

AUDIOVISUAL INTEGRATION IN THE
SACCADIC SYSTEM OF THE BARN OWL

by

ELIZABETH A. WHITCHURCH

A DISSERTATION

Presented to the Department of Biology
and the Graduate School of the University of Oregon
in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy

December 2006

“Audiovisual Integration in the Saccadic System of the Barn Owl,” a dissertation prepared by Elizabeth A. Whitchurch in partial fulfillment of the requirements for the Doctor of Philosophy degree in the Department of Biology. This dissertation has been approved and accepted by:

Dr. Janis C. Weeks, Chair of the Examining Committee

28 Nov. 2006

Date

Committee in Charge: Dr. Janis C. Weeks, Chair
 Dr. William W. Roberts
 Dr. Terry T. Takahashi
 Dr. Helen J. Neville
 Dr. Richard T. Marrocco

Accepted by:

Dean of the Graduate School

© 2006 Elizabeth A. Whitchurch

An Abstract of the Dissertation of

Elizabeth A. Whitchurch for the degree of Doctor of Philosophy

in the Department of Biology to be taken December 2006

Title: AUDIOVISUAL INTEGRATION IN THE SACCADIC SYSTEM OF THE BARN
OWLApproved: _____
Dr. Tetty T. Takahashi

Survival depends on our ability to detect and integrate sensory information from multiple modalities, allowing for the most efficient behavioral response. For example, barn owls must combine sights and sounds from the environment to localize potential prey. A vole scurrying through a drift of dried leaves is more likely to meet its doom if a nearby owl can both faintly see and hear it. How does the brain take two physically discreet inputs and combine them into a unified representation of the surrounding multisensory world? Moreover, how is this internal representation transformed into the most efficient behavioral response?

This dissertation comprises *original research* addressing these questions in the barn owl with two distinct approaches: First, Chapters II and III describe orientation behavior in response to auditory, visual, and audiovisual stimuli. Chapter II probes the effect of stimulus strength on saccadic behavior and the nature of audiovisual integration, and was

taken from a co-authored publication. Chapter III explores the behavioral consequence of an induced stimulus asynchrony in audiovisual integration and was taken from a co-authored manuscript being prepared for publication. The second experimental approach is described in Chapters IV and V. These chapters probe the physiological basis of saccadic behavior by measuring single-neuron responses to auditory, visual, and audiovisual stimuli. Chapter IV describes how auditory responses of neurons from the external nucleus of the inferior colliculus depend on sound pressure level. Chapter V describes activity of optic tectum neurons in response to auditory, visual, and audiovisual stimuli.

The behavioral findings described herein suggest that barn owls often incorporate both the speed of the auditory system and the accuracy of the visual system when localizing a multisensory stimulus, even when the two modalities are presented asynchronously. The physiological studies outlined in this dissertation show that sensory representations in the midbrain can be used to predict general trends in saccadic behavior: Neuronal thresholds were within the range of observed behavioral thresholds. Responses to multisensory stimuli were enhanced relative to unisensory stimuli, possibly corresponding to enhanced multisensory behavior. These data support fundamental rules in multisensory integration that may apply across species.

CURRICULUM VITAE

NAME OF AUTHOR: Elizabeth A. Whitchurch

PLACE OF BIRTH: Moscow, Idaho

DATE OF BIRTH: July 29, 1976

GRADUATE AND UNDERGRADUATE SCHOOLS ATTENDED:

University of Oregon
Utah State University
College of Southern Idaho

DEGREES AWARDED:

Doctor of Philosophy in Biology, 2006, University of Oregon
Bachelor of Science in Chemistry, 1998, Utah State University

AREAS OF SPECIAL INTEREST:

The Neural Basis of Behavior in a Multisensory Environment
Electrophysiological Techniques in Neuroscience

PROFESSIONAL EXPERIENCE:

Graduate Research Assistant, Department of Biology, Institute of Neuroscience,
University of Oregon, Eugene, 2001-2006

Graduate Teaching Assistant, Department of Biology, Institute of Neuroscience,
University of Oregon, Eugene, 2000

Graduate Research Assistant, Department of Chemistry, Utah State University,
Logan, 1998-1999

Graduate Teaching Assistant, Department of Chemistry, Utah State University,
Logan, 1998

Undergraduate Research Assistant, Department of Chemistry, Utah State University,
Logan, 1997-1998

GRANTS, AWARDS, AND HONORS:

Travel Award to the Gordon Research Conference in Oculomotor System Biology,
Bates College, Lewiston, Maine, 2005

Travel Award to the 7th International Congress of Neuroethology, Nyborg,
Denmark, 2005

Teacher Recognition Award, Department of Biology, University of Oregon, 2002

National Institute of General Medical Science Training Grant, Systems
Neuroscience, University of Oregon, 2001-2002.

Deans List, Utah State University, 1994-1998

Full-Tuition Award for Academic Achievement, Utah State University, 1994-1998

PUBLICATIONS:

Whitchurch EA, Takahashi TT (2006) Combined auditory and visual stimuli facilitate
head saccades in the barn owl (*Tyto alba*). *Journal of Neurophysiology* 96: 730-745.

ACKNOWLEDGMENTS

I wish to express sincere appreciation to my mentor and graduate advisor, Terry Takahashi, for his unceasing faith in my ideas and abilities. In addition, special thanks to Dr. Clifford Keller (a.k.a. Kip) whose friendship and guidance were invaluable throughout my graduate experience and beyond. I also thank the members of my graduate dissertation committee, Drs. Janis Weeks, Bill Roberts, Rich Marrocco, and Helen Neville, for their valuable input and oversight of my progress toward this degree. I am very grateful for the opportunity to work with talented and intelligent colleagues and mentors throughout my graduate career: Dr. Avinash Deep Singh Bala, Dr. Matt Spitzer, Dr. Brian Nelson, Dr. Kip Keller, Dr. Michael Spezio, Dr. Shin Yanagihara, Hanna Smedstad, Ted Cornforth, Dave Rowland, Pamela Johnston, and Hannah Dold. Finally, I want to thank my family and especially my husband and parents for being the role-models and pillars of support that I needed to get *this far*.

These investigations were supported in part by The National Institute on Deafness and Communication Disorders: Grant DC03925 to Terry Takahashi, and The National Institute of General Medical Sciences: Training Grant T32 GM07257 to the University of Oregon Institute of Neuroscience.

To my husband, Brian, and our children.

TABLE OF CONTENTS

Chapter	Page
I. INTRODUCTION.....	1
Overview	1
Anatomy of the Avian Optic Tectum.....	3
Auditory and Visual Sensory Processing.....	5
Summary of Data Chapters	8
II. COMBINED AUDITORY AND VISUAL STIMULI FACILITATE HEAD SACCADES IN THE BARN OWL (<i>TYTO ALBA</i>).....	13
Introduction.....	13
Methods.....	16
Apparatus	17
Acoustical Stimuli	17
Visual Stimuli.....	18
Behavioral Paradigm	18
Data Analyses	22
Results.....	27
Responses to Unisensory Stimuli.....	28
Responses to Audiovisual Stimuli	34
Saccade Reaction Times in Audiovisual Trials.....	34
Saccade Error in Audiovisual Trials.....	39
Discussion	47
Probability of Response.....	48
Saccade Reaction Time.....	49
Negative Race Model Violations in Saccade Reaction Time.....	51
Error.....	53
Neural Mechanisms.....	54
III. TIMING AUDITORY AND VISUAL STIMULI FOR OPTIMAL INTEGRATION	59
Introduction.....	59
Methods.....	62
Apparatus	63
Sensory Stimuli.....	63
Behavioral Paradigm	64
Data Analyses	66
Results.....	72
SRT.....	73
Error.....	78

Chapter	Page
Saccade Kinematics	85
Discussion	92
SRT	93
Error	94
The Significance of the Double-Step Saccade	95
What Happens if Audition Loses the Race	96
The Double-Step Model	98
IV. NEUROPHYSIOLOGICAL CORRELATES OF BEHAVIORAL CHANGES WITH SOUND PRESSURE LEVEL	100
Introduction	100
Methods	101
Apparatus and Stimuli	102
PDR Paradigm	103
Electrophysiological Paradigm	104
Data Analysis	105
Results	114
Cellular Thresholds	114
First-Spike Latency	115
Spatial Response Profiles	116
Discussion	118
Linking Cellular to Behavioral Thresholds	118
The Relevance of the First-Spike Latency	120
SPLs at the Eardrum Depend on Stimulus Location	121
V. AUDIOVISUAL INTEGRATION IN THE OPTIC TECTUM OF THE BARN OWL	124
Introduction and Background	124
Audiovisual Integration in the Barn Owl	125
Audiovisual Integration in the Mammalian SC	126
Audiovisual Integration in Awake vs. Anesthetized Preparations	129
Audiovisual Integration and the Sensory-Motor Transformation	129
Methods	131
Apparatus and Stimuli	132
Electrophysiological Paradigm and Data Analysis	133
Results and Discussion	133
VI. CONCLUSIONS	140
BIBLIOGRAPHY	142

