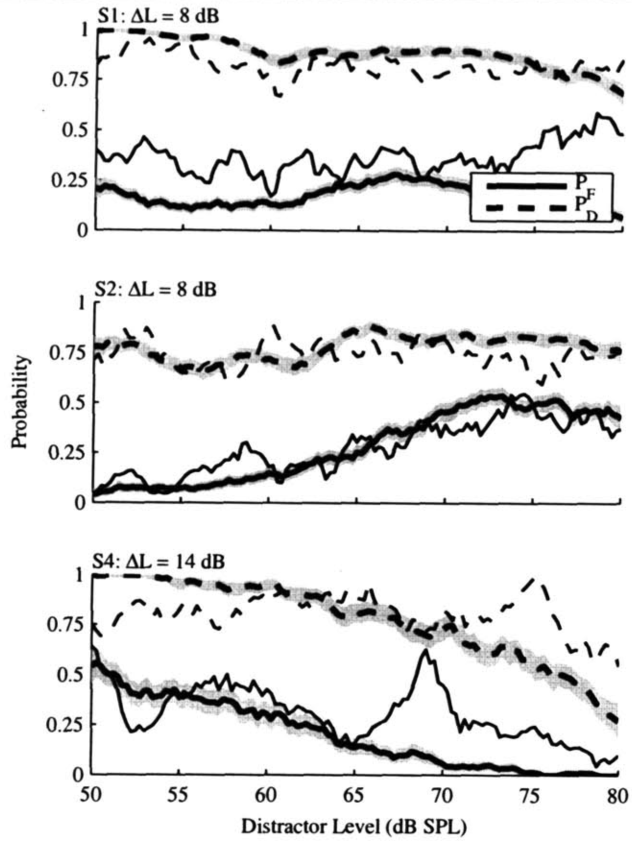

Figure 2. Each panel shows P_F (solid) and P_D (dashed) as a function of the distractor level in the roving-level condition (thick lines) for subjects S1, S2, and S4, respectively. The shaded regions around the functions represent the 95 percent confidence intervals based on a binomial distribution. Also shown as thin lines are P_F and P_D as a function of the distractor level in double-rove condition, for trials in which the distractor phase was near zero.

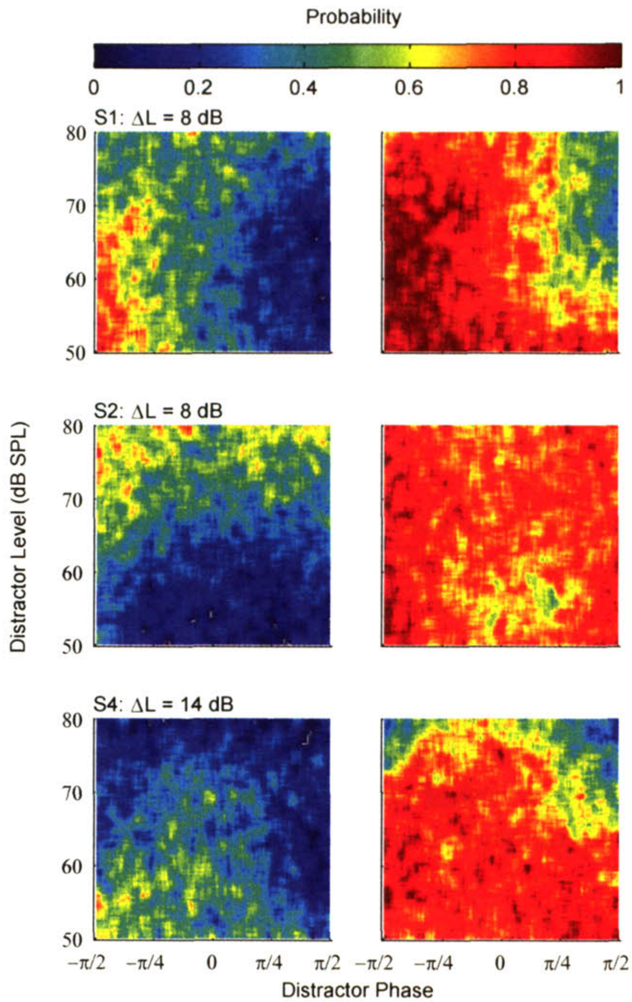

Figure 3. Two-dimensional surface plots of P_F (left column) and P_D (right column) as a function of the distractor level and phase in the double-rove condition for subjects S1 (top row), S2 (middle row), and S4 (bottom row). Areas of high probability are red and areas of low probability are blue. Each surface plots consist of approximately 20,000 overlapping bins with a width of 40 μs and a height of 2 dB.

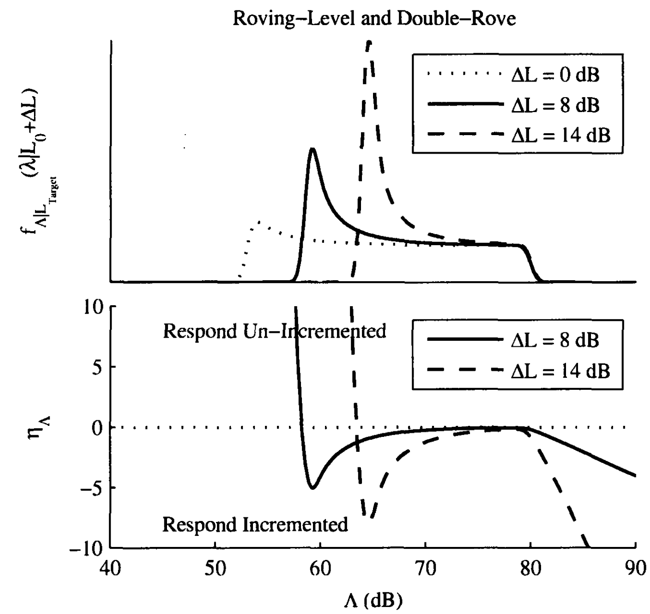

Figure 4. The top panel contains the probability density functions for the ideal observer of Λ alone in the roving-level and double-rove conditions with a ΔL of 0, 8, and 14 dB. The bottom panel contains the log-likelihood ratio for the ideal observer of Λ alone in the roving-level and double-rove conditions with a ΔL of 8 and 14 dB.

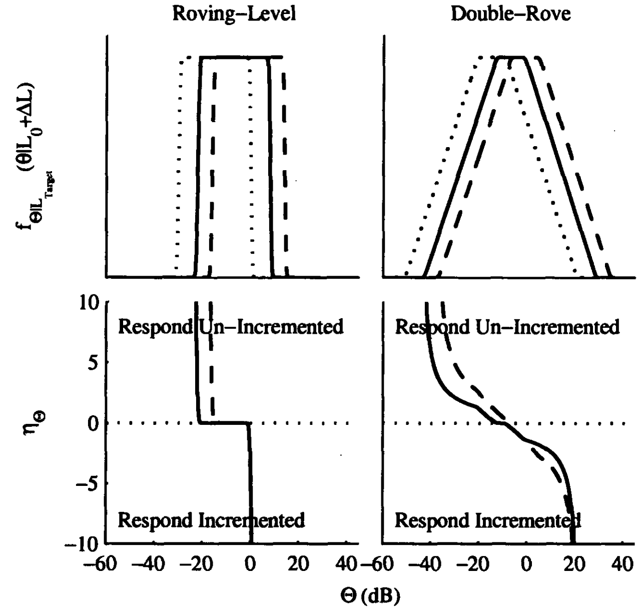

Figure 5. Similar to Fig. 4 except for the ideal observer of Θ alone. The probability density functions and log-likelihood ratio vary across the conditions. Results for the roving-level condition are in the left column and in the right column for the double-rove condition. The dotted, solid, and dashed lines correspond to values of ΔL of 0, 8, and 14 dB.

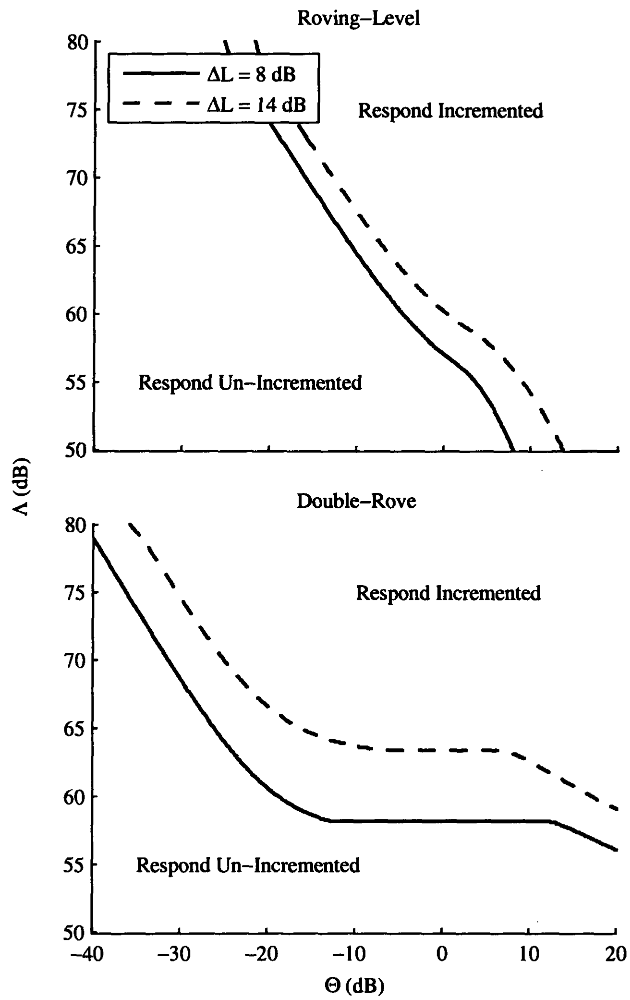

Figure 6. Decision boundaries of the ideal observer of Λ and Θ together for the roving-level condition (top panel) and double-rove condition (bottom panel) for values of ΔL of 8 dB (solid lines) and 14 dB (dashed lines).

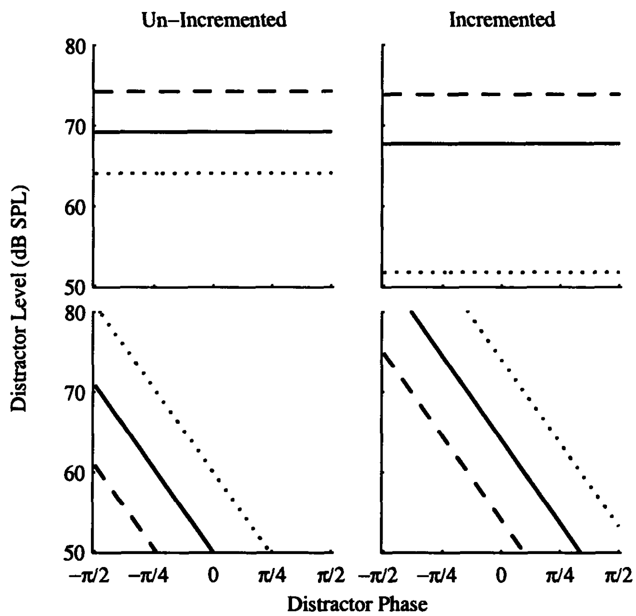

Figure 7. Contours of equal $\overline{\Lambda}$ (top row) and equal $\overline{\Theta}$ (bottom row) for the unincremented target (left column) and incremented target (right column). The target level increment ΔL is 14 dB. The contours are for values of $\overline{\Lambda}$ equal to 64.25 dB (dotted), 69.25 dB (solid), and 74.25 dB (dashed) and values of $\overline{\Theta}$ equal to −10 dB (dotted), 0 dB (solid), and 10 dB (dashed).

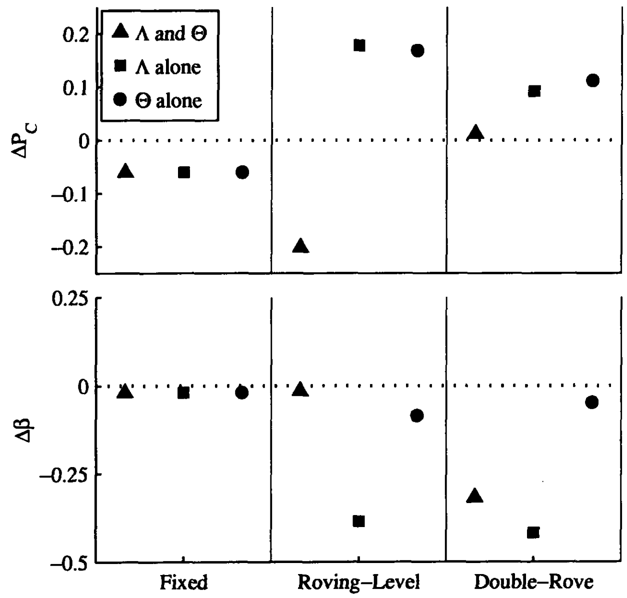

Figure 8. Differences between the average subject $\overline{P_C}$ (top panel) and $\overline{\beta}$ (bottom panel) and the ideal observer of Λ and Θ together (triangles), Λ alone (squares), and Θ alone (circles) in the three conditions. Negative values of $\overline{P_C}$ are indicative of the subjects having a lower probability of correct than the modeled observer. Negative values of $\overline{\beta}$ are indicative of the modeled observer responding *Incremented* with a greater probability than the subjects. In both the $\overline{P_C}$ and $\overline{\beta}$ panels the perfect match reference is shown as a dotted horizontal line at zero.

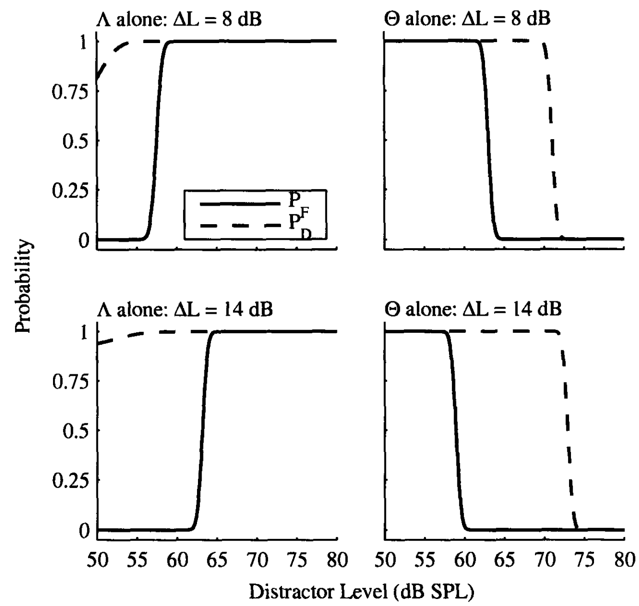

Figure 9. Predictions of the ideal observer of Λ alone (left column), and Θ alone (right column) in the roving-level condition with a ΔL of 8 dB (top row) and a ΔL of 14 dB (bottom row). The solid line is P_F and the dashed line is P_D as in Fig. 2. Predictions for the ideal observer of Λ and Θ together are not shown since performance is nearly perfect.