LIST OF FIGURES

Figure	Page
1-1. A slice from the barn owl midbrain.....	4
1-2. The tubular shape of an owl eye.....	6
1-3. The external ear.	7
2-1. Head saccade measurements.....	19
2-2. Cumulative probability distributions predicted by the race model.	25
2-3. Effects of stimulus strength on response percentage.	29
2-4. Effects of stimulus strength on saccade reaction time (SRT).....	30
2-5. Effects of target eccentricity and stimulus strength on saccadic error.	31
2-6. Effects of stimulus strength on azimuthal and elevational components of error.....	33
2-7. Average audiovisual SRTs compared to unisensory values for bird 'N'.....	35
2-8. Cumulative audiovisual SRT distributions with a quiet sound.	37
2-9. Cumulative audiovisual SRT distributions with a bright visual component.	38
2-10. Comparison of errors from auditory, visual, and audiovisual trials.	40
2-11. Multivariate analysis of saccadic behavior for one stimulus pair.....	43
2-12. Summary of mean SRT and error for birds 'N' (left) and 'J' (right).	45
3-1. Example of a normal saccade.	70
3-2. The effects of SOA on SRT.....	73
3-3. Observed SRT distributions vs. race model predictions – bright conditions.....	75
3-4. Observed SRT distributions vs. race model predictions – dim conditions.	76
3-5. Probability of winning the race to saccade initiation – Simulation results.....	78

Figure	Page
3-6. Saccade amplitude as a function of target eccentricity.....	79
3-7. Elevation error statistics.....	80
3-8. Summary of accuracy.....	81
3-9. Summary of precision.....	82
3-10. Fisher information – Examples from bird 'J'.....	84
3-11. Summary of MLE fit to observed data.....	85
3-12. Example of double-step saccade.....	86
3-13. Observed probability of double-step saccades.....	87
3-14. Timeline for parallel processing of two target signals.....	88
3-15. Delayed modality processing may occur after the completion of the saccade.....	89
3-16. Finding the optimal integration window.....	91
3-17. The best-fit prediction of double-step probability.....	92
4-1. Raw voltage traces from a PDR experiment.....	104
4-2. Free-field extrapolation of SPL.....	106
4-3. Measurements of free-field vs. virtual sounds from within the ear canal.....	107
4-4. Neuronal comparisons of free-field vs. virtual sounds.....	109
4-5. ICx neurons response to sound presented from the best area.....	111
4-6. Rate-level plot.....	112
4-7. Integrated PDR responses to stimuli of varying SPL.....	113
4-8. Population summary of cellular threshold.....	114
4-9. Behavioral discrimination thresholds.....	115
4-10. First-spike latencies of neurons with low threshold.....	116

Figure	Page
4-11. Spatial response profiles vary with SPL.....	117
4-12. HRTFs attenuate frequencies differentially, depending on spatial location.....	122
5-1. OT single unit responses in audiovisual trials.	134
5-2. OT neurons follow the inverse effectiveness principle.	136
5-3. OT responses are variable in time.....	138

CHAPTER I

INTRODUCTION

Overview

The animal kingdom is brimming with species that have evolved distinct sensory specializations. Whether it is the rattlesnake (an American species of pit viper) with its ability to sense infrared radiation and strike warm blooded animals moving through the tall grass, or a weakly electric fish with its electric organ and electro-receptors making possible communication and navigation in the murky waters of the Amazon, animals have responded to environmental pressures and evolved keenly efficient sensory systems for survival. The barn owl (*Tyto alba*) is no exception. This amazing nocturnal hunter is renowned for its ability to localize potential prey using sound alone (Payne and Drury, 1958; Payne, 1971; Konishi, 1973). But what becomes then, of these specialists' other senses? Do they become somewhat vestigial or can they be utilized to further elevate the chances of survival?

Both the viper and the electric fish maintain a keen visual sense in the presence of their individual specializations, and studies have shown that they are able to integrate across modalities. For example, the schooling behavior of the weakly electric mormyrid fish was found to be based equally on both its electric and visual senses, degrading in a linear fashion when one modality was hindered (Moller, 2002). Similarly, recent behavioral studies of pit vipers show that although these animals are able to hunt and successfully strike highly mobile

prey with little to no visual input, performance can be affected by complete visual deprivation or by altering visual input during development (Grace et al., 2001). Furthermore, the rattlesnake's optic tectum, an area of the brain important for orienting responses in most vertebrates, is known to have neurons that respond to both infrared and visible light even though the pit organ and the lateral eye are two physically separate receivers on the body (Hartline et al., 1978). These examples underline the fact that although modality specific sensory organs are both spatially and functionally distinct, somewhere in the process of sensory-motor integration they come together for the benefit of survival.

This dissertation focuses on the barn owl and its use of both vision and audition in orientation. When owls hunt, they spend much of their time perched relatively high above the ground surveying the surrounding area for signs of prey. When a sight or sound rises above the background noise, the animal initially responds with a quick head turn (also termed head saccade) to direct both the eyes and ears to a certain location. Saccades are ubiquitous in the animal kingdom as a way to orient toward the most interesting area in space at any given moment. Every day, humans make thousands of saccades, either in response to a new stimulus in the environment (reflexive saccades) or to purposefully direct their gaze toward a new location (voluntary saccade). Perhaps in part due to the relative simplicity and behavioral importance of saccades, the neural control of saccade generation is one of the best studied areas in sensory-motor integration (Jay and Sparks, 1987; Munoz et al., 2000; Sparks, 2002; Glimcher, 2003; Populin et al., 2004; Arai and Keller, 2005). However, in the barn owl, a careful examination of this behavior in response to various stimulus strengths and modalities has never been done.

In the following chapters, I present original work aimed at understanding how changes in sensory inputs might lead to changes in the motor output of the animal as a system. Therefore I focus mainly on the sensory side of the sensory-motor transformation that underlies the basic head saccade in the barn owl. Of particular interest is the way in which two sensory modalities interact in the process of saccade generation. The rest of the introduction will provide background for the following data chapters, including a review of the anatomy of the optic tectum and the visual and auditory sensory systems in the barn owl. A chapter by chapter summary of the data collected follows these reviews.

Anatomy of the Avian Optic Tectum

The optic tectum (OT) is an ideal neural substrate for the study of sensory-motor as well as audiovisual integration because it is one of the few structures known to have topographic representations of multiple sensory modalities, as well as a fundamental role in the motor outputs underlying orientation behavior (Stein and Gaither, 1981; Knudsen, 1982; Vanegas, 1984; du Lac and Knudsen, 1990; Stein and Meredith, 1993; Valentine et al., 2002). It is partly homologous to the mammalian superior colliculus (SC) in that both structures share connection patterns, cell types, physiology, and a highly stratified anatomy (Vanegas, 1984; Luksch, 2003). Thus many of the findings in the avian OT may be applicable to the SC and contribute to the general understanding of multisensory integration and saccade generation in vertebrates.

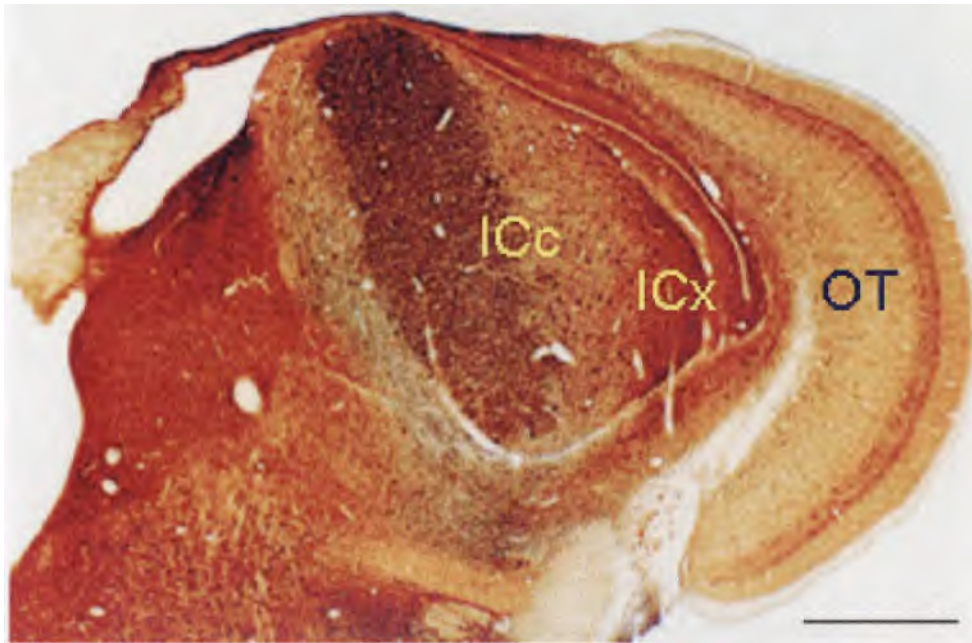


Figure 1-1. A slice from the barn owl midbrain. The slice shows immunoreactivity to calbindin. The optic tectum (OT), external nucleus of the inferior colliculus (ICx), and central nucleus of the inferior colliculus (ICc) are labeled respectively. Only one half of the midbrain is shown, with dorsal to the right and up, and ventral to the left and down. The scale bar at the lower right is 1 mm in length. (Courtesy Terry Takahashi, 2006)

The overall anatomy of the OT is similar to that of a halved onion in that it is curved and multilayered, located bilaterally at the level of the midbrain (Figure 1-1). The midbrain slice shown in Figure 1-1 has been stained for calbindin. Although the role this calcium-binding protein plays in auditory processing and saccade generation is not fully realized at this time, this particular staining illustrates the laminar structure of the OT as well as the distinct sub-nuclei of a key processing area for the computation of auditory space (inferior colliculus – IC, see below). In the barn owl, the deepest strata of the OT begin just dorsal to the external nucleus of the inferior colliculus (ICx). The 15 individual layers of the avian OT are anatomically and physiologically distinct, with input and output layers remaining characteristically separate. The superficial and deep layers are primarily input layers, whereas

the majority of output layers are arranged intermediately (Luksch, 2003). Topographically organized retinal inputs terminate in the 7 most dorsal layers of the avian OT (Ramon y Cajal, 1995; Luksch, 2003). There also exists a strong auditory input from the auditory system (ICx) to the deeper layers of the OT, primarily layer 13 (Knudsen and Knudsen, 1983). In the dorsal-most layers, physiological responses to sensory stimuli have been shown to be primarily visual in nature. However, both auditory and visual responses have been noted in deep and intermediate layers (Knudsen, 1982).

As has been shown in mammalian studies, single unit responses in the deep and intermediate OT layers in the barn owl can be evoked by either auditory or visual inputs (Knudsen et al., 1982; King and Palmer, 1985; Stein and Meredith, 1993). Stimulation experiments in the SC and OT of both non-human primates and barn owls have shown that activity in the intermediate and deep layers is sufficient for evoking gaze-orienting movements (Robinson, 1972; du Lac and Knudsen, 1990; Sparks, 2002). Descending projections from the intermediate layers of the OT, primarily layer 11 in the owl, are known to be linked to saccade generation. They do not reach the spinal cord however, but terminate in the mesencephalic reticular formation and the pontine nucleus, areas also shown to evoke head-saccades when stimulated (Masino and Knudsen, 1992, 1993).

Auditory and Visual Sensory Processing

The avian eye comes in many shapes and sizes. The owl's eye (Figure 1-2) is classified as being abnormally large relative to the size of the skull and tubular in shape, with the long axis of the eye protruding out the eye socket. The barn owl's eye is very similar to the eagle owl's eye, as shown in Figure 2 (Walls, 1972). The extra-ocular muscles generally found in reptiles and

other birds are very short in the barn owl and only attach to the convex portion of the eye, rendering them relatively ineffective (Walls, 1972).



Figure 1-2. The tubular shape of an owl eye. The owl eye is known for its unusual tubular shape. Shown here is a photo of a roughly dissected barn owl eye (*Tyto alba*). The lens is visible through the cornea (left) and the optic nerve is just visible at the right of the picture. A more explicit diagram of this tubular shaped eye may be seen in Walls, 1972.

The barn owls' eyes are thus relatively immobile within the skull, with at most 5 degrees observable rotation (du Lac and Knudsen, 1990). Because the eyes don't move in their sockets, gaze orientation may be accurately estimated by measuring head position. Following retinal processing, projections are sent from the retina directly to the contralateral optic tectum, where axons traverse the dorsal-most layer (stratum opticum) and turn inward to terminate in a retinotopic pattern as described above. The average time from visual stimulus presentation to first spike responses in the optic tectum ranges from approximately 50 to 80 msec (Knudsen 1982; Chapter V).

The pre-tectal processing of sound follows a much more circuitous anatomical path compared to vision. Because binaural cues are used in spatial mapping, monaural information from the two ears must be compared and combined. The external ear flaps, the shape of the ear canal, and the morphology of acoustically reflective facial feathers (facial ruff) act as filters to incoming sounds (Figure 1-3). Spectral components are differentially attenuated and

amplified, depending on the nature of these external filters, and thus the signals that reach each eardrum are spectrally distinct, making comparisons between the ears more informative (see below).



Figure 1-3. The external ear.

The barn owl's ear is large, relative to its head, and positioned just in front of the facial ruff, a distinct subclass of stiff feathers that differentially attenuate sound frequencies. The barn owl profile shown here is oriented with the beak to the left.

Because of these anatomical features, and because of the physics of sound travel in space, each ear then receives a distinct signal, different from one another in phase and loudness. Broad band noises are broken down into their constituent frequencies at the cochlea. From each internal ear then, the monaural signals important for later determining interaural time differences (ITD - azimuthal cue) and interaural level differences (ILD) - elevational cue) are sent to the ipsilateral cochlear nuclei, nucleus angularis and nucleus magnocellularis, respectively. Nucleus magnocellularis then sends monaural timing information bilaterally to the nucleus laminaris, where binaural timing relationships are determined. Nucleus laminaris projects the newly computed binaural ITD information to the central nucleus of the inferior colliculus, where it is combined across frequencies. Meanwhile, nucleus angularis has sent monaural level information bilaterally to the lateral lemniscus (LLVp) where ILD is computed. Projections from the LLVp lead to lateral shell of the inferior colliculus, where ILD cues are combined

across frequencies and then with ITD cues in the external nucleus of the inferior colliculus (ICx, Figure 1-2). It is in the ICx that we find the first representation of auditory space, or the auditory space map (Knudsen and Konishi, 1978).

The auditory space map is known to be influenced by visual feedback in developing and adult owls (Brainard and Knudsen, 1998; Bergan et al., 2005). Thus not only does the ICx project to the deeper layers of the OT, as mentioned above, but there is a reciprocal connection back to the ICx from the OT (Feldman and Knudsen, 1997). Taking into consideration the circuitous route from the peripheral sensory organs to the optic tectum, it is somewhat surprising to find that the first spike latency at the level of the OT in response to an auditory cue ranges from 10 to 25 msec, at least 25 msec before a simultaneously presented visual response may be detected in the OT by extracellular means.

Summary of Data Chapters

Chapter II, entitled "Combined auditory and visual stimuli facilitate head saccades in the barn owl (*Tyto alba*)," presents the first detailed description of audiovisual integration in non-mammalian orienting behavior. Moreover, this study was the first to parametrically probe the influence of unisensory stimulus strength on audiovisual behavior in any non-primate species. I measured head saccades evoked by auditory, visual, and simultaneous, co-localized audiovisual stimuli whose levels ranged from near to well above saccadic threshold. In accordance with previous human psychophysical findings, the owl's saccade reaction times (SRTs) and errors to unisensory stimuli were inversely related to stimulus strength. Auditory saccades characteristically had shorter reaction times, but were less accurate than visual

saccades. Audiovisual trials, over a large range of tested stimulus combinations, had auditory-like SRTs and visual-like errors, suggesting that barn owls are able to use both auditory and visual cues to produce saccades with the shortest possible SRT and greatest accuracy. These results support a model of sensory integration in which the faster modality initiates the saccade and the slower modality remains available to refine saccade trajectory. This work resulted in a publication co-authored with Terry Takahashi (Whitchurch and Takahashi, 2006). The present author was responsible for approximately 90% of the work involved in this publication.

Chapter III, "Timing auditory and visual stimuli for optimal integration in saccadic behavior," continues the work described in Chapter II. Previously, I found that barn owls use both auditory and visual cues to localize potential prey, and often the best characteristics of each unisensory response are present in saccades to audiovisual stimuli. In Chapter III, I probed the behavioral consequence of temporally misaligning audiovisual stimuli and compare these results to make predictions based on maximum likelihood estimates. As in Chapter II, saccadic reaction times to multisensory stimuli did not differ from predictions based on a race between the two modalities, even when these stimuli were presented asynchronously. Furthermore, for most asynchronous conditions, multisensory error was most like that seen in the visual-alone condition. These findings suggest: 1.) Saccadic reaction times in a divided attention task support the theory that auditory and visual streams progress in a parallel fashion toward separate activation thresholds in the barn owl saccade generation pathway; 2.) Even in the extreme condition in which there is a high probability that one modality will trigger a saccade well before the second is presented, orientation may incorporate some of the information from the later arriving stimulus. Evidence of target updating while the bird's head

is in motion, a theory described previously (Cornell et al., 2002) and in Chapter II, in the form of double saccades was detectable only when the two modalities were largely separated in time. Therefore, saccade kinematics and localization error suggest a time window for optimal integration of the two modalities, after which single saccades may either mimic attributes of the earlier presented modality or be replaced by double-step saccades. This work resulted in a manuscript currently being prepared, co-authored with Pamela Johnston, Hannah Dold, and Terry Takahashi. The present author was responsible for approximately 72.5% of the work involved in this manuscript.