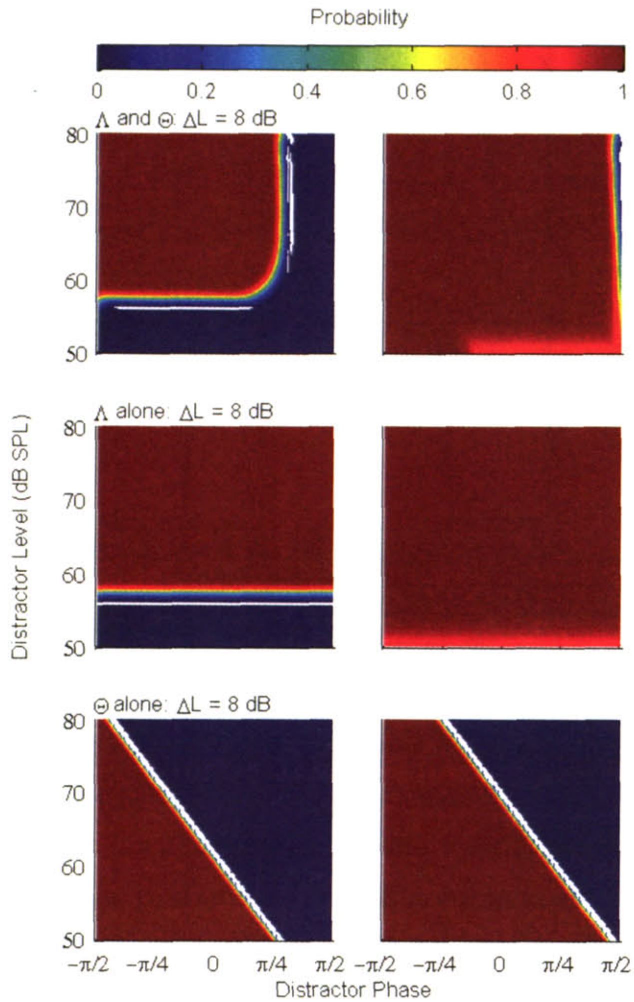

Figure 10. Predictions of the ideal observer of Λ and Θ together (top row), Λ alone (middle row), and Θ alone (bottom row) in the double-rove condition with a ΔL of 8 dB. The predictions consist of P_F (left column) and P_D (right column). Areas of high probability are red and areas of low probability are blue.

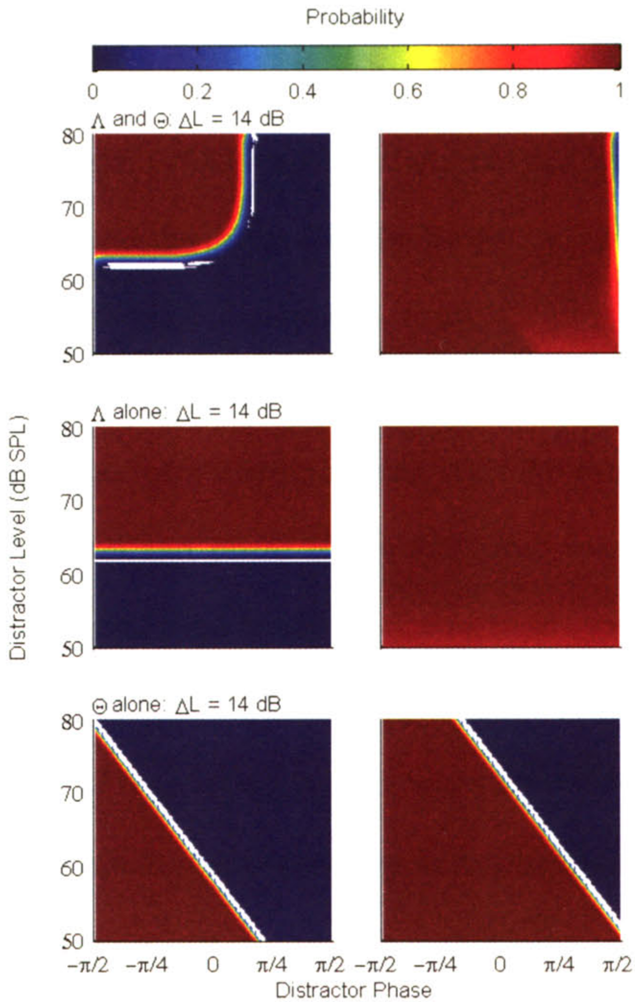

Figure 11. Same as Fig. 10 except with a ΔL of 14 dB.

CHAPTER IV. MODELING MONAURAL LEVEL DISCRIMINATION UNDER DICHOTIC CONDITIONS

I. Introduction

A three-dimensional model based on the perceived loudness and position of the dominant-image and the position of the time-image is used to predict the results of the psychophysical experiment from Chapter III. In that level-discrimination experiment, subjects were asked to judge the level of a target tone at one ear in the presence of a distractor tone with the same frequency at the other ear. When the level and phase of the distractor tone were randomized, discrimination of changes to the target level was much worse than when the level and phase were fixed. That chapter reported the results in terms of the conditional probability of responding that the target level was incremented given it was not, P_F, and the conditional probability of responding that the target level was incremented given it was, P_D, as a function of the distractor level and phase. The P_F and P_D functions were not accurately predicted by a two-dimensional model based on the loudness and position of the dominant-image. A natural extension of that model is to include a third dimension related to the position of the time-image. This chapter investigates the extent to which this three-dimensional model can predict the results. There are no new psychophysical results reported in this chapter.

Figure 1 presents a block diagram of the detection theoretic model based on a non-optimal observer of the observations of the loudness and position of the dominant-image (Λ and Θ, respectively) and the position of the time-image (Υ). The inputs to the model are the level of the left ear L_{Left}, the level of the right ear L_{Right}, and the interaural difference in time delay T. The model does not specify how these inputs might be extracted by the auditory system. The modeled psychophysical experiment uses long duration tonal stimuli and therefore the specifics of the extraction method are of limited importance. The inputs of the model are processed by an internal processor to generate observations of Λ, Θ, and Υ. Again the mechanisms by which the auditory system generates these observations/dimensions is not specified.

The model treats the Λ, Θ, and Υ dimensions as mathematical abstractions. The specifics of the generation of the dimensions are given below in the Model Overview section; note, however, that the dimensions include processing noise. The observations of Λ, Θ, and Υ are used to generate a maximum likelihood indicator function. The model decisions (whether the target level was incremented or not) are based upon a comparison of the observations and a jittered version of the maximum likelihood indicator function.

II. Psychophysical Experiment

This chapter investigates the extent to which a three-dimensional model (cf. Fig. 1 for the block diagram of the model) can predict the psychophysical results reported in Chapter III. Those methods and results are reviewed briefly, highlighting the aspects important for the model. A more general and complete treatment of the methods and results are given in the previous chapter.

A. Methods

The ability of three normal-hearing subjects (S1, S2, and S4)[1] to discriminate changes in the level of a target in the presence of a distractor was assessed with a one-interval, two-alternative-forced-choice (1I, 2AFC) constant-increment paradigm. The target was presented at the left ear and the distractor at the right ear. The target and distractor were presented simultaneously. Both the target and distractor were tones with frequencies of 600 Hz and durations of 300 ms. The phase of the distractor was chosen randomly from a uniform distribution between $-\pi/2$ and $+\pi/2$ and the distractor level (in decibels) was chosen randomly from a uniform distribution between 50 and 80 dB SPL. The phase of the target was fixed at zero and the level of the target was either un-incremented or incremented. The level of the un-incremented target was 50 dB SPL and the level of the incremented target was 58 dB SPL for S1 and S2 and 64 dB SPL for S4. In other words, the target level increment ΔL was 8 dB for S1 and S2 and 14 dB for S4. Subjects S1 and S2 each completed 30,000 trials and subject S4 completed 24,000 trials. All testing was done over headphones in a sound treated booth.

The results were reported as the probability of a false alarm P_F and the probability of a detection P_D as a function of the distractor level and phase. In calculating P_F and P_D, the data were binned according to the distractor level and phase. The bin size was 2 dB by 40 μs and the bins overlapped by 50 percent at each boundary. Although data for a large number of trials were collected, the ranges over which the distractor was roved were also large (30 dB in level and 830 μs) and therefore only a few trials (approximately 40 on average, but in some cases less than 10) were included in each bin.

B. Results

The empirical results to be predicted are shown in Fig. 2 with a separate panel for each subject and for each of the two probabilities (P_F and P_D). Specifically, Fig. 2 shows surface plot representations of P_F (left column) and P_D (right column) as a function of the distractor level and phase. Averages over the distractor stimuli showed (cf. Chapter III) that all the subjects had an average probability of responding correctly P_C of approximately 0.7 and the subjects were unbiased (they responded *"Un-Incremented"* as frequently as they responded *"Incremented"*). Thus, there was little across subject variability when performance was averaged across distractor level and phase. Unlike the average performance, the dependencies of P_F and P_D on the distractor level and phase are different for each subject (different rows). For subjects S1, S2, and S4, the P_F functions depend predominately on distractor phase, distractor level, and a combination of distractor level and phase, respectively. The P_D function depends on both distractor level and phase for subject S1 and predominately on distractor level for S4. For subject S2, the P_D function is nearly independent of both the distractor level and phase.

The correlation coefficient R can be used to quantitatively assess the across subject differences. The root mean squared (RMS) value of R across the subjects (considering both P_F and P_D simultaneously) is 0.70 reflecting the similarities in the average values of P_F and P_D for each subject (generally, P_F is low and P_D is high). The across-subject RMS value of R when the P_F and P_D functions are treated separately is only 0.36, quantifying the across-subject differences in the dependencies of P_F and P_D on the distractor level and phase.

The correlation coefficient can also be used to evaluate the predictions of a simple hypothetical model which only predicts the average P_F and P_D for each subject. This hypothetical model predicts no dependence of P_F and P_D on the distractor level and phase, but has an R value of 0.81. Models which have values of R greater than 0.81 are predicting some of the dependence of P_F and P_D on the distractor level and phase.

III. Modeling

The model is based on a non-ideal observer of a three-dimensional decision space in which the dimensions loosely correspond to the perceived loudness and position of the dominant-image and the position of the time-image, respectively denoted Λ, Θ, and Υ. The model is evaluated on the monaural level discrimination task described above. Specifically the model is used to make predictions of P_F and P_D as a function of the distractor level and phase. The free parameters of the model are adjusted to predict the dependence of P_F and P_D on the distractor level and phase and the substantial individual differences.

A. Model Overview

The Λ, Θ, and Υ dimensions of the model are defined in terms of their dependence on the level of the left ear L_{Left} (in decibels), the level of the right ear L_{Right} (in decibels), the interaural difference in time delay Τ (in seconds), the intensity-time trading ratio k (in decibels per seconds), and three hypothetical internal noises N_Λ, N_Θ, and N_Υ (in decibels). Specifically, the dimensions are defined as

$$\Lambda = 10\log_{10}\left(10^{\frac{L_{Left}}{10}} + 10^{\frac{L_{Right}}{10}}\right) + N_\Lambda \tag{1}$$

$$\Theta = L_{Left} - L_{Right} + k\mathrm{T} + N_\Theta \tag{2}$$

$$\Upsilon = k\mathrm{T} + N_\Upsilon . \tag{3}$$

Note that Λ and Θ are defined the same as in Chapters II and III. Both Λ and Θ have units of decibels and for consistency the position of the time-image Υ is defined in decibels as well. The Λ and Θ dimensions were chosen since they correlate with the dominate perception of the stimuli. The Υ dimension has been proposed by others (Hafter and Carrier 1972; Yost 1972) and also leads to a mathematically tractable solution.[2]

Throughout the modeling, the intensity-time trading k and the internal noises N_{Λ}, N_{Θ}, and N_{Υ} are held fixed. The intensity-time trading ratio k is held fixed at 1 dB per 20 μs (cf. Blauert 1997). The internal noises are zero-mean Gaussian and statistically independent across the dimensions. The standard deviations of N_{Λ}, N_{Θ}, and N_{Υ} (denoted σ_{Λ}, σ_{Θ}, and σ_{Υ}, respectively) are held fixed at 0.5 dB, chosen to be consistent with previous studies on the discrimination of overall level (Viemeister 1988), ILD, ITD, and the time-image (Hafter and Carrier 1972; Yost 1972).

The variables L_{Left}, L_{Right}, and T depend on the target level L_{Target}, the distractor level A, and the distractor phase Φ, respectively. Note that, on any trial, L_{Target}, A, and Φ are randomly chosen. On each trial the model observer responds either "*Un-Incremented*" or "*Incremented*" based on observations of Λ, Θ, and Y. The conditional probabilities of responding *Incremented* (i.e., P_F and P_D) as a function of the distractor level and phase are defined in terms of the conditional probability density for Λ, Θ, and Y, given L_{Target}, A, and Φ and denoted $f_{\Lambda,\Theta,\Upsilon|A,\Phi,L_{Target}}$, and the conditional probability of responding *Incremented* given Λ, Θ, and Y and denoted $P_{"Incremented"|\Lambda,\Theta,\Upsilon}$. Noting that L_{Target} is equal to L_0 (in decibels) when the target level is un-incremented and that L_{Target} is equal to the sum of L_0 and ΔL (both in decibels) when the target level is incremented, P_F and P_D as a function of the distractor level and phase are defined as

$$P_F(a,\phi) = \iiint \begin{bmatrix} P_{\text{"}Incremented\text{"}|\Lambda,\Theta,\Upsilon}\left(\text{"}Incremented\text{"}|\lambda,\theta,\upsilon\right) \\ \times f_{\Lambda,\Theta,\Upsilon|A,\Phi,L_{Target}}\left(\lambda,\theta,\upsilon|a,\phi,L_0\right) \end{bmatrix} d\lambda d\theta d\upsilon,$$

and

$$P_D(a,\phi) = \iiint \begin{bmatrix} P_{\text{"}Incremented\text{"}|\Lambda,\Theta,\Upsilon}\left(\text{"}Incremented\text{"}|\lambda,\theta,\upsilon\right) \\ \times f_{\Lambda,\Theta,\Upsilon|A,\Phi,L_{Target}}\left(\lambda,\theta,\upsilon|a,\phi,L_0+\Delta L\right) \end{bmatrix} d\lambda d\theta d\upsilon. \quad (4)$$

In calculating $P_{\text{"}Incremented\text{"}|\Lambda,\Theta,\Upsilon}$, an indicator function is defined. The maximum likelihood indicator function equals unity (i.e., indicates *Incremented*) when the probability of the observed Λ, Θ, and Y is greater with an incremented target level than with an un-incremented target level. The conditional probability of Λ, Θ, and Y given L_{Target} is denoted $f_{\Lambda,\Theta,\Upsilon|L_{Target}}$. As was the case in the two-dimensional models of the previous chapters, a closed form analytical expression of $f_{\Lambda,\Theta,\Upsilon|L_{Target}}$ was not found. Rather, a mixture of analytic and numerical techniques is again used to approximate $f_{\Lambda,\Theta,\Upsilon|L_{Target}}$. A detailed derivation of $f_{\Lambda,\Theta,\Upsilon|L_{Target}}$ is presented in Appendix A. Note that in approximating $f_{\Lambda,\Theta,\Upsilon|L_{Target}}$ no Monte-Carlo type simulations were used; the approximations were limited to using numerical techniques to solve for the integrals of continuous functions.