Chapter IV is titled “Neurophysiological correlates of behavioral changes with sound pressure level.” In Chapter II, saccadic responses are shown to degrade systematically as targets become quieter. For instance, response probability decreased dramatically when near-threshold sounds were presented (Figure 2-3 & 2-4). Moreover, saccades in response to quiet auditory targets were initiated much later and resulted in much greater error than saccades to louder stimuli (Figure 2-4 & 2-5). To probe the neural correlates of these behaviors, I recorded the extra-cellular responses of single units in the anesthetized barn owl’s auditory space map (ICx – see above description). The noises used in this study were identical in fine structure to the ones used in Chapters II and III, and thus said to be “frozen”. These frozen noises were used to map neuronal rate-level-functions, the first-spike-latencies, and the spatial-response-profiles at various sound pressure levels (SPLs). Fitted rate-level-functions were used to describe the relationship between a neuron’s observed spike rate and the stimulus SPL. Of the 149 cells recorded, I completed rate-level tests on 121 cells. Cellular thresholds were defined as the first statistically significant increases in firing rate across increasing SPL. All

firing rates were compared to the observed spontaneous firing rate. The cellular thresholds in these 121 cells ranged from -15 to +17 dB-SPL_A, and 41% were below -5 dB. Thus 59% of the cells recorded did not respond to sounds in this range. This corresponds with the observed decrease in the probability that a bird will respond with a head saccade to stimuli presented at or below -6 dB (Chapter II), as well as the detection threshold (data presented in this chapter). Next I observed that the neuronal first spike latency (hereafter simply referred to as "latency") was longer for increasingly attenuated sounds: The average latency to sounds presented at -5 dB was 30.7 ms, whereas the average latency of those same cells to sounds presented at 25 dB was 13.7 ms. Although this difference in latency is striking, it does not directly account for the almost 100 ms difference in SRT seen in behavior (Chapter II). Finally, there was a distinctive change in the spatial response profile with SPL. Spatial response profiles describe the average neural response to sounds as a function of the sound's spatial location. With the use of virtual auditory space technology, I randomly presented sounds of varying SPL from a large number of spatial locations to a subset of neurons. I found that when quiet sounds were used, neurons responded more vigorously to locations that were more centralized than peripheral. This essentially translated into a predictable "shift" in the spatial response profile, or receptive field (RF). Although the behavioral implications of this apparent shift are not known, this change in spatial response may be correlated with the increase in behavioral error (the greater undershoot, see Chapter II) seen in response to quieter sounds. The present author is responsible for 98% of the work described in this chapter. The remaining was performed by Dr. Avinash Bala in a collaborative effort to determine the barn owl's detection threshold using the pupil dilation response (PDR).

Chapter V, "Audiovisual integration in the optic tectum of the barn owl," summarizes data from extracellular recordings of single unit responses to auditory, visual, and audiovisual stimuli in an anesthetized preparation. Cellular latencies to loud auditory stimuli ranged from 8 to 25 ms, whereas those to bright visual stimuli were longer, ranging from 70 to 120 ms. Audiovisual first-spike latencies approximated auditory latencies across SPLs.

Unisensory spike rates to stimuli presented in the peak area of the cell were obtained and summed to predict audiovisual spike rates. Observed audiovisual spike rates were then compared to these predicted linear values and found to range from near-linear to supra-linear in nature. Interestingly, for one cell, the initial spike rate of the audiovisual response was significantly higher than the initial spike rate of the auditory response, suggesting that the visual stimulus affected spike rates prior to what was temporally predicted based on the visual latency. These findings were not consistent across cells, nor were they consistent within a given cell. I attribute this variability to the fact that nitrous oxide was used to maintain a light state of anesthesia. As was evident by the position of the facial feathers and the movement of the bird, I suspected that the birds' state of arousal fluctuated throughout the course of individual experiments. The optic tectum is strongly connected with the forebrain and such fluctuations in behavioral states would undoubtedly affect the nature of audiovisual integration in the OT, as well as the responsiveness to stimuli in general. Future recordings from awake birds will lead to more concrete assertions regarding the neural basis of the audiovisual integration observed in behavior. The present author is responsible for 100% of the work described in this chapter.

CHAPTER II

COMBINED AUDITORY AND VISUAL STIMULI FACILITATE HEAD SACCADDES
IN THE BARN OWL (*TYTO ALBA*)**Introduction**

This chapter was published and produced in collaboration with Terry Takahashi (Whitchurch and Takahashi, 2006). I completed 99% of the experimental design, 100% of the data collection, 100% of the analysis, and 70% of the writing.

Objects in nature are often both visible and audible, and studies have shown that adding an auditory component to a visual stimulus decreases reaction time (Todd, 1912; Raab, 1962; Miller, 1982; Welch and Warren, 1986; Engelken and Stevens, 1989; Hughes et al., 1994; Corneil and Munoz, 1996; Harrington and Peck, 1998; Diederich et al., 2003; Colonius and Diederich, 2004) and improves detection and orientation (Stein et al., 1988; Stein et al., 1989; Wilkinson et al., 1996; Corneil et al., 2002; Jiang et al., 2002). How are audition and vision, two peripherally independent senses, integrated to enhance behavior?

A certain amount of improved performance in multisensory tasks may be due to simple probability. In the same way that flipping two coins instead of one increases the chances of obtaining at least one “head”, the probability of either hearing or seeing a weak multisensory stimulus is greater than the probability afforded by either modality alone. Raab (1962) extended this idea of statistical facilitation to reaction times. According to his “race model”,

the two sensory signals race along separate pathways toward a common response generation site, and the winning modality triggers the response. If stimulus conditions are such that the two modalities have an equal chance of winning the race, the race model predicts that the multisensory reaction time will be, on average, 0.6 standard deviations earlier than the mean unisensory reaction time (Raab, 1962). Multisensory reaction times that are quicker than predicted are said to *positively* violate the race model (Miller, 1982; Hughes et al., 1994). Positive violations have been taken as evidence that instead of remaining separate, the two modalities “converge”, giving rise to reaction times that are quicker than the race model allows (Hughes et al., 1994; Nozawa et al., 1994; Harrington and Peck, 1998; Hughes et al., 1998; Arndt and Colonius, 2003). The convergence model is thus distinguished from the race model by its assumption that the two modalities converge to either form a new internal sensory representation, enhance the efficiency of the ensuing motor response, or both.

The study of audiovisual responses in mammalian systems has elucidated three general principles of multisensory integration: for review, (Stein and Meredith, 1993; Calvert et al., 2004; Spence and Driver, 2004). First, when the auditory and visual components of an audiovisual stimulus are aligned in space, the probability of correctly localizing the target is increased relative to trials in which they are misaligned (Stein et al., 1988; Stein et al., 1989; Jiang et al., 2002). Conversely, spatial misalignment lengthens saccadic reaction times beyond what would be seen with one stimulus alone (Corneil and Munoz, 1996; McSorley et al., 2005). Second, auditory and visual *information* must be aligned in time. If the timing of the auditory and visual stimuli compensates for the difference in processing times between the two modalities, facilitated behavioral (Corneil et al., 2002) or neural responses (Meredith

et al., 1987) are more likely. Third, multisensory facilitation is more likely to occur when the unisensory components are presented at low amplitude or signal-to-noise ratio. This has been termed “inverse effectiveness” (Stein and Meredith, 1993). Most recently, Stanford et al. (2005) showed that neural responses to audiovisual stimuli were most likely to be super-additive (i.e. facilitated beyond the predicted linear sum of the unisensory components) when near-threshold unisensory stimuli were presented. Similarly, feline detection/orientation behavior in response to audiovisual stimuli was most enhanced when unisensory stimuli of low efficacies (dim and quiet) were combined (Stein et al., 1989). Moreover, the saccades of human subjects to auditory, visual, and audiovisual stimuli in the presence of background auditory and visual noise were initiated as rapidly with SRTs typical of unisensory auditory trials and high accuracy typical of unisensory visual trials (Cornel et al., 2002). To my knowledge, this tendency to observe the “best” features of acoustically- and visually- guided saccades had not been previously noted in studies using higher signal-to-noise ratios.

The present study takes an ethological approach to multisensory integration, using the barn owl (*Tyto alba*) as a model. When the owl hears or sees an object of interest, it naturally turns its head to face the object. This stereotyped saccadic behavior, like primate eye movements, can be measured with precision (Knudsen et al., 1979; Knudsen and Konishi, 1979).

Moreover, the barn owl, renowned for its ability to localize sound, has a precise spatiotopic map of auditory space in the external nucleus of its inferior colliculus (ICx) that projects topographically to the retinotopic map in its optic tectum, generating a physiologically accessible multisensory map of auditory and visual space (Knudsen and Konishi, 1978;

Knudsen, 1982). The barn owl thus provides us with a model system in which to explore the integration of information from two modalities in the context of its natural behavior.

Here, I measured saccadic reaction time (SRT) *and* saccadic error in auditory, visual, and simultaneous, spatially aligned audiovisual conditions. The targets spanned a wide range of stimulus locations and strengths, including those unisensory stimuli that were significantly less effective in eliciting fast, accurate saccades. For most audiovisual combinations, SRTs were no earlier than the earlier of the two modalities, typically, audition, and were no more accurate than the more accurate of the two, typically, vision. However, saccades to audiovisual targets had the short latencies typical of saccades to auditory targets and the accuracy typical of those to visual targets, thus revealing the best features of both modalities, in agreement with a previous human audiovisual study (Corneil et al., 2002). These results are not consistent with the race model which predicts that the winning modality would *determine both* the SRT and error. Instead, these results support an “updating” hypothesis of sensory convergence in which the first modality to reach threshold triggers the saccade, while information from the second modality is available to update and refine the target position (Corneil et al., 2002; Van Opstal and Munoz, 2004).

Methods

All experiments were carried out under protocols approved by the University of Oregon Institutional Animal Care and Use Committee and in compliance with *The Guide for the Care and Use of Laboratory Animals*, (Institute of Laboratory Animal Resources (U.S.) and NetLibrary Inc., 1997).