The specifics of the maximum likelihood indicator function are presented below in the Maximum Likelihood Indicator Function section. In general terms, the maximum likelihood indicator function divides the $\{\Lambda,\Theta,\Upsilon\}$ space into regions of indicate-*Un-Incremented* and indicate-*Incremented*. For this model, the resulting indicator functions are such that for each value of $\{\Theta, \Upsilon\}$, there is only one *Un-Incremented/Incremented* boundary value of Λ; the boundary can be represented by a single valued function of Θ and Y, denoted $\psi_{\Lambda,\Theta,\Upsilon}$. The indicator function indicates *Un-Incremented* for values of Λ less than $\psi_{\Lambda,\Theta,\Upsilon}$ and *Incremented* for values of Λ greater than $\psi_{\Lambda,\Theta,\Upsilon}$. In other words, for a given Θ and Y the ideal observer responds *Incremented* whenever the Λ (the loudness) is greater than some threshold value; this threshold value is a function of Θ and Y, but not Λ.

Estimates of the stimulus attributes (L_{Left}, L_{Right}, and T) from the Λ, Θ, and Y dimensions are only limited by the internal noises N_Λ, N_Θ, and N_Υ. One could represent any set of N-dimensions based on combinations (linear and non-linear) of Λ, Θ, and Y. For example, transforming the Y dimension (or the Λ or Θ dimension) into a monaural energy dimension has no effect on the predictions of the ideal observer. Similarly, one could transform the Y dimension (or again the Λ or Θ dimension) into dimensions which correspond to the spatial width or interaural level difference and again have no effect. In regards to the psychophysical experiment, the ideal observer of Λ, Θ, and Y performs nearly perfectly (P_F of zero and P_D of unity) with the values of ΔL used in the experiment (8 and 14 dB) and the chosen values of σ_Λ, σ_Θ, and σ_Υ.

In preliminary investigations, three methods of reducing the performance of the ideal observer were explored. The first is a jittered decision boundary (this is the method used in the reported results), the second is related to divided attention (modeled as an increased cost for attending to multiple dimensions), and the third is *ad hoc* manipulations of the decision boundary. One *ad hoc* manipulations considered was for the indicator function to be unity (indicate *Incremented*) when the loudness (Λ) was in excess of some criterion level independent of Θ and Υ and the maximum likelihood indicator functions. Similar manipulations were also considered for Θ and Υ. Combining *ad hoc* manipulations of the decision boundary with either the divided attention observer or the jittered observer, increased the predictive power of the model only slightly. The predicative power of the jittered observer and the divided attention observer were approximately equal. The added predictive power of combining the two observers was minimal.

The performance of the ideal observer was limited by jittering the maximum likelihood indicator function. The jittered observer was chosen since the maximum likelihood indicator function is complex (cf. the Maximum Likelihood Indicator Function section) and conceptually the authors believe that subjects could have problems implementing such a complex indicator function. The divided attention observer was not included since Bonnel and Hafter (1998) reported minimal costs for attending to multiple auditory dimensions simultaneously. Further, the increase in the predictive power was minimal given the increase in the number of free parameters. The increases in the predicative power from *ad hoc* manipulations was also insufficient to warrant their inclusion in the reported model.

The derivation of $P_{"Incremented"|\Lambda,\Theta,\Upsilon}$ and $f_{\Lambda,\Theta,\Upsilon|A,\Phi,L_{Target}}$, which are used to calculate P_F and P_D, for the jittered observer are presented in Appendix B. Closed-form analytic expressions of $P_{"Incremented"|\Lambda,\Theta,\Upsilon}$ and $f_{\Lambda,\Theta,\Upsilon|A,\Phi,L_{Target}}$ were not found; rather, they were approximated with a mixture of analytical and numerical techniques. As was the case with the $f_{\Lambda,\Theta,\Upsilon|L_{Target}}$, the approximations were limited to using numerical techniques to solve for integrals of continuous functions.

B. Maximum Likelihood Indicator Function

The maximum likelihood indicator function divides the $\{\Lambda,\Theta,\Upsilon\}$ space into regions of indicate-"*Un-Incremented*" and indicate-"*Incremented*". Although, the decision boundary is only an intermediate stage in the calculation of P_F and P_D (i.e., the model predictions) understanding the boundary is useful for understanding the predictions. The maximum likelihood indicator function is defined in terms of a log-likelihood ratio $\eta_{\Lambda,\Theta,\Upsilon}$ which depends on $f_{\Lambda,\Theta,\Upsilon|L_{Target}}$. Specifically $\eta_{\Lambda,\Theta,\Upsilon}$ is defined as

$$\eta_{\Lambda,\Theta,\Upsilon}(\Lambda,\Theta,\Upsilon,\Delta L)=10\log_{10}\left(\frac{f_{\Lambda,\Theta,\Upsilon|L_{Target}}(\lambda,\theta,\upsilon|L_{Target}=L_0)}{f_{\Lambda,\Theta,\Upsilon|L_{Target}}(\lambda,\theta,\upsilon|L_{Target}=L_0+\Delta L)}\right).$$

The indicator function indicates *Incremented* when $\eta_{\Lambda,\Theta,\Upsilon}$ is less than zero. In this model, the indicator function can be characterized by a decision boundary $\Psi_{\Lambda,\Theta,\Upsilon}$ such that the indicator function indicates *Incremented* when the observed Λ is greater than $\Psi_{\Lambda,\Theta,\Upsilon}$ evaluated at the observed Θ and Υ. Note that $\eta_{\Lambda,\Theta,\Upsilon}$ and hence $\Psi_{\Lambda,\Theta,\Upsilon}$ depend on ΔL.

Figure 3 shows $\Psi_{\Lambda,\Theta,\Upsilon}$ for the values of ΔL of 8 dB (top panel) and 14 dB (bottom panel). In each panel there are seven lines corresponding to $\Psi_{\Lambda,\Theta,\Upsilon}$ evaluated at values Υ equal to -15, -10, -5, 0, 5, 10, and 15 dB (the expected value of Υ with a 300 μs ITD is 15 dB). The boundary function $\Psi_{\Lambda,\Theta,\Upsilon}$ systematically shifts with Υ. The line corresponding to $\Psi_{\Lambda,\Theta,\Upsilon}$ evaluated at the largest value of Υ is the upper-right most line, while the line corresponding to $\Psi_{\Lambda,\Theta,\Upsilon}$ evaluated at the smallest (most negative) value of Υ is the lower-left most line. For a given Υ, the maximum likelihood indicator function indicates *Incremented* for values of $\{\Lambda,\Theta\}$ which fall above and to the right of the appropriate line. In the three-dimensional $\{\Lambda,\Theta,\Upsilon\}$ space, the boundary is a smooth surface. This surface is shifted by changes in ΔL; the boundary function $\Psi_{\Lambda,\Theta,\Upsilon}$ for a ΔL of 8 dB is approximately a shifted version (down and left) of $\Psi_{\Lambda,\Theta,\Upsilon}$ for a ΔL of 14 dB.

C. Non-Ideal Observer

An ideal observer of the three dimensional decision space is only limited by the internal noises N_{Λ}, N_{Θ} and N_{Υ}. A non-ideal observer is used to further limit discrimination performance. We have chosen a non-ideal observer that bases its decisions on observations of Λ, Θ, and Υ (denoted λ, θ, and υ, respectively) relative to a jittered version of the maximum likelihood indicator function (cf. Fig 4). The non-ideal model observer responds *Incremented* when λ is greater than $\Psi_{\Lambda,\Theta,\Upsilon}(\theta+J_{\Theta},\upsilon+J_{\Upsilon})+J_{\Lambda}$, where J_{Λ}, J_{Θ}, and J_{Υ} are the criterion jitters. The criterion jitters are assumed to be zero-mean Gaussian random variables which are independent across the dimensions. The standard deviation of the jitters J_{Λ}, J_{Θ}, and J_{Υ} are denoted $\sigma_{J_{\Lambda}}$, $\sigma_{J_{\Theta}}$, and $\sigma_{J_{\Upsilon}}$. The standard deviation of the jitters were treated as free parameters and adjusted for each subject independently to predicted P_F and P_D.

In fitting the model, $\sigma_{J_{\Lambda}}$, $\sigma_{J_{\Theta}}$, and $\sigma_{J_{\Upsilon}}$ were allowed to vary between zero and infinity. Setting all three jitter standard deviations equal to infinity results in chance performance. When all three jitter standard deviations are equal to zero, the non-ideal observer is the ideal (maximum likelihood) observer. Setting the jitter standard deviation of one dimension to infinity and the other dimensions to zero, does not simply reduce the model to a two-dimensional model.

Figure 4 shows the maximum likelihood decision boundaries for the $\{\Lambda,\Theta,\Upsilon\}$ space (top panel) and for the $\{\Lambda,\Theta\}$ space (bottom panel) with a ΔL of 8 dB. The decision boundary for the $\{\Lambda, \Theta, \Upsilon\}$ space, denoted $\Psi_{\Lambda,\Theta,\Upsilon}$, shown in Fig. 4 is the same the one shown in the top panel of Fig. 3. The decision boundary for the $\{\Lambda, \Theta\}$ space, denoted $\Psi_{\Lambda,\Theta}$ was derived in Chapter III. It can also be derived from $f_{\Lambda,\Theta,\Upsilon|L_{Target}}$ by integrating over all Y to get $f_{\Lambda,\Theta|L_{Target}}$ and creating a log-likelihood function. It does not appear that the boundary $\Psi_{\Lambda,\Theta}$ can be obtained by a simple combination of $\Psi_{\Lambda,\Theta,\Upsilon}$ evaluated at the different values of Y.

The criterion jitter is fundamentally different than the internal noise incorporated into the indicator function. Internal noise results in variable observations given a fixed stimulus. Criterion jitter results in variable responses given a fixed observation. In the uni-dimensional Gaussian case internal noise and criterion jitter are mathematically inseparable. In this model, as discussed above, there appears to be a mathematical difference. Appendix C proves that there is a mathematical difference between internal noise and criterion jitter for the ideal observer of Λ alone.

A comparison between the predictions of a model with only coding noise and one with both coding and decision noise is not made. Intuitively it seems reasonable that the same stimulus does not always need to give rise identical observations (supporting the inclusion of internal noise). It also seems reasonable that given the same observations that subjects may not react identically; criterion jitter incorporates this second effect. By including decision noise which differs across the subjects, some of the inter-subject variability can be captured. It seems reasonable that given the same observations that subjects would make different decisions. It is not clear if it is valid to assume that the coding noise varies across the subjects and therefore the coding noise is held fixed.

D. Data Analysis

The predictions are evaluated with three statistical metrics: the (1) mean difference, (2) correlation coefficient R, and (3) root of the mean squared (RMS) difference between the psychophysical measurements and the predictions. In calculating the statistics no regard was given to the confidence of the estimates of the P_F and P_D at each distractor level and phase (the confidence depends on the conditional probability and the number of trials). The mean difference quantifies the average difference between the measured and predicted data (P_F and P_D) without regard to the dependencies on the distractor level and phase. The correlation coefficient quantifies the similarity of the dependencies on the distractor level and phase between the measured and predicted data without regard to the average difference. The RMS difference weights both the average difference between the measured and predicted data and the dependencies on the distractor level and phase. Given the complexity of the measured and predicted data functions, no single statistic captures all the aspects of the data.

Large mean differences would indicate a fundamental flaw in the model. A model which on average (across distractor level and phase) predicts a probability of a correct response P_C or a bias that is substantially different than the subjects has limited predictive power. A model which predicts the average

P_C and bias (or equivalently the average P_F and P_D) for each subject has substantial predictive power; as mentioned in the presentation of the psychophysical results such a model has an R value of 0.81 (it predicts 2/3 of the variance of the data). An even better model would not only predict the average P_F and P_D for each subject, but would also predict some the dependence of P_F and P_D on the distractor level and phase.

In evaluating the model predictions, recall that although data for a large number of trials were collected for each subject, P_F and P_D for each distractor level and phase bin was based on only a small number of trials. Each estimate of P_F and P_D has a different reliability due to the underlying trial-to-trial variability (Bernoulli). Davidson et al. (2006) described how to calculate this percentage of the variance, expanding upon that we state that the maximum correlation coefficient R_{max} is

$$R_{max} = \sqrt{1 - \frac{E_{a,\phi}\left[\dfrac{P_D(a,\phi)(1-P_D(a,\phi))}{n_{inc,inc}(a,\phi)}\right] + E_{a,\phi}\left[\dfrac{P_F(a,\phi)(1-P_F(a,\phi))}{n_{inc,un-inc}(a,\phi)}\right]}{\sigma_{Tot}{}^2}}$$

where $\sigma_{Tot}{}^2$ is the total variance of the psychophysical data, $n_{inc,inc}(a,\phi)$ is the number of trial for which the response was *"Incremented"* and the target level was incremented and $n_{inc,un-inc}(a,\phi)$ is the number of trial for which the response was *"Incremented"* and the target level was un-incremented. Given the data from the psychophysical experiment R_{max} is 0.97.

E. Model Implementation

The free parameters (σ_{J_Λ}, σ_{J_Θ}, and σ_{J_Υ}) of the model were adjusted to minimize the RMS difference between the measured and predicted data. The fitting was done for each subject independently. Note that minimizing the RMS difference does not minimize the mean difference. The mean difference with the parameters that minimized the RMS difference, however, was small enough that it was not necessary to consider more complex fitting algorithms.

The reported predictions are for the parameter values returned by the **fminsearch** function included in Matlab. Due to the computational complexity of the model, it is not possible to be assured that the reported values of the RMS difference are truly the minimum values. Further, the parameter search was performed using a poorly sampled version of the indicator function (1 dB step size). The final model predictions were made with a finely sampled indicator function (0.1 dB step size). The RMS difference in the predictions of P_F and P_D with the poorly and finely sampled indicator function was 0.17. Using a more finely sampled indicator function (e.g., 0.05 dB) did not substantially change the predictions (i.e., RMS difference in the predicted P_F and P_D with the 0.1 and 0.05 dB sampling was negligible).

The minimization routine was run on a single processor computer and took approximately 8 hours of computing time. Once the "best fitting" parameters were obtained with the minimization routine, the values of P_F and P_D, were obtained with the finely sampled version of the indicator function. The predictions were obtained on a 54 processor super-computer and also took approximately 8 hours of time. Conducting a more extensive parameter search with the finely sampled indicator function might lead to a different set of parameters which further reduce the RMS difference between the measured and predicted data.

In the preliminary analysis parameter values which maximized R were also found. The parameter values which maximize R are substantially different than those which minimize the RMS difference. The R value obtained with the parameter values which maximize R is much greater than the R value obtained for the parameter values which minimize the RMS difference, but the change in the RMS difference is small. The improved R value comes at the cost of large errors in the mean differences. Since the reported predictions are for the parameter values which minimize the RMS difference, the reported values of the mean difference and R are not the best values which can be obtained. In our opinion, the values of the parameters which minimize the RMS difference (or maximized R) do not give the best visual fit between the measurements and the predictions. The quality of the visual fit, however, is highly subjective and was not considered in adjusting the parameters.

F. Results

Figure 5 shows the model predictions of P_F (left column) and P_D (right column) for subjects S1 (top row), S2 (middle row), and S4 (bottom row). As with the psychophysical results (cf. Fig. 2), the predicted P_F and P_D are functions of the distractor level and phase. The dependence of the predicted P_F and P_D on the distractor level and phase varies across the subjects. The prediction of P_F for subject S1 depends predominately on the distractor phase. The predictions of P_F for subjects S2 and S4 depend predominately on the distractor level. For subject S2, the predicted P_F increases with increasing distractor level, but for subject S4, the predicted P_F decreases with increasing distractor level. The predicted P_D functions predominately depend on level and phase (S1), and level alone (S2 and S4). For subject S1, the predicted P_D is minimal for positive distractor phases and high distractor levels. For subjects S2 and S4, the predicted P_D decreases with increasing distractor level. For a given distractor level and phase, the predicted P_D is higher than the predicted P_F.

The summary statistics comparing the predictions and the data are presented in Table 1. The standard deviation of the criterion jitters (σ_{J_Λ}, σ_{J_Θ}, and σ_{J_Υ}) used for the predictions are substantially different across the subjects. Given the differences in the P_F and P_D functions, differences in the criterion jitters are not surprising. Subject S1 accurately places the decision boundary in the Λ and Θ dimensions and inaccurately in the Y; the subject uses the information in the loudness and position of the dominant-image more than the information in the position of the time-image. Subject S2 more equally weights the Λ, Θ, and Y dimensions, but puts the least importance on the position of the dominant-image. Subject S4 weights the position of the time-image the most and the loudness the least.

The relative weightings of the dimensions in the model quantify how accurately the decision boundary is located. The relative weightings do not quantify on how accurately the subjects could rate the loudness and position of the dominant-image and the position of the time-image. For example, although not directly measured, but based on the data reported in Chapter II it is probable that, all the subjects can discriminate changes in the loudness and position of the dominant-image and the position of the time-image equally well even though the subjects weighted the dimensions differentially. It is hypothesized that the differences across the subjects arise due to the multidimensional nature (little information is contained in any dimension individually) of the modeled task.

The mean differences between the predictions and the data in P_C and the bias are small for all the subjects. The model is accurately predicting the average performance of each subject. The average performance, however, does not vary substantially across the subjects. Subject S1 had the lowest value of P_C of 0.69 and subject S2 and S4 had values of P_C of 0.72. The minimum bias was 0.02 (S4) and the maximum bias was 0.07 (S1).

The R value between the data and the predictions is 0.87. This R value is less than the R_{max} value of 0.97, but is higher than 0.81 which is the R value for a model which predicts the average P_F and P_D independent of the distractor level and phase. Therefore the model is predicting some (but not all) of the dependence of P_F and P_D on the distractor level and phase. The ability to predict the fine structure of the data is quantified by considering the R value for each subject for P_F alone and P_D alone. The model better predicts the distractor level and phase dependence of P_F than that of P_D. The model does similarly well predicting P_F for subjects S1 and S2 with R values of 0.73 and 0.74, and slightly worse for S4 with an R value of only 0.54. The R values for P_D are much more variable across the subjects. The predictions are best for subject S1 with an R value of 0.64 and worse for subject S2 with an R value of 0.10. Note that the extent of variability with respect to the distractor level and phase in the P_D function for subject S2 is much less than the variability for subject S1. The low R value is a result of this lack of dependence (there is essentially no dependence to predict).

Although the correlations between the model and the data are relatively high and the model predicts some of the dependence of P_F and P_D on the distractor level, there are systematic discrepancies between the model and the data. Figure 6 shows the differences between the measured and predicted P_F (left column) and P_D (right column) functions for each subject (different rows). The distribution of the errors with an ideal model would be randomly scattered. The magnitude of the errors of the ideal model would depend on the measured

P_F and P_D (probabilities near zero and unity have higher confidences than probabilities near 0.5). The actual differences, although generally small (cf. Table 1) are highly clustered. This clustering is to be expected since the predicted dependence, on distractor level and phase, of P_F and P_D is less than the measured dependence.

The dependence of the errors on the distractor level and phase vary across the subjects. In general, the magnitude of the transitions in the predicted probabilities (in both P_F and P_D) are less than the magnitude of the measured transitions. When the measured probability is high the predicted probability is not high enough and when the measured probability is low, the predicted probability is not low enough. The failure of the model to accurately predict the transitions is consistent across the subjects, even though the dependence of P_F and P_D varies across the subjects.

IV. Discussion

One of the most interesting aspects of the psychophysical results is the finding that all three subjects performed similarly on average, but seem to have used different strategies. In the previous chapter it was shown that the magnitude of the difference between the psychophysical data and the predictions of the average probability of a correct response for the ideal observers of Λ alone, Θ alone and Λ and Θ together were as large as 0.2. Further, the predicted dependence of the P_F and P_D functions on the distractor level and phase was in disagreement with the data. By introducing a third dimension (Y) and considering a non-ideal observer, the revised model predicts the average psychophysical performance better (mean differences less than 0.05 in both P_C and bias). The P_F and P_D functions, however, are not fully predicted. The RMS difference between the measured data and predictions for each P_F and P_D function is similar (between 0.11 and 0.14). The R values, however, are highly variable with a maximum of 0.73 and a minimum of 0.10. The minimum R value (worst "fit") occurs for the same case (subject S2 P_D) for which the minimum RMS difference (best "fit") occurs. This discrepancy highlights the danger in assessing the quality of the predictions based on a single statistic.

In general the model correctly predicts the trends in the P_F and P_D functions. The predictions, however, do not show as strong a contrast between areas of high and low probability as do the psychophysical data. For subject S1, both the measured and predicted P_F depend predominately on the distractor phase. The measured P_F functions almost spans the entire ranges between zero and unity (between 0.0 and 0.94), but the predicted function varies only between 0.15 and 0.69. Similarly for subject S1, both the measured and predicted P_F depend on the distractor level and phase, and again the measured function varies between 0.22 and 1.0 while the predicted function varies only between 0.60 and 1.0. Although more difficult to assess visually, this same trend is also observed for subjects S2 and S4.

The model makes different predictions for the different subjects because the parameters are adjusted for each subject individually. Given the complexity of the task, it is not unreasonable to assume that the subjects would implement different listening and response strategies. The framework of the model allows for the effects of divided attention and criterion jitter to be analyzed separately. Only the effects of criterion jitter on the predictions were formally evaluated.

This decision to only consider the effects of criterion jitter was based on three reasons. The first reason was the increase in computational time needed to account for divided attention; even when not accounting for divided attention, computing P_F and P_D required considerable computer time and computing power due to the range over which the distractor level and phase varied relative to the sensitivity of normal-hearing subjects. The second reason was that the preliminary attempts at modeling both divided attention and criterion jitter did not substantially increase the predictive power of the model. The third reason was that studies of divided attention in auditory tasks (e.g., Bonnel and Hafter 1998) have reported limited effects of attending to multiple dimensions.

The failure of the model to predict all of the dependence on the distractor level and phase may be related to the simplifying assumptions incorporated into the model. As will be discussed below, some of the assumptions are in disagreement with psychophysical results. Most of the assumptions are not unique to our model. For example the assumption that the internal noise in the Θ

dimension is independent of the ILD and ITD disagrees with the psychophysical results of Domnitz (1973) and Domnitz and Colburn (1973) who demonstrated a small effect. The perceived loudness, position, and time-image are obviously more complicated than the assumed model dimensions. It is also assumed that subjects can generate and utilize a maximum likelihood indicator function. Finally it is assumed that subjects jitter the maximum likelihood indicator function in a Gaussian manner. Given the number of simplifying assumptions, it is not surprising that the model fails to predict all of the dependence on the distractor level and phase.

Reducing the number of simplifying assumptions could increase the amount of the variance in the data predicted by the model. The current model predicts 76% of the variance in the data. Given the intrinsic variability in the data attributable to the underlying Bernoulli trials, only 94% of the variance in the data is predictable. The changes to the model needed to account for this remaining 18% are not obvious. Some changes (e.g., changes to the definition to the dimensions or an increased dimensionality) will increase the computational complexity of the model (quite possibly to a level for which predictions cannot be made). Other changes will not increase the computational complexity of the model (e.g., *ad hoc* manipulations to the indicator function), but have little support to justify their inclusion. Overall a model based on the loudness, position, and time-image can predict the average performance in a monaural level discrimination task under dichotic conditions. Further the model can predict some of the individual differences in the dependencies of the P_F and P_D functions on the distractor level and phase.

V. Summary

Applying a sub-optimal detection theoretic model based on observations of the loudness and position of the dominant-image and the position of the time-image gives insight into the decision process of subjects under conditions in which contra-aural interference is observed. In the modeled psychophysical experiment subjects are unable to attend exclusively to the ear at which the target is presented. This model demonstrates, that although the subjects are not attending exclusively to the target ear, they are behaving reasonably. The relatively high correlation between the model and data suggests that subjects are making non-optimal use of the modeled dimensions. By allowing different relative weightings of the three dimensions the model predicts a considerable portion of the inter-subject variability.

Acknowledgements

This research was supported by NIH/NIDCD Grant DC00100.

Appendix A

In this appendix, analytical and numerical techniques are used to approximate the conditional joint probability density function of Λ, Θ, and Υ, as defined in Eqs. 1, 2, and 3, given a target level L_{Target} equal to the sum of the reference level L_0 and an increment ΔL. The conditional joint probability density function is denoted as $f_{\Lambda,\Theta,\Upsilon|L_{Target}}$. The derivation of $f_{\Lambda,\Theta,\Upsilon|L_{Target}}$ is similar to the derivation of the two dimensional conditional joint probability density function of Λ and Θ given L_{Target} derived in Chapter II.

Before the derivation of $f_{\Lambda,\Theta,\Upsilon|L_{Target}}$ begins, Eqs. 1, 2, and 3 are restated in terms of the variables of the psychophysical experiment. Specifically, L_{Left} is equal to L_{Target} and L_{Right} is equal to the distractor level expressed as a sum of L_0 and a random variable A. The negative of the distractor phase Φ divided by the radian frequency ω is equal to T. Making these substitutions into equations 1, 2, and 3 results in

$$\Lambda = 10\log_{10}\left(10^{\frac{L_{Target}}{10}} + 10^{\frac{L_0+A}{10}}\right) + N_\Lambda, \tag{A1}$$

$$\Theta = L_{Target} - \left(L_0 + A\right) - \frac{k}{\omega}\Phi + N_\Theta, \tag{A2}$$

and

$$\Upsilon = -\frac{k}{\omega}\Phi + N_\Upsilon. \tag{A3}$$

Note that to match the psychophysical experiment the random variables A and Φ are uniformly distributed between a_{min} and a_{max} and ϕ_{min} and ϕ_{max}, respectively.

The derivation of $f_{\Lambda,\Theta,\Upsilon|L_{Target}}$ begins by using the definition of conditional probability to expand $f_{\Lambda,\Theta,\Upsilon|L_{Target}}$. With this expansion and noting the independence of Y on Λ and L_{Target}, $f_{\Lambda,\Theta,\Upsilon|L_{Target}}$ can be rewritten as

$$f_{\Lambda,\Theta,\Upsilon|L_{Target}} = f_{\Lambda|L_{Target}}\left(\lambda|L_0+\Delta L\right) f_{\Upsilon}(\upsilon) f_{\Theta|\Upsilon,\Lambda,L_{Target}}\left(\theta|\upsilon,\lambda,L_0+\Delta L\right). \qquad \text{(A4)}$$

The $f_{\Theta|\Upsilon,\Lambda,L_{Target}}$ term is then rewritten in terms of its marginal distributions with respect to A and Φ as

$$f_{\Theta|\Upsilon,\Lambda,L_{Target}} = \iint f_{\Theta|\Lambda,\Upsilon,A,\Phi,L_{Target}}\left(\theta|\lambda,\upsilon,a,\phi,L_0+\Delta L\right) f_{A,\Phi|\Lambda,\Upsilon,L_{Target}}\left(a,\phi|\lambda,\upsilon,L_0+\Delta L\right) d\phi da.$$

Making a substitution based on Eq. A2 into $f_{\Theta|\Lambda,\Upsilon,A,\Phi,L_{Target}}$ gives

$$f_{\Theta|\Upsilon,\Lambda,L_{Target}} = \iint f_{N_\Theta}\left(x_\Theta(a,\phi)\right) f_{A,\Phi|\Lambda,\Upsilon,L_{Target}}\left(a,\phi|\lambda,\upsilon,L_0+\Delta L\right) d\phi da$$

where $x_\Theta(a,\phi)$ is equal to $\frac{k}{\omega}\phi-\Delta L+a+\theta$. Then using the definition of conditional probability to expand $f_{A,\Phi|\Lambda,\Upsilon,L}$, noting that $f_{\Phi|A,\Lambda,\Upsilon,L_{Target}}$ is independent of A, Λ, and L_{Target} and $f_{A|\Lambda,\Upsilon,L_{Target}}$ is independent of Υ, and simplifying gives

$$f_{\Theta|\Upsilon,\Lambda,L_{Target}} = \iint f_{N_\Theta}\left(x_\Theta(a,\phi)\right) f_{A|\Lambda,L_{Target}}\left(a|\lambda,L_0+\Delta L\right) f_{\Phi|\Upsilon}\left(\phi|\upsilon\right) d\phi da.$$