Apparatus

All experiments were conducted in a double-wall, sound-isolating anechoic chamber (Industrial Acoustics Company, IAC), the properties of which have been described previously (Spitzer et al., 2003). The ambient noise inside the chamber was below 18 dB SPL_A, between 2 and 10 kHz. Head-aim was recorded and sampled at 468 Hz with a custom-built magnetic search coil system (Remmel Labs, Ashland, MA). The head coil was mounted on a post that was cemented to the skull prior to training. The head coil system was calibrated over a $\pm 50^\circ$ range in azimuth and elevation at the beginning of each session. This compensated for any day-to-day changes in the imposed magnetic field and enabled the system to remain accurate within 1° to 2° .

Acoustical Stimuli

One of thirty pre-synthesized, 100 msec auditory stimuli was randomly presented in each auditory or audiovisual trial. Stimuli were broad band noises, flat between 2 and 12 kHz and differed only in fine structure. The noises were digitally synthesized, ramped on and off by a 5 msec cosine envelope, and sampled at 30 kHz (Tucker-Davis Technologies PowerDAC PD1). The analog signal was fed to a programmable attenuator (PA4, Tucker Davis Technologies), amplified (HB6, Tucker-Davis Technologies), and presented through one of seven lightweight speakers (Peerless, 5.08 cm cone tweeters). The speakers' frequency response characteristics were flat (within 3 dB) between 2 and 12 kHz. Auditory levels ranged from -12 dB to 6 dB, in increments of 3 dB on an A-weighted SPL scale. The lower levels were extrapolated from attenuation/SPL relationships recorded at detectable levels

with a ½ inch microphone (Brüel & Kjaer model 1760) and sound-level meter (Brüel & Kjaer model 2235).

Visual Stimuli

The visual stimulus was a single, stationary 100 msec flash from a red (635 nm) light emitting diode (LED), presented in the otherwise completely darkened, light-sealed chamber.

Luminous intensities were manipulated by ranging the applied voltage. Six intensities were used (5.2 cd, 0.052 cd, 0.0053 cd, 0.0013 cd, 0.00077 cd, and 0.00031 cd). Each LED was suspended directly in front of one of the seven speakers.

Behavioral Paradigm

Two adult barn owls (*Tyto alba*) were hand-reared and trained to make head saccades to auditory, visual, and simultaneously presented audiovisual stimuli. Because the owl's eyes are relatively immobile (du Lac and Knudsen, 1990), gaze was estimated by head-aim. Stimulus synthesis, data acquisition, and data analysis were all carried out with Matlab 6.5 (The Mathworks). Animals were given food rewards and maintained at 85% free-feeding body weight.

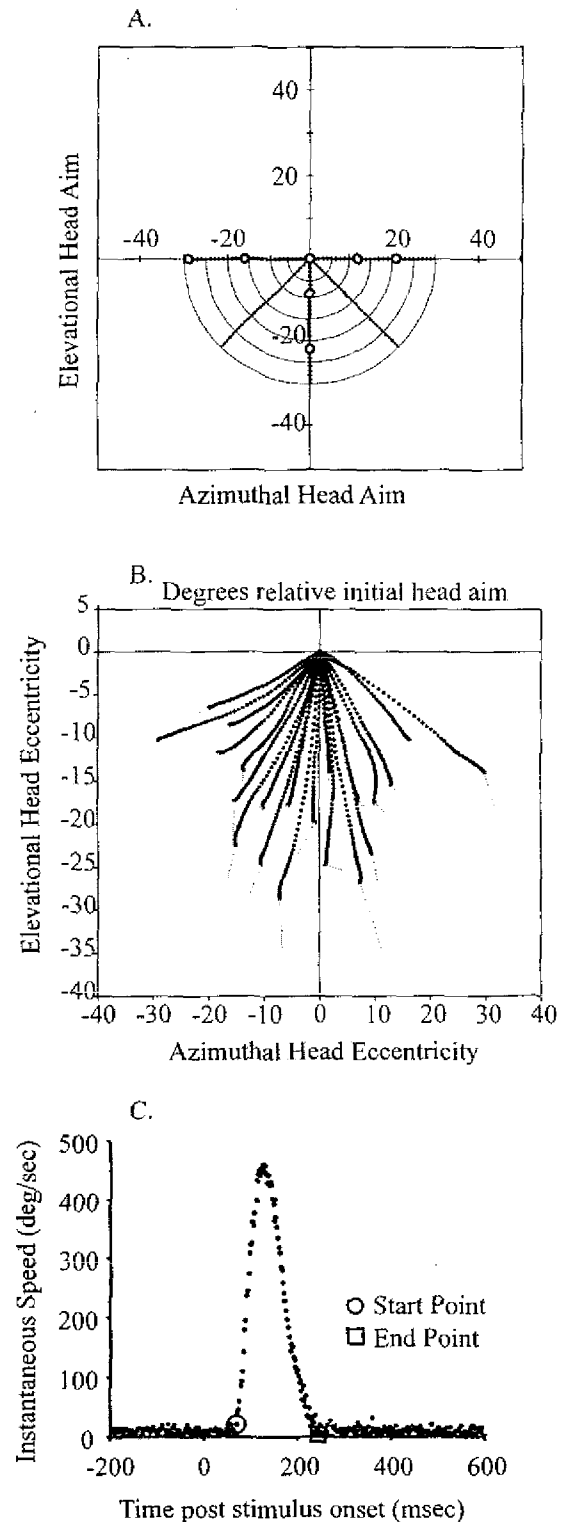
Trials were conducted while the bird was tethered to a perch located near the center of an isolated, completely darkened room. Food rewards (mouse bits) were presented from a remotely controlled feeder located in front of the bird, at the level of the perch. Each session, seven targets were pseudo-randomly positioned on either a "T" shaped or

“inverted-V” shaped frame, separated by at least 7° (Figure 2-1A, black and grey orientations, respectively).

Figure 2-1. Head saccade measurements.

A) Targets were presented in the lower hemi-field on either a frame shaped like a “T” (black) or an inverted “V” (dark grey). The circles represent speakers, each with an LED at its center. The speakers are shown on the T-shaped frame, but were also placed on the inverted-V frame on different days in order to increase the uncertainty about speaker location from session to session. During each session, 7 targets were arranged within 30° deg of the origin (i.e., the apex of the inverted-V frame, or the intersection of the stem and arms of the T frame) and separated by at least 7° deg. All trials accepted for analysis were triggered and terminated within the calibrated region, which subtended 50° deg from the origin (full area of the graph). **B)** Example saccades, from trials with bright visual stimuli, are shown, (data from bird ‘N’ and bird ‘J’). Saccade eccentricity is plotted relative to initial head position. Light grey lines show distance between saccade end-position and target. **C)** A typical 40° deg head saccade is plotted in black, (data from bird ‘J’). Latency points (circle) and end points (square) were determined by velocity criteria (see Methods).

The frame was alternated between the “T” and “inverted-V” orientations on a session by session basis. Target locations ranged from -30° to 30° in azimuth and 0° to -30° in elevation, with 0° azimuth and 0° elevation being at the intersection of the owl’s midsagittal plane and eye-level.



Birds were trained to localize both auditory and visual stimuli over the course of approximately 4 months. During initial training, only clearly audible and visible stimuli were used (30 dB, 5.2 cd), and the birds were rewarded for localization attempts directed toward the quadrant of space containing the target. As performance gradually improved, the reward criteria became more stringent. Eventually, the bird localized clearly audible and visible stimuli within 5° on more than 90% of the trials. Weaker stimuli (lower SPL or luminous intensity) were then introduced. The training period ended when the accuracy and precision of the birds remained stable for each given modality and stimulus-strength. Data from the training period are not reported here.

The 11 months following training yielded 97 and 94 test sessions for birds 'N' and 'J' respectively. There were 40 possible rewards within a given session. The session ended when all rewards were dispensed. Sometimes the bird's performance would warrant a reward on every trial. In these sessions, there were only 40 trials. Other times, the bird's behavior fell short of the reward criteria and would not be rewarded for up to 15 trials. Such sessions were ended at trial number 65. Analysis included both rewarded and unrewarded trials. In general, rewards were given for head turns terminating within 5° of the target. Trials with very low level stimuli, however, were rewarded when localization attempts were directed in the quadrant of the target, so as to maintain motivation. Unlike earlier head-turn experiments in owls (Knudsen et al., 1993; Poganiatz et al., 2001), I did not require the initial head-aim to be within some radius of a central starting location. Instead, a trial was manually triggered when the bird's spontaneous head turns ceased and its head came to rest, typically within 40° of being aligned with its body, as seen with an infrared camera. Post-hoc analysis

discarded any trials during which the head was moving faster than $10^\circ/\text{sec}$ prior to stimulus onset, thus avoiding saccades with head motion immediately prior to and during stimulus presentation. The angular distance to the target is expressed relative to the initial head-aim and is referred to as *target eccentricity*.

Head-aim was recorded for 3 seconds, a time-span comprising a 200 msec pre-stimulus window, the 100 msec stimulus, and a 2,700 msec post-stimulus window. The inter-stimulus interval varied from 15 to 90 sec. A total of 9,872 trials were recorded, 2,779 of which were excluded from analysis, as described below.

Of excluded trials, 66% were those in which the head moved more than 2° in azimuth or elevation in the 200 msec pre-stimulus window, or before 40 msec post-stimulus onset.

Head turns initiated before 40 msec post-stimulus onset were almost never directed towards the target and thus not stimulus driven. Trials were also excluded if, at any point during the 3 second recording, the head was outside the calibrated region of the coil system ($100^\circ \times 100^\circ$ perimeter, Figure 2-1A). Seventeen percent of excluded trials were outside of the calibrated region. The remaining exclusions were those trials in which the targets were located above the initial head-aim and/or either too close or too far away; In training and preliminary testing, both birds showed a bias against responding to stimuli closer than 12° and farther than 55° away, as well as those targets presented in the upper visual hemi-field.

The 7,093 accepted trials were categorized as responses or non-responses. Non-responses were relevant in measuring the probability of response, especially in the trials with very low

stimulus intensities. Each trial contained an auditory, visual, or audiovisual stimulus, and therefore any head turn with a SRT falling between 40 and 600 msec of stimulus onset was scored as a response. More than 99% of scored responses were directed toward the quadrant of space containing the target. Those trials in which the SRT was greater than 600 msec were instances in which the birds typically did not make any head turn at all and were thus scored as non-responses.

I tested 7 sound pressure levels (SPLs - henceforth also referred to as levels for simplicity) and 6 luminous intensities, resulting in 13 unisensory conditions and 42 audiovisual combinations. For each bird, 70 – 120 trials were recorded for each unisensory condition and 30 – 60 trials for each audiovisual condition. In each session, trials varied randomly in stimulus-strength, location, and modality (auditory, visual, and audiovisual). Each session had a random proportion of unisensory and multisensory trials; however the average session consisted of 30% unisensory and 70% multisensory trials. The luminances, SPLs, and combinations of luminance and SPL also pseudorandomly varied on a session by session basis, and each session included very strong as well as very weak stimuli.

Data Analyses

Saccade Reaction Time (SRT)

Figure 2-1B shows 20 randomly selected, visually evoked, saccades from both birds 'N' and 'J'. Instantaneous speed from one 40° saccade (Figure 2-1C) was computed by dividing the distance between consecutively sampled locations by the time between sampling points.

SRT was defined relative to stimulus onset as the first point in time at which the

instantaneous head speed continuously exceeded two standard deviations of the average speed measured in the 200 msec pre-stimulus window (Figure 2-1C, open circle).

Preliminary analysis showed a negative correlation between target eccentricity and SRT for auditory trials in bird 'N' ($r(566) = -0.1228, p < 0.01$) and bird 'J' ($r(550) = -0.3126, p < 0.0001$). However, correlations between SRT and target eccentricity were not significant in visual trials for bird 'N' ($r(579) = 0.005, p = 0.9135$) and bird 'J' ($r(604) = -0.0627, p = 0.1608$). No significant correlation was found in audiovisual trials between SRT and target eccentricity for bird 'N' ($r(1514) = -0.0087, p = 0.73$), but this correlation was significant in audiovisual trials for bird 'J' ($r(1558) = -0.1465, p < 0.0001$). Accounting for this correlation in an analysis of covariance, however, did not affect the overall trends reported below.

Thus, trials were pooled across target eccentricities in all further SRT analysis. An analysis of variance (ANOVA) was used to test for main effects of stimulus-strength on SRT in unisensory conditions. Mean SRT changed across conditions in a sigmoidal fashion and was fit by the least squares method for display. *A-priori* pair-wise comparisons (auditory vs. audiovisual, visual vs. audiovisual) of mean SRTs were conducted utilizing the Bonferroni correction for non-orthogonal test design ($\alpha = (0.05/\text{number of comparisons})$).

Observed audiovisual SRT distributions were compared to theoretical race model distribution boundaries computed assuming: 1.) The auditory and visual information remain separate. 2.) Either the auditory OR the visual stream may trigger a response. 3.) The first signal to arrive at the motor control center will determine the SRT (Raab, 1962). In addition, any particular model from the class of models based on these basic 'race-model' assumptions

will also assume a certain level of unisensory dependence, or independence (Hughes et al., 1994).

Regardless of the level of dependence however, the upper boundary of SRT facilitation is statistically predicted by the *Miller Inequality*, which sums the unisensory cumulative probabilities at each SRT (Miller boundary – Eqn. 2-1, Figure 2-2, dot-dash line) (Miller, 1982; Hughes et al., 1994; Townsend and Wenger, 2004):

$$P(T_{\min} \leq t \mid AV) \leq P(T_A \leq t) + P(T_V \leq t), \quad \text{Eqn. 2-1}$$

where t is time relative to stimulus onset, and T_A and T_V are times associated with processing auditory and visual streams, respectively, drawn randomly from the distribution of unisensory trials. T_{\min} is the smaller of the two values, T_A and T_V . If a significant proportion of the observed multisensory SRT distribution falls to the left of this upper boundary of the race-model predictions, the distribution *positively* violates the race model; the response time is shorter than statistically expected, were the birds to act on the signal contributed from the earlier-arriving modality.

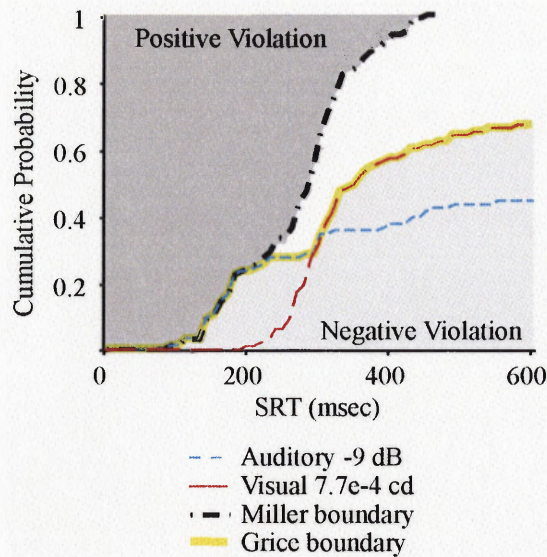


Figure 2-2. Cumulative probability distributions predicted by the race model. The dot-dashed line shows the Miller boundary, which is the upper boundary of any race model prediction. When audiovisual SRT distributions are significantly to the left of the Miller boundary, SRTs are generally shorter than predicted by the race-model (dark grey area). The yellow line shows the Grice boundary, which is the lower boundary of the race model. The race model is said to be negatively violated when the observed audiovisual SRT distributions are significantly longer than race model predictions (light grey area). Actual auditory and visual SRT distributions are shown for near-threshold conditions from bird 'J'.

The lower boundary of race-model SRT predictions can be found by using the *Grice inequality*, which takes the maximum unisensory probability at each SRT value (Grice boundary – Eqn. 2-2, Figure 2-2, yellow line) (Grice et al., 1984; Hughes et al., 1994; Townsend and Wenger, 2004):

$$P(\Gamma_{\min} \leq t \mid AV) \geq \max [P(\Gamma_A \leq t), P(\Gamma_V \leq t)]. \quad \text{Eqn. 2-2}$$

Audiovisual SRT distributions showing a significant proportion of their distribution to the right of the Grice boundary are considered examples of *negative* race-model violations, in which the SRTs are longer than predicted by any version of the original race model.

Race model predictions and observed audiovisual SRT distributions were compared using a one-tailed, one-dimensional, two-sample Kolmogorov-Smirnov test (KS-test) at $\alpha = 0.05$. Bonferroni corrections were applied for multiple comparisons. Observed distributions were subtracted from theoretically predicted distributions and plotted as difference graphs (insets, Figures 2-8, 2-9). Significantly positive differences between the Miller boundary and observed SRT distributions indicate conditions in which the observed audiovisual SRT was shorter than predicted by the race model. Significantly negative differences between the Grice boundary and observed SRT distributions represent conditions in which the observed audiovisual SRT was longer than predicted by the race model.

Response Percentage

Response percentage was defined for each condition as the percentage of accepted trials in which there was an obvious response (see SRT criteria above) toward the target within 600 msec of stimulus onset. The vast majority of trials in which the bird did not respond before 600 msec had no obvious saccade throughout the 3 sec recording period. Theoretical audiovisual response percentages were predicted from a probabilistic combination of the observed unisensory response percentages (Eqn. 2-3):

$$P_{AV} = (P_A + P_V) - (P_A \times P_V), \quad \text{Eqn. 2-3}$$

where P_{AV} , P_A , and P_V are the probabilities for responding to an audiovisual, auditory, and visual stimulus respectively.

Localization Error

The end point of a saccade was designated when the head speed dropped below 5°/sec, (Figure 2-1). Across saccades, this criterion for head speed accurately delineated the end position of the saccade, (e.g. Figure 2-1C). Errors in azimuth and elevation were measured, respectively, as the horizontal and vertical distance between the saccade endpoint and the target. Total error was computed by vector summation of the horizontal and vertical components.

ANCOVA's were conducted to assess total errors co-varying with target eccentricity, as a function of stimulus-strength. Slopes and marginal means for the fitted lines are reported for each condition. The marginal mean error is the mean of the total error, adjusted for the linear relationship between total error and target eccentricity. The marginal mean errors for each condition were computed at a value of 28° from the target, the general midpoint of tested target eccentricities. Marginal means were plotted against stimulus-strength and fit by least-squares with a sigmoid. Pair-wise comparisons of marginal means were conducted for unisensory conditions, with Bonferroni corrections at $\alpha = 0.05$.