Using the definition of conditional probability on $f_{A|\Lambda,\Upsilon,L_{Target}}$ and $f_{\Phi|\Upsilon}$, then noting the independence of $f_{A|L_{Target}}$ on L_{Target}, and simplifying gives

$$f_{\Theta|\Upsilon,\Lambda,L_{Target}} = \iint f_{N_\Theta}\left(x_\Theta(a,\phi)\right) \frac{f_A(a) f_{\Lambda|A,L_{Target}}\left(\lambda|a,L_0+\Delta L\right)}{f_{\Lambda|L_{Target}}\left(\lambda|L_0+\Delta L\right)} \frac{f_{\Upsilon|\Phi}\left(\upsilon|\phi\right) f_\Phi(\phi)}{f_\Upsilon(\upsilon)} d\phi da$$

Substituting $f_{\Theta|\Upsilon,\Lambda,L_{Target}}$ into Eq. A4 and simplifying gives

$$f_{\Lambda,\Theta,\Upsilon|L_{Target}} = \int f_A(a) f_{\Lambda|A,L_{Target}}\left(\lambda|a,L_0+\Delta L\right) \int f_{N_\Theta}\left(x_\Theta(a,\phi)\right) f_{\Upsilon|\Phi}\left(\upsilon|\phi\right) f_\Phi(\phi) d\phi da.$$

Making use of the uniform probability density functions of A and Φ, $f_{\Lambda,\Theta,\Upsilon|L_{Target}}$ can be rewritten as

$$f_{\Lambda,\Theta,\Upsilon|L_{Target}} = \kappa \int_{a_{min}}^{a_{max}} f_{\Lambda|A,L_{Target}}\left(\lambda|a,L_0+\Delta L\right) \int_{\phi_{min}}^{\phi_{max}} f_{N_\Theta}\left(x_\Theta(a,\phi)\right) f_{\Upsilon|\Phi}\left(\upsilon|\phi\right) d\phi da,$$

where κ is equal to $1 \big/ (a_{max}-a_{min})(\phi_{max}-\phi_{min})$. Making a substitution based on Eq. A3 into $f_{\Upsilon|\Phi}$ results in

$$f_{\Lambda,\Theta,\Upsilon|L_{Target}} = \kappa \int_{a_{min}}^{a_{max}} f_{\Lambda|A,L_{Target}}\left(\lambda|a,L_0+\Delta L\right) \int_{\phi_{min}}^{\phi_{max}} f_{N_\Theta}\left(x_\Theta(a,\phi)\right) f_{N_\Upsilon}\left(x_\Upsilon(\phi)\right) d\phi da,$$

where $x_\Upsilon(\phi)$ is equal to $\upsilon + \frac{k}{\varpi}\phi$. Substituting the Gaussian density functions of N_Θ and N_Υ, combining, and simplifying gives

$$f_{\Lambda,\Theta,\Upsilon|L_{Target}} = \kappa \int_{a_{min}}^{a_{max}} \mathrm{H}(a,\theta,\upsilon) f_{\Lambda|A,L_{Target}}\left(\lambda|a,L_0+\Delta L\right) \int_{\phi_{min}}^{\phi_{max}} e^{\frac{-(\phi-\mu(a,\theta,\upsilon))^2}{2\sigma^2}} d\phi da,$$

where the constant σ and the functions $\mathrm{H}(a,\theta,\upsilon)$ and $\mu(a,\theta,\upsilon)$ are defined in Appendix D. Defining the integral of an exponential squared function as

$$\mathrm{G}(\alpha,\beta,\mu,\sigma) = \int_{\alpha}^{\beta} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx, \qquad \text{(A5)}$$

allows for $f_{\Lambda,\Theta,\Upsilon|L_{Target}}$ to be simplified as

$$f_{\Lambda,\Theta,\Upsilon|L_{Target}} = \kappa \int_{a_{min}}^{a_{max}} \mathrm{H}(a,\theta,\upsilon) f_{\Lambda|A,L_{Target}}\left(\lambda|a,L_0+\Delta L\right) \mathrm{G}\left(\phi_{min},\phi_{max},\mu(a,\theta,\upsilon),\sigma\right) da.$$

Finally a substitution into $f_{\Lambda|A,L_{Target}}$ based on Eq. A1 is made to get

$$f_{\Lambda,\Theta,\Upsilon|L_{Target}} = \frac{K}{\sqrt{2\pi\sigma_\Lambda^2}} \int_{a_{min}}^{a_{max}} \mathrm{H}(a,\theta,\upsilon) e^{-\left(\frac{x_\Lambda(\lambda,a)^2}{2\sigma_\Lambda^2}\right)} \mathrm{G}(\phi_{min},\phi_{max},\mu(a,\theta,\upsilon),\sigma)\, da,$$

where $x_\Lambda(\lambda,a)$ is equal $\lambda - 10\log_{10}\left(10^{\frac{L_0+\Delta L}{10}} + 10^{\frac{L_0+a}{10}}\right)$.

Further analytical manipulations of $f_{\Lambda,\Theta,\Upsilon|L_{Target}}$ do not appear to reduce the complexity of the solution, but at this stage, $f_{\Lambda,\Theta,\Upsilon|L_{Target}}$ can be approximated with numerical methods. The first step is to approximate the definite integral with summation. Let us denote $a[n]$ as a sampled version of the continuous random variable A. Further let $a[1]$ equal a_{min} and $a[N]$ equal a_{max}. The probability density functions $f_{\Lambda,\Theta,\Upsilon|L_{Target}}$ can then be numerically approximated as

$$f_{\Lambda,\Theta,\Upsilon|L_{Target}} \approx \frac{1}{\sqrt{2\pi\sigma_\Lambda^2}} \frac{1}{N(\phi_{max}-\phi_{min})} \times \sum_{n=1}^{N} \mathrm{H}(a[n],\theta,\upsilon) e^{-\left(\frac{x_\Lambda(\lambda,a[n])^2}{2\sigma_\Lambda^2}\right)} \mathrm{G}(\phi_{min},\phi_{max},\mu(a[n],\theta,\upsilon),\sigma).$$

Appendix B

This appendix contains a derivation of the conditional probabilities of responding *Incremented* given an incremented target level (i.e., P_D) as a function of the distractor level and phase. This appendix only outlines the analytical and numerical techniques used in approximating P_D, but the derivation of P_F is extremely similar. Decisions are based upon the observation of Λ, Θ, and Υ, as defined in Eqs. A1, A2, and A3, respectively. As defined in Eq. 4, P_D depends on two conditional probabilities. The first is the conditional probability of Λ, Θ, and Υ given the target level L_{Target}, the distractor level increment *A*, and the distractor phase Φ and is denoted $f_{\Lambda,\Theta,\Upsilon|A,\Phi,L_{Target}}$. The second is the conditional probability of responding *Incremented* given Λ, Θ, and Υ and is denoted $P_{\text{"}Incremented\text{"}|\Lambda,\Theta,\Upsilon}$.

The conditional probability $f_{\Lambda,\Theta,\Upsilon|A,\Phi,L_{Target}}$ can be expanded by iteratively using the definition of conditional probability such that

$$\begin{aligned} f_{\Lambda,\Theta,\Upsilon|A,\Phi,L_{Target}} &= f_{\Lambda|A,\Phi,L_{Target}}\left(\lambda|a,\phi,L_0+\Delta L\right) \\ &\times f_{\Upsilon|\Lambda,A,\Phi,L_{Target}}\left(\upsilon|\lambda,a,\phi,L_0+\Delta L\right) f_{\Theta|\Upsilon,\Lambda,A,\Phi,L_{Target}}\left(\theta|\upsilon,\lambda,a,\phi,L_0+\Delta L\right) \end{aligned}.$$

Noting that (1) Λ is independent of Φ, (2) Υ is independent of *A*, Λ and L_{Target}, and (3) Θ is conditionally independent of Υ and Λ when *A* and Φ are given (i.e., $f_{\Theta|\Upsilon,\Lambda,A,\Phi,L_{Target}}$ is equal to $f_{\Theta|A,\Phi,L_{Target}}$), allows us to simplify $f_{\Lambda,\Theta,\Upsilon|A,\Phi,L_{Target}}$ to

$$f_{\Lambda,\Theta,\Upsilon|A,\Phi,L_{Target}} = f_{\Lambda|A,L_{Target}}\left(\lambda|a,L_0+\Delta L\right) f_{\Upsilon|\Phi}\left(\upsilon|\phi\right) f_{\Theta|A,\Phi,L_{Target}}\left(\theta|a,\phi,L_0+\Delta L\right).$$

Making substitutions based on Eqs. A1, A2, and A3 and then substituting the Gaussian density functions of N_Λ, N_Θ and N_Υ and simplifying results in

$$f_{\Lambda,\Theta,\Upsilon|A,\Phi,L_{Target}} = \gamma e^{\frac{(\lambda-\mu_\Lambda(a))^2}{2\sigma_\Lambda{}^2}} e^{\frac{-(\upsilon-\mu_\Upsilon(\phi))^2}{2\sigma_\Upsilon{}^2}} e^{\frac{-(\theta-\mu_\Theta(a,\phi))^2}{2\sigma_\Theta{}^2}}, \qquad \text{(B1)}$$

where $\mu_\Lambda(a)$, $\mu_\Theta(a,\phi)$, $\mu_\Upsilon(\phi)$, and γ are equal to $10\log_{10}\left(10^{\frac{L_0+\Delta L}{10}}+10^{\frac{L_0+a}{10}}\right)$,

$\Delta L - a - \frac{k}{\omega}\phi$, $-\frac{k}{\omega}\phi$, and $\frac{1}{(2\pi)^{\frac{3}{2}}\sigma_\Lambda\sigma_\Upsilon\sigma_\Theta}$, respectively.

The conditional probability of responding *Incremented* given Λ, Θ, and Υ $P_{\text{"Incremented"}|\Lambda,\Theta,\Upsilon}$ depends on the criterion jitters and the indicator function. The criterion jitters for the Λ, Θ, and Υ dimensions are denoted J_Λ, J_Θ, and J_Υ, respectively. The criterion jitters are both Gaussian and statistically independent across the dimensions. The indicator function divides the $\{\Lambda,\Theta,\Upsilon\}$ space into regions of indicate-*Un-Incremented* and indicate-*Incremented*. The boundaries which arise from the psychophysical experiment are such that for each value of $\{\Theta,\Upsilon\}$, there is only one *Un-Incremented/Incremented* boundary value of Λ (cf., Fig. 3); in other words each boundary can be represented by a single valued function of $\{\Theta,\Upsilon\}$, denoted $\Psi_{\Lambda,\Theta,\Upsilon}$. The indicator function indicates *Un-Incremented* for values of Λ less than $\Psi_{\Lambda,\Theta,\Upsilon}$ and *Incremented* for values of Λ greater than $\Psi_{\Lambda,\Theta,\Upsilon}$.

The probability of Λ being greater than the jittered boundary is equal to $P_{\text{"Incremented"}|\Lambda,\Theta,\Upsilon}$. Expressing $P_{\text{"Incremented"}|\Lambda,\Theta,\Upsilon}$ in terms the observations of Λ, Θ, and Υ, the criterion jitters J_Λ, J_Θ, and J_Υ, and the decision boundary $\Psi_{\Lambda,\Theta,\Upsilon}$, gives

$$P_{\text{"Incremented"}|\Lambda,\Theta,\Upsilon} = P_{\Lambda<\Psi_{\Lambda,\Theta,\Upsilon}(\Theta+J_\Theta,\Upsilon+J_\Upsilon)+J_\Lambda|\Lambda,\Theta,\Upsilon}\left(\lambda<\Psi_{\Lambda,\Theta,\Upsilon}(\theta+J_\Theta,\upsilon+J_\Upsilon)+J_\Lambda\,|\lambda,\theta,\upsilon\right).$$

Treating $P_{\Lambda<\Psi_{\Lambda,\Theta,\Upsilon}(\Theta+J_\Theta,\Upsilon+J_\Upsilon)+J_\Lambda|\Lambda,\Theta,\Upsilon}$ as a marginal distribution of $P_{\Lambda<\Psi_{\Lambda,\Theta,\Upsilon}(\Theta+J_\Theta,\Upsilon+J_\Upsilon)+J_\Lambda,\Theta+J_\Theta,\Upsilon+J_\Upsilon|\Lambda,\Theta,\Upsilon}$ gives

$$P_{\text{"Incremented"}|\Lambda,\Theta,\Upsilon} = \iint P_{\Lambda<\Psi_{\Lambda,\Theta,\Upsilon}(\Theta+J_\Theta,\Upsilon+J_\Upsilon)+J_\Lambda,\Theta+J_\Theta,\Upsilon+J_\Upsilon|\Lambda,\Theta,\Upsilon}\left(\lambda<\Psi_{\Lambda,\Theta,\Upsilon}(q,u)+J_\Lambda,q,u\middle|\lambda,\theta,\upsilon\right)dqdu.$$

Using the definition of conditional probability to expand $P_{\Lambda<\Psi_{\Lambda,\Theta,\Upsilon}(\Theta+J_\Theta,\Upsilon+J_\Upsilon)+J_\Lambda,\Theta+J_\Theta,\Upsilon+J_\Upsilon|\Lambda,\Theta,\Upsilon}$, noting the independence of $f_{\Theta+J_\Theta|\Lambda,\Theta,\Upsilon,J_\Upsilon}$ on J_Υ, and simplifying yields

$$P_{\text{"Incremented"}|\Lambda,\Theta,\Upsilon} = \iint\left[\begin{array}{l} f_{J_\Theta}(q-\theta)f_{J_\Upsilon}(u-\upsilon) \\ \times P_{J_\Lambda<\Lambda-\Psi_{\Lambda,\Theta,\Upsilon}(\Theta+J_\Theta,\Upsilon+J_\Upsilon)|\Lambda,\Theta,\Upsilon,\Theta+J_\Theta,\Upsilon+J_\Upsilon}\left(\lambda-J_\Lambda<\Psi_{\Lambda,\Theta,\Upsilon}(q,u)\middle|\lambda,\theta,\upsilon,q,u\right)\end{array}\right]dqdu.$$

Rewriting $P_{J_\Lambda<\Lambda-\Psi_{\Lambda,\Theta,\Upsilon}(\Theta+J_\Theta,\Upsilon+J_\Upsilon)|\Lambda,\Theta,\Upsilon,\Theta+J_\Theta,\Upsilon+J_\Upsilon}$ as a cumulative distribution results in

$$P_{\text{"Incremented"}|\Lambda,\Theta,\Upsilon} = \iint f_{J_\Theta}(q-\theta)f_{J_\Upsilon}(u-\upsilon)\int_{-\infty}^{\Psi_{\Lambda,\Theta,\Upsilon}(q,u)} f_{J_\Lambda}(\lambda-j_\Lambda)dj_\Lambda dqdu.$$

Making a substitution of the Gaussian density functions of J_Λ, J_Θ, and J_Υ results in

$$P_{\text{"Incremented"}|\Lambda,\Theta,\Upsilon} = \nu\int e^{\frac{-(q-\theta)^2}{2\sigma_{J_\Theta}^2}}\int e^{\frac{-(u-\upsilon)^2}{2\sigma_{J_\Upsilon}^2}}\int_{-\infty}^{\Psi_{\Lambda,\Theta,\Upsilon}(q,u)} e^{\frac{-(\lambda-j_\Lambda)^2}{2\sigma_{J_\Lambda}^2}}dj_\Lambda dqdu, \qquad \text{(B2)}$$

where ν is equal to $\dfrac{1}{(2\pi)^{\frac{3}{2}}\sigma_{J_\Lambda}\sigma_{J_\Upsilon}\sigma_{J_\Theta}}$.

Restating Eq. 4 in terms of Eqs. B1 and B2 and rearranging gives

$$P_D(a,\phi) = \gamma v$$

$$\times \int_u \int_q \int_{-\infty}^{\Psi_{\Lambda,\Theta,\Upsilon}(q,u)} \int_v e^{\frac{-(u-v)^2}{2\sigma_{J_\Upsilon}^2}} e^{\frac{-(v-\mu_\Upsilon(a))^2}{2\sigma_\Upsilon{}^2}} \int_\theta e^{\frac{-(q-\theta)^2}{2\sigma_{J_\Theta}^2}} e^{\frac{-(\theta-\mu_\Theta(a,\phi))^2}{2\sigma_\Theta{}^2}} \int_\lambda e^{\frac{-(\lambda-j_\Lambda)^2}{2\sigma_{J_\Lambda}^2}} e^{-\left(\frac{(\lambda-\mu_\Lambda(a))^2}{2\sigma_\Lambda{}^2}\right)} d\lambda d\theta dv dj_\Lambda dq du$$

Calculating the integrals over λ, θ, and υ results in

$$P_D(a,\phi) = \gamma v \int_u e^{\frac{-(u-\mu_\Upsilon(a))^2}{2(\sigma_\Upsilon{}^2+\sigma_{J_\Upsilon}^2)}} \int_q e^{\frac{-(q-\mu_\Theta(a,\phi))^2}{2(\sigma_\Theta{}^2+\sigma_{J_\Theta}^2)}} \int_{-\infty}^{\Psi_{\Lambda,\Theta,\Upsilon}(q,u)} e^{-\left(\frac{(j_\Lambda-\mu_\Lambda(a))^2}{2(\sigma_\Lambda{}^2+\sigma_{J_\Lambda}^2)}\right)} dj_\Lambda dq du$$

Making a substitution of variables to replace the non-intuitive j_Λ, q, and u notation with the more familiar λ, θ, and υ notation gives

$$P_D(a,\phi) = \gamma v \int_v e^{\frac{-(v-\mu_\Upsilon(a))^2}{2(\sigma_\Upsilon{}^2+\sigma_{J_\Upsilon}^2)}} \int_\theta e^{\frac{-(\theta-\mu_\Theta(a,\phi))^2}{2(\sigma_\Theta{}^2+\sigma_{J_\Theta}^2)}} \int_{-\infty}^{\Psi_{\Lambda,\Theta,\Upsilon}(\theta,v)} e^{-\left(\frac{(\lambda-\mu_\Lambda(a))^2}{2(\sigma_\Lambda{}^2+\sigma_{J_\Lambda}^2)}\right)} d\lambda d\theta dv$$

Replacing the definite integral over with λ with Eq. A5 yields

$$P_D(a,\phi) = \gamma v \int e^{\frac{-(v-\mu_\Upsilon(a))^2}{2(\sigma_\Upsilon{}^2+\sigma_{J_\Upsilon}^2)}} \int e^{\frac{-(\theta-\mu_\Theta(a,\phi))^2}{2(\sigma_\Theta{}^2+\sigma_{J_\Theta}^2)}} \mathrm{G}\left(-\infty, \Psi_{\Lambda,\Theta,\Upsilon}(\theta,v), \mu_\Lambda(a), \sqrt{\sigma_\Lambda{}^2+\sigma_{J_\Lambda}^2}\right) d\theta dv.$$

Since the $\Psi_{\Lambda,\Theta,\Upsilon}$ which arises from the psychophysical experiment is relatively smooth and can be approximated as a constant over a small range of $\{\Theta,\Upsilon\}$ the integrals over θ and υ to be approximated by summation. This allows for P_D to be re-written into a form which can then be computed:

$$P_D(a,\phi)=\gamma\nu\sum_n\left[\begin{array}{l}G\left(\upsilon[n],\upsilon[n+1],\mu_\Upsilon(a),\sqrt{\sigma_\Upsilon{}^2+\sigma_{J_\Upsilon}^2}\right)\\ \times\sum_m\left[\begin{array}{l}G\left(\theta[m],\theta[m+1],\mu_\Theta(a,\phi),\sqrt{\sigma_\Theta{}^2+\sigma_{J_\Theta}^2}\right)\\ \times G\left(-\infty,\Psi_{\Lambda,\Theta,\Upsilon}(\theta,\upsilon),\mu_\Lambda(a),\sqrt{\sigma_\Lambda{}^2+\sigma_{J_\Lambda}^2}\right)\end{array}\right]\end{array}\right],$$

where $\theta[m]$ and $\upsilon[n]$ are sampled versions of the continuous Θ and Υ.

Appendix C

This appendix proves that a coding noise N_C and a decision noise N_D cannot be summed together into a single equivalent coding noise N_{Eq}. The appendix begins by demonstrating that for some probability distributions the maximum likelihood criterion c is dependent on the standard deviation of a zero-mean Gaussian noise. Figure 7 shows the log-likelihood ratio of Λ (cf. Eq. 1) in the double-rove with a ΔL of 14 dB for three different noise standard deviations (0.5, 2, and 8 dB). The criterion value (the value of Λ for which the log-likelihood ratio is equal to zero) shifts systematically with changes to the noise standard deviation. For a standard deviation of 0.5 dB (the value used throughout this work) the criterion is 63.4 dB; for standard deviations of 2 dB and 8 dB the criterion is 62.1 and 61.6 dB, respectively.

To prove that N_C and N_D cannot be summed together into a single equivalent coding noise N_{Eq}, we consider the theoretical cases in which decisions are based on observations of the decision variable X (including separate coding and decision noises) or X' (single "equivalent" coding noise). The arbitrary event A (which can be conceptualized as a response of "un-incremented") occurs if X is less than C, the sum of the criterion value μ_{Coding} and the decision noise N_D. Similarly, the event A' occurs if X' is less than μ_{Eq}. The following derivation shows that the probabilities of A and A' are generally only equal if μ_{Coding} is equal to μ_{Eq}. Since, as discussed above, in the double-rove condition the criterion depends on the noise standard deviation, μ_{Coding} will not be equal to μ_{Eq}.

The random variable X is the sum of a random variable Y (which depends on the stimulus) and the coding noise N_C (which is independent of the stimulus) and X' is the sum of Y and N_{Eq}. We begin the derivation by formally stating the probabilities of A and A' as

$$P_A = \int_{-\infty}^{\infty} P\left(x < \mu_{Coding} + N_D \middle| x\right) f_X(x)\, dx \tag{C1}$$

and

$$P_{A'} = \int_{-\infty}^{\infty} P\left(x' < \mu_{Eq} \middle| x'\right) f_{X'}(x')\, dx'. \tag{C2}$$

Noting that $P\left(x' < \mu_{Eq} \middle| x'\right)$ is either equal to zero or unity gives

$$P_{A'} = \int_{-\infty}^{\mu_{Eq}} f_{X'}(\eta)\, d\eta .$$

Making use of the fact that X' is the sum of Y and N_{Eq} gives

$$P_{A'} = \int_{-\infty}^{\mu_{Eq}} \int_{-\infty}^{\infty} f_Y(\xi) f_{N_{Eq}}(\eta - \xi)\, d\xi\, d\eta . \tag{C3}$$

The probability of A can be expressed in a form similar to that of Eq. C3 by expressing $P\left(x < \mu_{Coding} + N_D \middle| x\right)$ as the integral of a probability density function such that one obtains

$$P_A = \int_{-\infty}^{\infty} f_X(x) \int_{-\infty}^{\mu_{Coding}} f_{N_D}(x-\eta)\, d\eta dx .$$

Then by expanding $f_X(x)$ (it is the sum of Y and N_D) one gets

$$P_A = \int_{-\infty}^{\mu_{Coding}} \int_{-\infty}^{\infty} f_Y(\xi) \int_{-\infty}^{\infty} f_{N_C}(x-\xi) f_{N_D}(x-\eta)\, dx d\xi d\eta .$$

Recognizing the innermost integral as the result of the summation of N_D and N_C allows for the integral to be replaced by N_{Eq} such that

$$P_A = \int_{-\infty}^{\mu_{Coding}} \int_{-\infty}^{\infty} f_Y(\xi) f_{N_{Eq}}(\eta-\xi)\, d\xi d\eta . \tag{C4}$$

Comparing Eqs. C3 and C4 one can see that P_A and $P_{A'}$ are generally only equal when μ_{Coding} and μ_{Eq} are equal. Since the criterion value in the double-rove condition depends on the noise standard deviation, one cannot simply sum the coding noise and decision noise into a single equivalent noise. Rather, coding noise and decision noise affect the probability of a particular response differently.

Appendix D

In this appendix the values of σ, $\mathrm{H}(a,\theta,\upsilon)$, and $\mu(a,\theta,\upsilon)$ which satisfy

$$\mathrm{H}(a,\theta,\upsilon)e^{\frac{-(\phi-\mu(a,\theta,\upsilon))^2}{2\sigma^2}} = \frac{1}{\sqrt{2\pi\sigma_\Theta^{\,2}}}e^{\frac{-\left(\frac{k}{\varpi}\phi+\theta-\Delta L+a\right)^2}{2\sigma_\Theta^{\,2}}}\frac{1}{\sqrt{2\pi\sigma_\Upsilon^{\,2}}}e^{\frac{-\left(\frac{k}{\varpi}\phi+\upsilon\right)^2}{2\sigma_\Upsilon^{\,2}}},$$

are found. We begins by rearranging and combing the terms on the right hand side of the equality to get

$$\mathrm{H}(a,\theta,\upsilon)e^{\frac{-(\phi-\mu(a,\theta,\upsilon))^2}{2\sigma^2}} = \frac{1}{2\pi\sigma_\Theta\sigma_\Upsilon}e^{\frac{-\left(\phi-\frac{\varpi}{k}(\theta-\Delta L+a)\right)^2}{2\left(\frac{\varpi}{k}\sigma_\Theta\right)^2}+\frac{-\left(\phi-\frac{\varpi}{k}\upsilon\right)^2}{2\left(\frac{\varpi}{k}\sigma_\Upsilon\right)^2}}.$$

Letting μ_1 equal $\frac{\omega}{k}(\Delta L-\theta-a)$, μ_2 equal $-\frac{\omega}{k}\upsilon$, σ_1 equal $\frac{\omega}{k}\sigma_\Theta$, and σ_2 equal $\frac{\omega}{k}\sigma_\Upsilon$ allows the above equation to be rewritten as

$$\mathrm{H}(a,\theta,\upsilon)e^{\frac{-(\phi-\mu(a,\theta,\upsilon))^2}{2\sigma^2}} = \frac{1}{2\pi\sigma_1\sigma_2}\left(\frac{\omega}{k}\right)^2 e^{\frac{-(\phi-\mu_1)^2}{2\sigma_1^2}+\frac{-(\phi-\mu_2)^2}{2\sigma_2^2}}.$$

Expanding the exponential term gives

$$\mathrm{H}(a,\theta,\upsilon)e^{\frac{-(\phi-\mu(a,\theta,\upsilon))^2}{2\sigma^2}} = \frac{1}{2\pi\sigma_1\sigma_2}\left(\frac{\omega}{k}\right)^2 e^{-\left(\frac{\sigma_1^2\sigma_2^2}{2\sigma_1^2\sigma_2^2}\phi^2-\frac{\mu_1\sigma_2^2+\mu_2\sigma_1^2}{\sigma_1^2\sigma_2^2}\phi+\frac{\sigma_2^2\mu_1^2+\sigma_1^2\mu_2^2}{2\sigma_1^2\sigma_2^2}\right)}, \quad \text{(D1)}$$

Recalling that

$$A^2x^2-2ABx+B^2+C=(Ax-B)^2+C=\frac{\left(x-\frac{B}{A}\right)^2}{2\left(\sqrt{\frac{1}{2A^2}}\right)^2}+C,$$

and expressing A, B, and C in terms of μ_1, μ_2, σ_1, and σ_2 gives

$$A^2x^2 = \frac{\sigma_1^{\,2}+\sigma_2^{\,2}}{2\sigma_1^{\,2}\sigma_2^{\,2}}x^2 \quad\rightarrow\quad A=\sqrt{\frac{\sigma_1^{\,2}+\sigma_2^{\,2}}{2\sigma_1^{\,2}\sigma_2^{\,2}}}$$

$$-2ABx = -\frac{\mu_1\sigma_2^{\,2}+\mu_2\sigma_1^{\,2}}{\sigma_1^{\,2}\sigma_2^{\,2}}x \quad\rightarrow\quad B=\frac{\mu_1\sigma_2^{\,2}+\mu_2\sigma_1^{\,2}}{2\sigma_1^{\,2}\sigma_2^{\,2}A} \quad ,$$

$$B^2+C = \frac{\sigma_2^{\,2}\mu_1^{\,2}+\sigma_1^{\,2}\mu_2^{\,2}}{2\sigma_1^{\,2}\sigma_2^{\,2}} \quad\rightarrow\quad C=\frac{\sigma_2^{\,2}\mu_1^{\,2}+\sigma_1^{\,2}\mu_2^{\,2}}{2\sigma_1^{\,2}\sigma_2^{\,2}}-B^2$$

allowing for equation D1 to be rewritten as

$$\mathrm{H}(a,\theta,\upsilon)e^{\frac{-(\phi-\mu(a,\theta,\upsilon))^2}{2\sigma^2}} = \frac{1}{2\pi\sigma_1\sigma_2}\left(\frac{\omega}{k}\right)^2 e^{\frac{-\left(\phi-\frac{B}{A}\right)^2}{2\left(\sqrt{\frac{1}{2A^2}}\right)^2}-C} .$$