Results

I analyzed 7,093 head turns toward auditory, visual, and audiovisual stimuli presented in the lower frontal hemisphere across a wide range of stimulus intensities. In audiovisual trials, the light and sound had synchronous onsets and equal durations and were presented from the same location. Data from the two birds, 'N' and 'J', are individually reported and plotted, and the trends described herein are not significantly different between birds, unless

otherwise noted. The statistical significance of each analysis was determined separately for each bird (Figures 2-1 – 2-10), with the exception of the multivariate analysis illustrated in Figure 2-11.

Responses to Unisensory Stimuli

Auditory and visual unisensory trials were randomly interleaved with multisensory trials in each daily test session. The measures of response percentage, SRT, and error were monitored for each head saccade and individually reported below. The effectiveness of stimuli at evoking a response of the seven auditory levels and six visual intensities ranged from near to well above threshold. At the highest auditory and visual stimulus levels, both birds responded at least 99.9% of the time. The weakest auditory stimuli were minimally effective, in that both birds responded to the stimulus only 20% of the time. The weakest visual stimulus evoked a response 75% and 35% of the time for birds 'J' and 'N' respectively. Predictably, response percentage in unisensory trials increased as auditory and visual stimulus-strength increased (Figure 2-3).

Mean unisensory SRTs decreased toward an asymptote as auditory and visual stimulus-strength increased (Figure 2-4A, C, respectively; lighter shades = weaker stimuli). At the highest levels tested, SRTs of auditory trials were significantly shorter than those of visual trials (bird 'N': auditory: 72 ± 2 msec; visual: 146 ± 2 msec; *t*-test, $p < 0.001$).

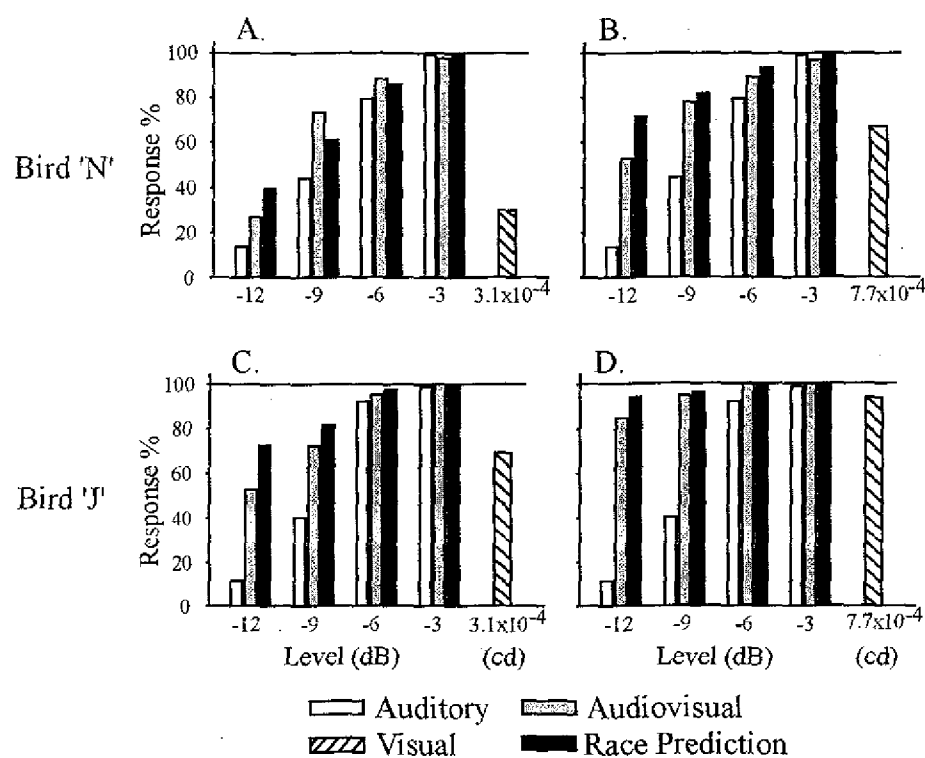


Figure 2-3. Effects of stimulus strength on response percentage.

Data from the two birds are shown (bird 'N', A&B; bird 'J', C&D). Response probabilities obtained in auditory (white bars) and audiovisual (grey bars) trials are plotted for the SPLs tested at the two weakest luminances tested (cross-hatched bars, 0.00031 cd, A&C; 0.00077 cd, B&D). The black bars represent the predictions of the race model. Observed audiovisual response percentage was either equal to or greater than unisensory values. Observed response percentage generally followed race model predictions for most stimulus pairs.

Figures 2-4B and 2-4D show the cumulative probability distributions for auditory and visual SRTs for each condition. For both modalities, the variance, which may be visualized by the slopes and ranges of the cumulative distribution functions, increased for weaker stimuli. Response percentages, taken from the cumulative response probability at 600 msec post stimulus onset, also decreased systematically with stimulus-strength.

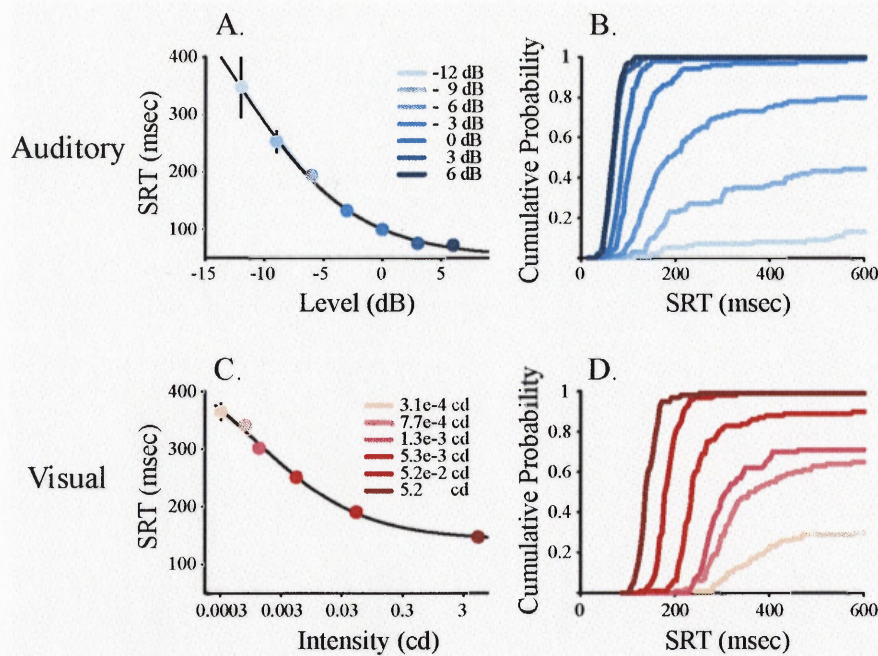


Figure 2-4. Effects of stimulus strength on saccade reaction time (SRT). Mean (\pm SEM) auditory SRTs from bird 'N' are shown (A) to decrease with increasing SPL, (Data shown bird 'N': $F(6,560)=84.69, p < 0.001$; Data not shown bird 'J': $F(6,526)=73.49, p < 0.001$;). Mean visual SRTs (C) also decreased with increasing intensity, (bird 'N': $F(5,466)=188.07, p < 0.001$; bird 'J': $F(5,487) = 21.14, p < 0.001$;). Cumulative probability distributions obtained in auditory (B) and visual (D) trials show that distribution variance increased and the probability of response decreased stimulus strength decreased. The response percentage is the cumulative probability at 600 msec after stimulus onset, multiplied by 100.

Figure 2-5 shows the analysis of the head-turn accuracy to auditory (A-C) and visual (D-F) targets. Figures 2-5A and 2-5D demonstrate that saccade error depended on target eccentricity (angular distance to the target, relative to the initial head aim; See Methods) as well as stimulus strength. I therefore report the marginal means (B, E) and the slopes (C, F) of the regressions shown in Figures 2-5A and 2-5D. Just as SRT and response percentage improved systematically with stimulus strength, saccade error decreased as SPL and luminous intensity increased (B, E).