One can then readily determine that

$$\sigma=\sqrt{\frac{1}{2A^2}}=\sqrt{\frac{1}{2\left(\sqrt{\frac{\sigma_1^{\,2}+\sigma_2^{\,2}}{2\sigma_1^{\,2}\sigma_2^{\,2}}}\right)^2}}=\sqrt{\frac{1}{2\frac{\sigma_1^{\,2}+\sigma_2^{\,2}}{2\sigma_1^{\,2}\sigma_2^{\,2}}}}=\sqrt{\frac{\sigma_1^{\,2}\sigma_2^{\,2}}{\sigma_1^{\,2}+\sigma_2^{\,2}}},$$

$$\mu=\frac{B}{A}=\frac{\mu_1\sigma_2^{\,2}+\mu_2\sigma_1^{\,2}}{4\sigma_1^{\,2}\sigma_2^{\,2}A^2}=\frac{\mu_1\sigma_2^{\,2}+\mu_2\sigma_1^{\,2}}{2\sigma_1^{\,2}\sigma_2^{\,2}}\sigma^2,$$

and

$$\mathrm{H}(a,\theta,\upsilon)=\frac{1}{2\pi\sigma_1\sigma_2}\left(\frac{\omega}{k}\right)^2 e^{-C}=\frac{1}{2\pi\sigma_1\sigma_2}\left(\frac{\omega}{k}\right)^2 e^{-\left(\frac{\sigma_2^{\,2}\mu_1^{\,2}+\sigma_1^{\,2}\mu_2^{\,2}}{2\sigma_1^{\,2}\sigma_2^{\,2}}-B^2\right)}$$

$$=\frac{1}{2\pi\sigma_1\sigma_2}\left(\frac{\omega}{k}\right)^2 e^{-\left(\frac{\sigma_2^{\,2}\mu_1^{\,2}+\sigma_1^{\,2}\mu_2^{\,2}}{2\sigma_1^{\,2}\sigma_2^{\,2}}-(\mu A)^2\right)}=\frac{1}{2\pi\sigma_1\sigma_2}\left(\frac{\omega}{k}\right)^2 e^{-\left(\frac{\sigma_2^{\,2}\mu_1^{\,2}+\sigma_1^{\,2}\mu_2^{\,2}}{2\sigma_1^{\,2}\sigma_2^{\,2}}-\frac{\mu^2}{2\sigma^2}\right)}.$$

Finally, re-substituting the original values for μ_1, μ_2, σ_1, and σ_2 gives

$$\sigma=\frac{\omega}{k}\sqrt{\frac{\sigma_\Theta^{\,2}\sigma_\Upsilon^{\,2}}{\sigma_\Theta^{\,2}+\sigma_\Upsilon^{\,2}}},$$

$$\mu(a,\theta,\upsilon)=\left(\frac{k}{\omega}\right)\frac{(\Delta L-\theta-a)\sigma_{\Upsilon}^{2}-\upsilon\sigma_{\Theta}^{2}}{2\sigma_{\Theta}^{2}\sigma_{\Upsilon}^{2}}\sigma^{2},$$

and

$$\mathrm{H}(a,\theta,\upsilon)=\frac{1}{2\pi\sigma_{1}\sigma_{2}}\left(\frac{\omega}{k}\right)^{2}e^{\frac{\sigma_{\Upsilon}^{2}(\Delta L-a-\theta)^{2}+\sigma_{\Theta}^{2}\upsilon^{2}}{2\sigma_{\Theta}^{2}\sigma_{\Upsilon}^{2}}-\frac{\mu^{2}}{2\sigma^{2}}}.$$

---

[1] No data was reported for S3.

[2] The mathematical complexity of the model drastically increases when all three dimensions depend on the target level and the distractor level and phase. The complexity is reduced by defining dimensions such that Θ depends on the target level and the distractor level and phase, that Λ depends on the target level and distractor level, and that Υ depends on only the distractor phase.

References

Blauert, J. (1997). "Spatial hearing: The psychophysics of human sound localization," (MIT Press, Cambridge).

Bonnel, A. M., and Hafter, E. R. (1998). "Divided attention between simultaneous auditory and visual signals," Percept Psychophys 60, 179-90.

Davidson, S. A., Gilkey, R. H., Colburn, H. S., and Carney, L. H. (2006). "Binaural detection with narrowband and wideband reproducible noise maskers.

III. Monaural and diotic detection and model results," J Acoust Soc Am 119, 2258-75.

Domnitz, R. (1973). "The interaural time jnd as a simultaneous function of interaural time and interaural amplitude," J. Acoust. Soc. Am. 53, 1549-52.

Domnitz, R. H., and Colburn, H. S. (1977). "Lateral position and interaural discrimination," J. Acoust. Soc. Am. 61, 1586-98.

Hafter, E. R., and Carrier, S. C. (1972). "Binaural interaction in low-frequency stimuli: The inability to trade time and intensity completely," J. Acoust. Soc. Am. 51, 1852-62.

Viemeister, N. F. (1988). "Psychophysical aspects of auditory intensity coding" in Auditory function: Neurobiological bases of hearing, Vol. edited by G. M. Edelman, W. E. Gall and W. M. Cowan.(John Wiley & Sons, NYC)

Yost, W. A. (1972). "Tone-on-tone masking for three binaural listening conditions," J. Acoust. Soc. Am. 52, 1234-7.

Figures and Tables

	S1	S2	S4
ΔL (dB)	8	8	14
σ_{J_Λ}, σ_{J_Θ}, and σ_{J_Υ} (dB)	1.0, 2.1, and 15.5	0.0, 7.0, and 5.8	12.6, 7.0, and 1.1
Mean Difference P_C	0.04	0.01	0.00
Mean Difference Bias	0.01	0.04	0.04
RMS Difference for P_F	0.13	0.14	0.13
RMS Difference for P_D	0.14	0.11	0.14
RMS Difference for P_F and P_D	0.14	0.13	0.14
R for P_F	0.73	0.74	0.54
R for P_D	0.64	0.10	0.48
R for P_F and P_D	0.86	0.89	0.87

Table 1 Model Parameters and Performance

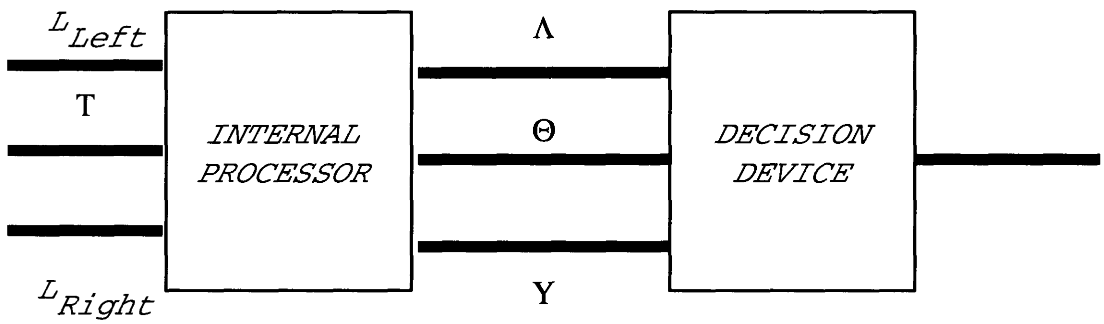

Figure 1. Block diagram of the model based on a non-optimal observer of Λ, Θ, and Y. The model inputs are the level of the left ear L_{Left}, the level of the right ear L_{Right} and the interaural difference in time delay T. The model does not specify mechanisms for estimating L_{Left}, L_{Right}, and T.

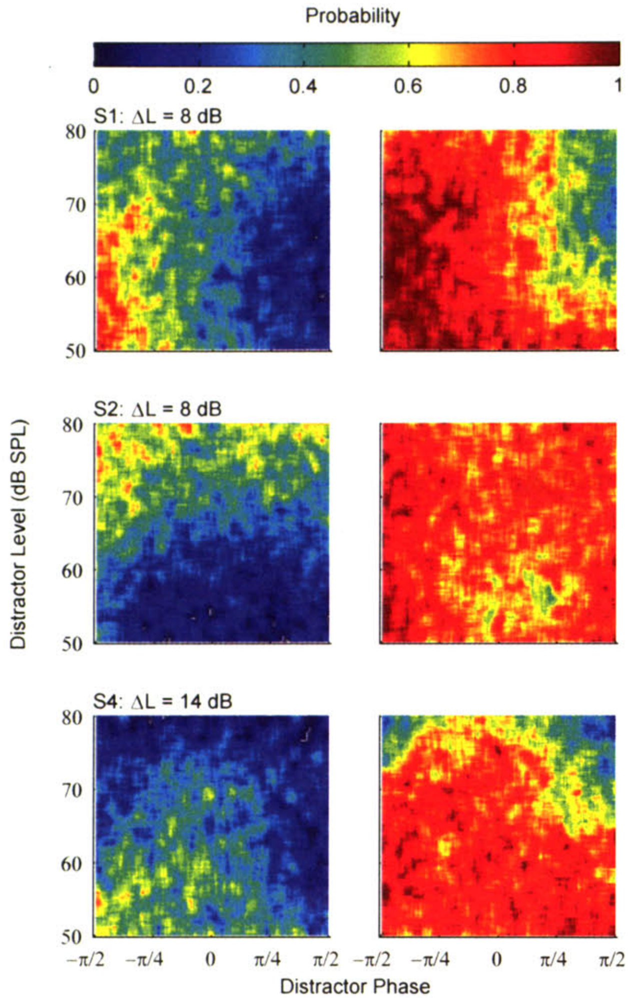

Figure 2. Two-dimensional surface plots of P_F (left column) and P_D (right column) as a function of the distractor level and phase in the double-rove condition for subjects S1 (top row), S2 (middle row), and S4 (bottom row). Areas of high probability are red and areas of low probability are blue. Each surface plots consist of approximately 20,000 overlapping bins with a width of 40 μs and a height of 2 dB.

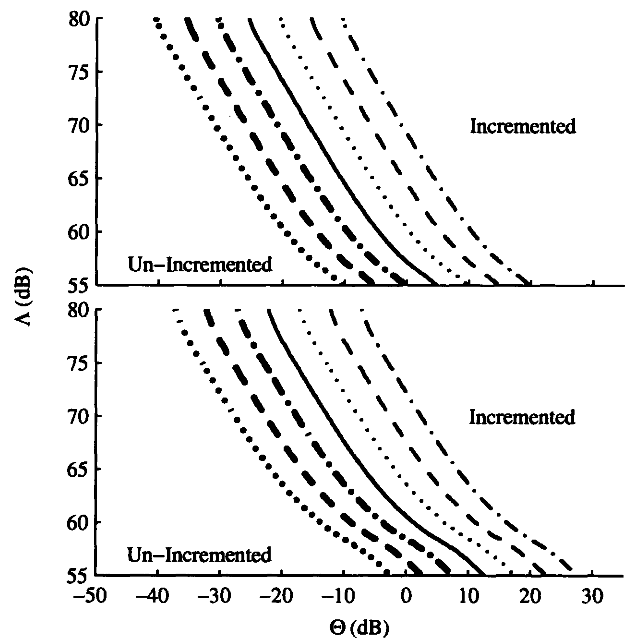

Figure 3. Maximum likelihood indicator functions $\Psi_{\Lambda,\Theta,\Upsilon}$ for the $\{\Lambda,\Theta,\Upsilon\}$ space with a ΔL of 8 dB (top panel) and 14 dB (bottom panel). Each line corresponds to $\Psi_{\Lambda,\Theta,\Upsilon}$ for a particular value of Υ. The upper-right most line is for Υ of -15 dB (an ITD of -300 μs) and the lower-left most line is for a Υ of 15 dB.

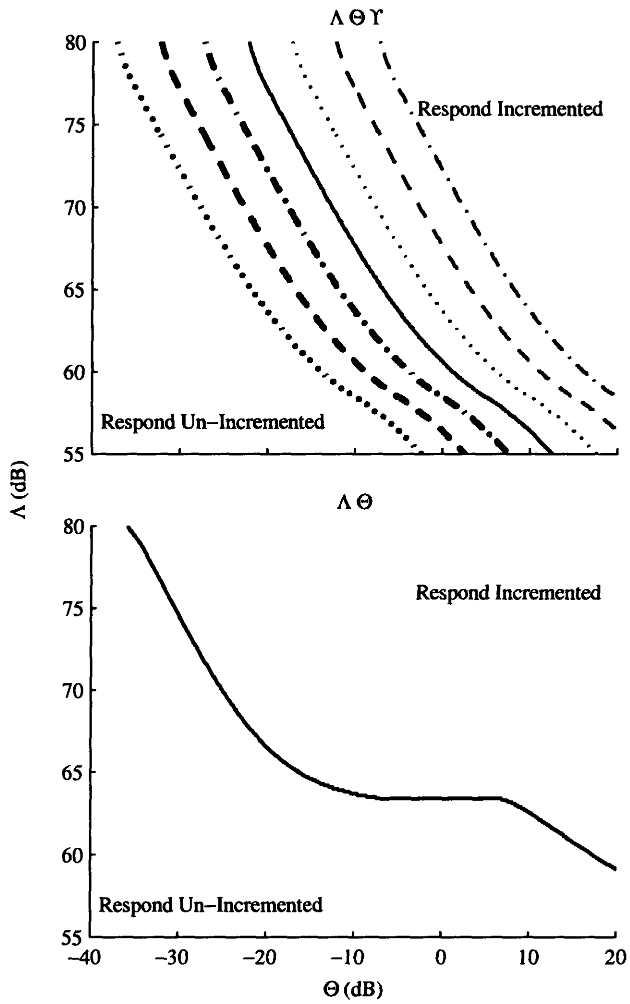

Figure 4. Maximum likelihood indicator function $\Psi_{\Lambda,\Theta,\Upsilon}$ for the $\{\Lambda,\Theta,\Upsilon\}$ space (top panel) and $\Psi_{\Lambda,\Theta}$ for the $\{\Lambda,\Theta\}$ space (bottom panel) with a ΔL of 8 dB. The $\Psi_{\Lambda,\Theta,\Upsilon}$ functions is the same as in Fig. 3. The indicator function $\Psi_{\Lambda,\Theta}$ cannot be obtained by a simple weighted combination of $\Psi_{\Lambda,\Theta,\Upsilon}$ at different values of Y.

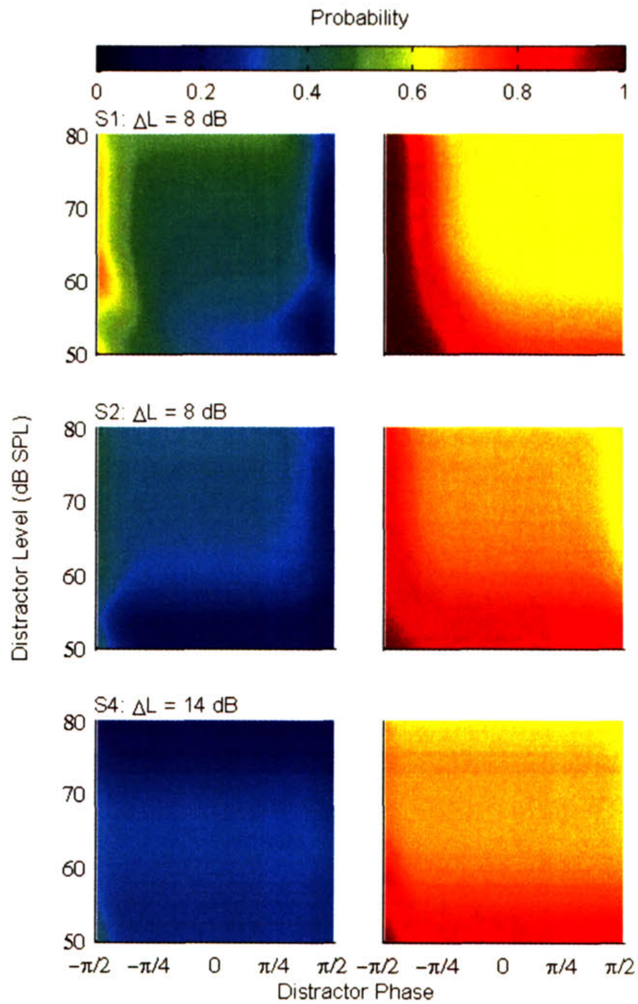

Figure 5. Two-dimensional surface plots of the predicted P_F (left column) and P_D (right column) as a function of the distractor level and phase for subjects S1 (top row), S2 (middle row), and S4 (bottom row). Areas of high probability are red and areas of low probability are blue. The model parameters σ_{J_Λ}, σ_{J_Θ}, and σ_{J_Υ} were adjusted for each subject individually.

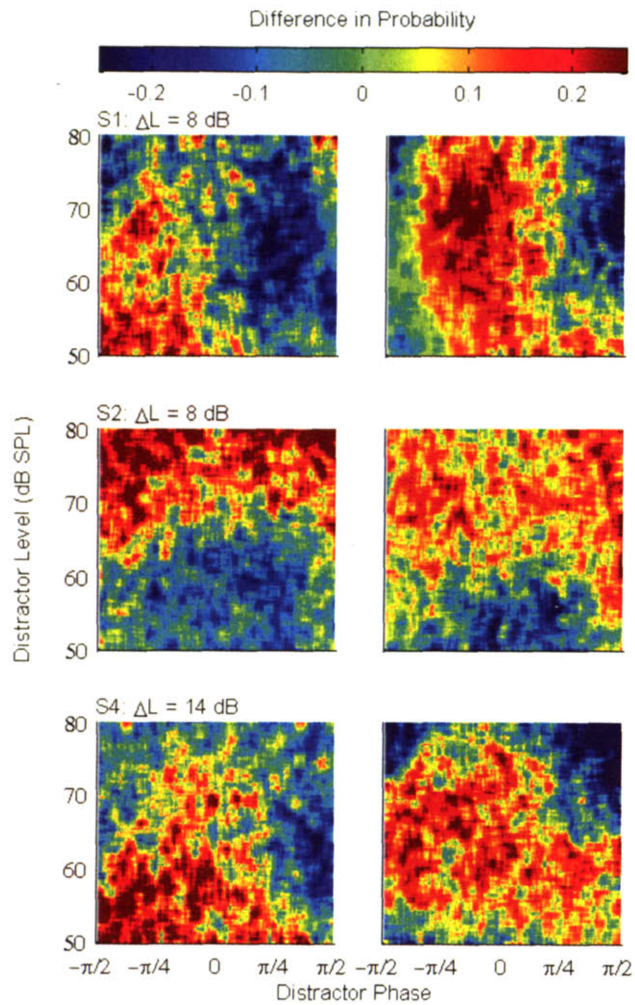

Figure 6. Two-dimensional surface plots of the difference between the measured and predicted P_F (left column) and P_D (right column) as a function of the distractor level and phase for subjects S1 (top row), S2 (middle row), and S4 (bottom row). Areas in which the predicted is less than the measured are blue and areas with higher predicted probability are red. Note the expanded color scale.

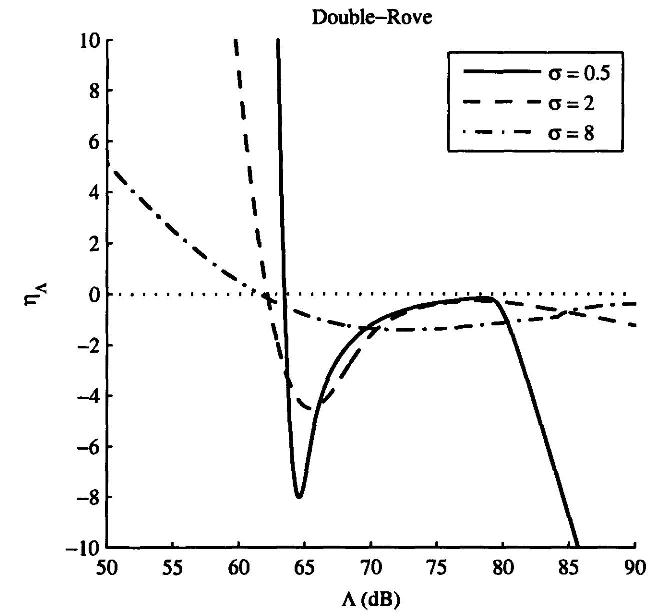

Figure 7. Log-likelihood ratios for Λ in the double-rove condition with a ΔL of 14 dB and three different coding noise standard deviations. The criterion value of Λ changes as a function of the standard deviation. This dependence of the criterion value on the standard deviation of the coding noise is required for separating the effects of coding noise and the decision noise.

CHAPTER V. CONCLUSIONS

I. Introduction

This chapter summarizes the results of Chapters II, III, and IV and discusses some future work that follows upon the understanding achieved by the "thesis" and would expand our understanding of contra-aural interference. In summarizing the results of chapters II, III, and IV emphasis is given to integrating the results across the chapters. The future work section is divided into two parts: psychophysics and modeling.

The set of psychophysical experiments reported in Chapters II and III involved the discrimination of the level of a monotic target (600-Hz tone) in the presence of a monotic distractor (600-Hz tone) presented simultaneously but contra-aurally to the target. The stimuli were such that the dominant perception was a fused image with a salient loudness and position. The reliabilities of the loudness and position in the discrimination of changes in the target level were manipulated by parametrically varying the distractor level and phase across the experimental conditions. In all conditions the information carried by the target was unchanged and therefore attending exclusively to the ear at which the target was presented would lead to identical performance across conditions. The target and distractor stimuli were chosen, however, such that they perceptually fused; attending to the loudness and position of this fused image would result in variable performance across conditions.

In the simplest condition the distractor level and phase were fixed and therefore both the loudness and position carried useful information for discriminating the target level. In the most complex condition, both the distractor level and phase were randomly chosen on every presentation and therefore the loudness, the position, and their combination did not carried sufficient information for discriminating the target level. In addition to the information carried by the loudness and position of the dominant-image, additional secondary images (e.g., the "time-image") may have also carried information.

The modeling work considers a theoretical observer of simple representations of the perceived loudness and position of the dominant-image and the position of the time-image. Even in the most complex experimental condition tested (random distractor level and phase), the ideal observer of loudness, position, and time-image together is only limited by internal noise. Therefore both optimal and non-optimal integration of the information across the modeled perceptual dimensions is considered. Although the model is only applied to the experimental conditions tested in this thesis, the similarity of the model to traditional models (Hafter 1971; Yost 1972; Stern and Colburn 1978) is such that it is believed that the model will predict a wider set of psychophysical results.

II. Summary

The paradigms of the psychophysical experiments in Chapter II (multi-interval adaptive) and Chapter III (single-interval constant-increment) were different, but the most general findings were similar. Discrimination performance decreased when the distractor variability reduced the reliability of both the loudness and position for discriminating changes in the target level. Since the distractor was presented contra-aurally to the target, this decrease in discrimination shows contra-aural interference and provides evidence that subjects cannot attend exclusively to a single ear even when it is desirable. In control conditions in which both the loudness and position, or only the loudness, or only the position carried the information, subjects performed similar to previous studies. The conditions in which the reliability of both the loudness and position were reduced was modeled extensively.

In both Chapters II and III a model based on optimal use of the loudness and position was used to interpret the results. The predictive power of this model was limited. In some cases the model failed to predict the extent of the contra-aural interference (the loudness and position carried too much information). The model also failed to predict the dependence of the probability of a false alarm P_F and the probability of a detection P_D as a function of the distractor level and phase. In Chapter IV the scope of the model was expanded; a model based on the non-optimal use of the loudness and position of the dominant-image and the position of the time-image partially predicted the dependence of P_F and P_D on the distractor level and phase.

Predictions of the measured contra-aural interference presented here were limited to a modification of traditional binaural models. In the modified model, the monaural processors were removed and the decision process was implemented as a non-ideal observer. It is possible that a model with monaural processors could lead to similar (or better) predictions of the contra-aural interference. In such a model, the ability to use the relevant monaural information, or the information itself, would need to be degraded by the introduction of the distractor.

One structure for such a corrupted monaural model could be based on the activity of efferent neural pathways that are stimulated by the contralateral stimulus. (Guinan 1996) Incorporating the efferent neural pathway into the monaural (and binaural) processors might degrade the information appropriately. It is not obvious, however, how such a model would predict the inter-subject variability in the dependence of P_F and P_D on the distractor level and phase. The current model predicts the inter-subject variability by assuming that each subject integrates the information across multiple perceptual dimensions differently.

This thesis investigates overall-level discrimination of a low-frequency tonal stimulus and demonstrates that the ability to discriminate small changes in the level of a target stimulus presented at one ear can be adversely affected by a distractor stimulus presented simultaneously at the ear contralateral to the target. The thesis focuses only on conditions in which the target and distractor are perceptually fused. The introduction of a distractor stimulus contra-aural to the target, which decreases the reliability of the perceived loudness and position for discriminating changes in the target level, increased discrimination thresholds substantially. The results are consistent with a model based on non-optimal integration of the information carried by the loudness, position, and time-image. The non-optimality assumed that subjects had difficulties reliably using the multi-dimensional maximum likelihood decision rule; this idea was implemented as criterion jitter. The modeling of the psychophysical results demonstrates that, although contra-aural interference is not predicted by traditional models, contra-aural interference is consistent with a reasonable use of the information carried by binaural perceptual attributes.

III. Future Work

A. Psychophysical

Future work on the psychophysical aspects of contra-aural interference could focus on the perceptual fusion of the target and the distractor. In the experiments reported in this thesis, the target and distractor had the same frequencies and were presented simultaneously. The target and distractor were perceptually fused. The perceptual fusion was maintained throughout the psychophysical experiments. Understanding how perceptual fusion affects contra-aural interference will give insight into the mechanisms by which the stimuli at the two ears are integrated (non-optimally). Traditional binaural models assume that the information from each ear is available individually and are obviously incorrect. The current model assumes an obligatory non-optimal integration of the information; this integration is independent of the perceptual fusion. Future psychophysical work on how perceptual fusion affects the integration is crucial for testing the model.

B. Modeling

The model presented in Chapter IV only made predictions for the condition in which the distractor level and phase were randomized. No predictions for other psychophysical tasks were made. The reports of contra-aural interference by Taylor and colleagues (Taylor and Clarke 1971; Taylor et al. 1971a; Taylor et al. 1971b) work on Monaural Detection with Contralateral Cue (MDCC) are conceptually inconsistent with the model; in those studies the "loudness", as defined in the model, was reliable, yet subjects were unable to use the "loudness" to detect a target. To our knowledge the results of Taylor and Colleagues have never been satisfactorily conceptualized or modeled. Shub and Colburn (2005) suggested that modifications to the "loudness" coupled with the position variable model of Stern and Colburn (1978) might be able to predict the contra-aural reported in the MDCC studies. The effects on the predictions of the MDCC studies of incorporating the recent work by Edmonds and Culling (2006) on the dependence of the perceptual binaural loudness on the interaural correlation into the model should be explored. This exploration would consist of redefining the loudness dimension and re-deriving the model to make predictions of the MDCC results.

The model also mathematically separates the effects of internal coding noise and criterion jitter. The multi-dimensional nature of the model allows for these effects to be evaluated separately. The internal coding noise determines the extent to which the perceptions of identical stimuli vary while the criterion jitter determines the extent to which the responses to identical perceptions vary. Although the model presents a way of separating these effects (for both contra-aural interference and perception in general), only a cursory examination was conducted. Future work on the differential effects of internal coding noise and criterion jitter may prove helpful for a wide range of psychological studies.

References

Edmonds, B., A., and Culling, J., F. (**2006**). "Interaural correlation and loudness," The Journal of the Acoustical Society of America **119,** 3235-6.

Guinan, J. J. J. (**1996**). "The physiology of olivocochlear efferents" in *The cochlea,* Vol. 8 edited by P. Dallos, A. N. Popper and R. R. Fay.(Springer, New York)

Hafter, E. R. (**1971**). "Quantitative evaluation of a lateralization model of masking-level differences," J. Acoust. Soc. Am. **50,** 1116-22.

Shub, D., E., and Colburn, H. S. (**2005**). "A biologically inspired binaural approach to monaural modeling," The Journal of the Acoustical Society of America **117,** 2563-4.

Stern, R. M., Jr., and Colburn, H. S. (**1978**). "Theory of binaural interaction based in auditory-nerve data. IV. A model for subjective lateral position," J. Acoust. Soc. Am. **64,** 127-40.

Taylor, M. M., and Clarke, D. P. J. (**1971**). "Monaural detection with contralateral cue (MDCC). II. Interaural delay of cue and signal," J. Acoust. Soc. Am. **49,** 1243-53.

Taylor, M. M., Clarke, D. P. J., and Smith, S. M. (**1971a**). "Monaural detection with contralateral cue (MDCC). III. Sinusoidal signals at a constant performance level," J. Acoust. Soc. Am. **49,** 1795-804.

Taylor, M. M., Smith, S. M., and Clarke, D. P. (**1971b**). "Monaural detection with contralateral cue (MDCC). IV. Psychometric functions with sinusoidal signals," J. Acoust. Soc. Am. **50,** 1151-61.

Yost, W. A. (**1972**). "Tone-on-tone masking for three binaural listening conditions," J. Acoust. Soc. Am. **52,** 1234-7.

BIOGRAPHICAL NOTE

Daniel Eric Shub
Born November 23, 1974, Saugus, Massachusetts

Education:
1993 Holderness School
1997 University of Pennsylvania, B.S.E in Bioengineering
2001 Massachusetts Institute of Technology, M.S. in Electrical Engineering

Teaching:
Massachusetts Institute of Technology
 Acoustics of Speech and Hearing (HST.714)
 Psychoacoustics Project Laboratory (6.182)
University of Pennsylvania
 Bioengineering Lab IV (BE 310)

Fellowships and Grants:
2005 University of California Los Angeles, Institute for Pure and Applied Mathematics travel award
2004 National Institutes of Health, National Research Service Award Individual Predoctoral Fellowship
2001 Acoustical Society of America travel award

Professional Societies:
Acoustical Society of America
Association for Research in Otolaryngology