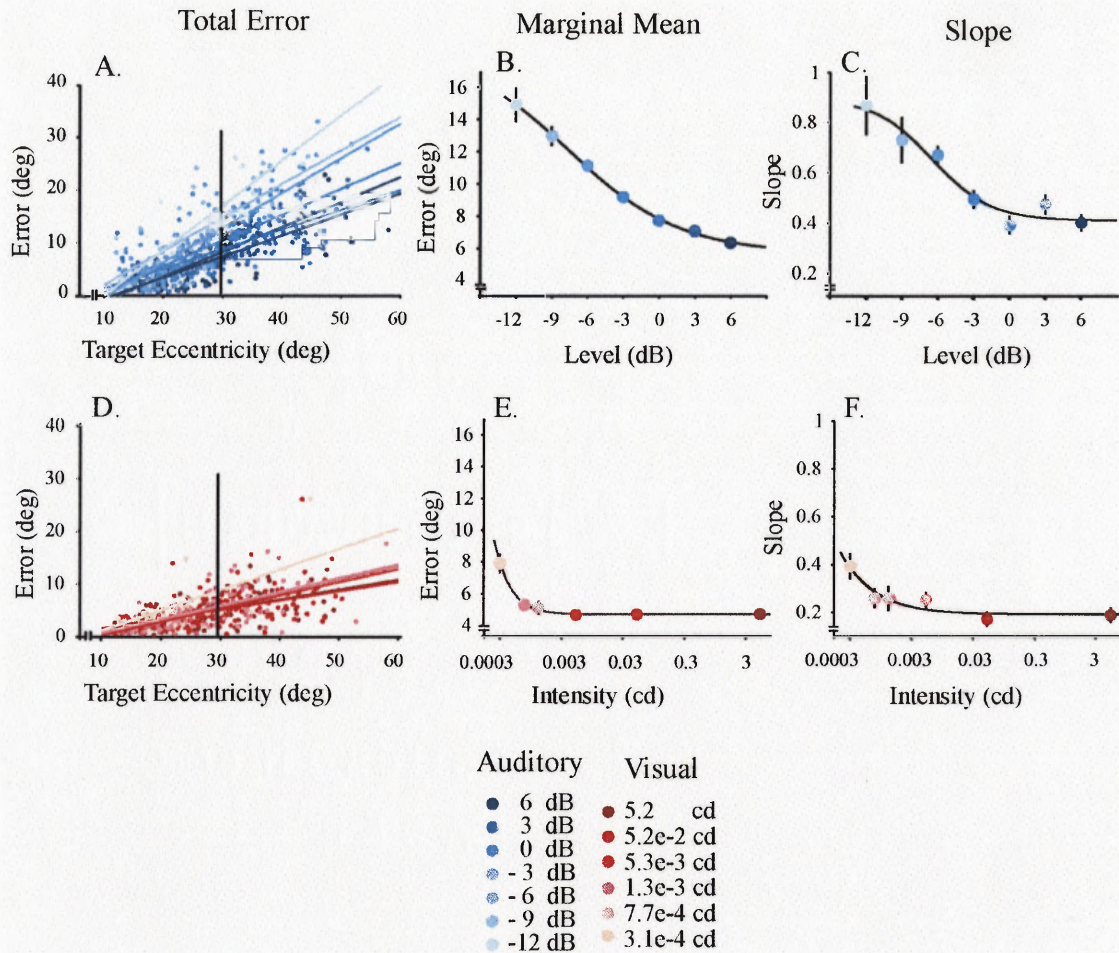


Figure 2-5. Effects of target eccentricity and stimulus strength on saccadic error. Panels A and D show that auditory and visual saccades to more distant targets had greater errors (solid lines: regression fits). An ANCOVA analysis, with initial distance to the target (eccentricity) as a covariate, verified that error depended on stimulus strength for both auditory and visual conditions (Data shown bird 'N' -- Aud: $F(6,560)=31.97, p < 0.001$, Vis: $F(5,466)=3.8, p < 0.01$; Data not shown bird 'J' -- Aud: $F(6,526)=22.98, p < 0.001$, Vis: $F(5,487)=7.47, p < 0.001$). Marginal mean errors were computed for the mean target eccentricity (vertical line at 28°) and plotted in B&E. Auditory marginal error (B) decreased as a function of increasing SPL (\pm SEM). Only the dimmest visual stimulus produced a significantly different marginal error (E, pair wise comparison, Bonferroni corrected, $p < 0.05$). All but the dimmest visual conditions gave rise to saccades that were more accurate than those of the auditory conditions (t -test of marginal means and slope, Bonferroni corrected, $p < 0.05$). The slopes from the regression fits shown in A&D are plotted in C&F. The degree to which total error depended on target eccentricity decreased as a function of stimulus strength (\pm SE of slope). For the majority of conditions, visual saccades were less dependent on target eccentricity.

The error for visual trials, however, reached an asymptote more quickly than those of auditory trials for the range of magnitudes tested. The degree to which error depended on target eccentricity also diminished as stimulus strength increased (slope; Figure 2-5C, F).

To better understand the nature of the errors, saccade total error was divided into its azimuthal and elevational components. Figure 2-6 shows the azimuthal and elevational components of saccade length for strong and weak unisensory stimuli plotted against target eccentricity. Perfect performance in either dimension would thus yield points along the unity slope line (dashed black line). Stronger auditory stimuli, (black dots, see Figure 2-4 for relative stimulus effectiveness), consistently resulted in accurate saccades (Figure 2-6A, B). Interestingly, when auditory targets of small eccentricity and high SPL were presented, this bird had a distinct overshoot in elevation (Figure 2-6B). The regression line for the azimuthal and elevational components of the saccades to the weaker auditory stimuli (grey circles) tended to have a shallower slope compared to the regression line for the louder condition and to the unity line. As stimulus-strength decreased, this tendency to undershoot was more pronounced for elevation than for azimuth (Figure 2-6A vs. B, but note differences in scales). Furthermore, the scatter about the regression line was larger for elevation than for azimuth, especially for the quieter sound (Figure 2-6, r^2 values in insets). Thus accuracy and precision in both azimuth and elevation were better with louder stimuli.

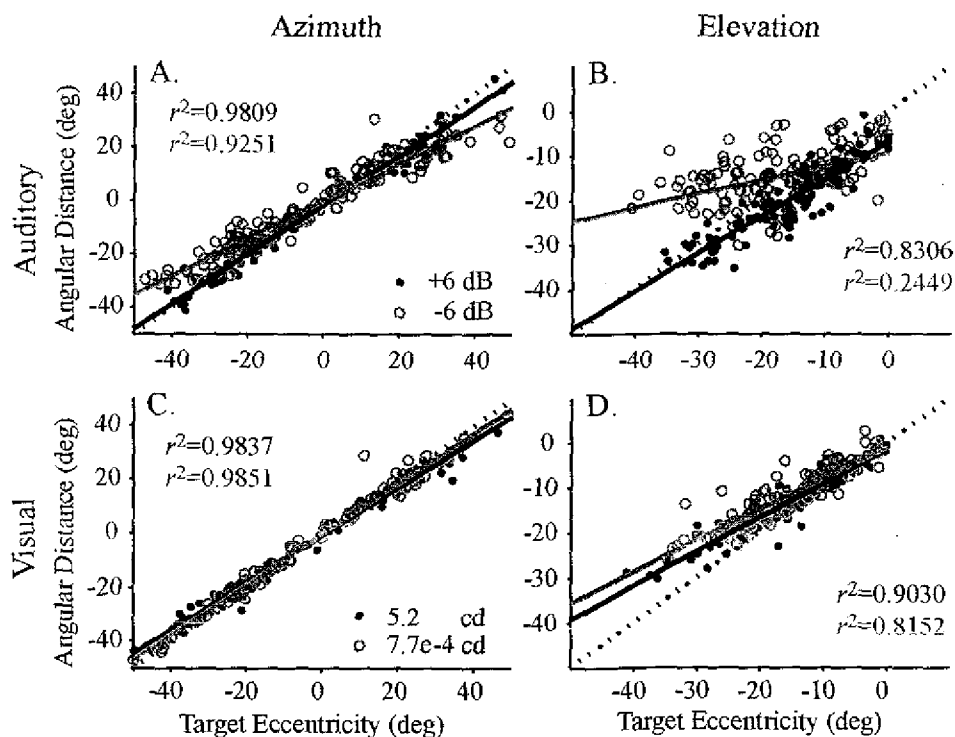


Figure 2-6. Effects of stimulus strength on azimuthal and elevational components of error. A&B Saccade length (angular distance) is plotted as a function of initial distance to target (target eccentricity) for relatively loud (6 dB, black solid) and quiet (-6 dB, grey open) auditory trials. Perfectly accurate azimuthal or elevational components of a saccade fell along the unity line (dashed line). Regression lines (solid lines) show that in quieter auditory trials, under-shooting in both dimensions was common. This undershooting was exaggerated in the elevational component and more prevalent for larger target eccentricities. **C&D** No significant difference was found between error trends of saccades to bright (5.2 cd, black solid) versus the dimmer (7.7e-4 cd, open grey) visual stimuli. Data from bird 'J'.

In contrast, for visual intensities that elicited equivalent response percentages, there was no significant difference between head turns to dim and bright stimuli in azimuth or elevation (compare open vs. closed circles in Figure 2-6C, D; Note: The dim visual condition plotted here was not the dimmest condition tested). There was, however, a slightly greater tendency to undershoot in elevation than in azimuth. These results are also consistent with the data in Figures 2-5E – F, which shows a significant increase in error for only the dimmest stimulus (at 0.00033 cd, pair-wise Bonferroni corrected, $p < 0.05$). Thus, while visual SRT increases systematically with decreasing stimulus-strength, visual accuracy remains virtually unaffected.

Responses to Audiovisual Stimuli

Below, I compare results from the multisensory trials with those of unisensory trials.

Measured audiovisual response percentages, SRTs, and errors are then compared to race model predictions.

When weaker auditory and visual stimulus conditions were combined in audiovisual trials, the probability of observing a head turn often increased as compared to unisensory levels (Figure 2-3, grey). Given the auditory and visual unisensory response probabilities, predicted audiovisual response percentages were computed assuming the race model (Eqn. 2-3) and plotted (Figure 2-3, black). In general, the observed probability of responding to an audiovisual stimulus was indistinguishable from the theoretical probability of responding to the light, the sound, or both. Only one stimulus combination yielded a response probability that was noticeably higher than predicted out of 42 conditions for 2 birds (Figure 2-3A, -9 dB, 3.1×10^{-4} cd, bird 'N').

Saccade Reaction Times in Audiovisual Trials

Mean audiovisual SRTs typically followed the shorter of the two unisensory SRT means. In Figure 2-7A – F, the mean SRT (\pm SEM) is plotted against SPL for auditory (short-dash) and audiovisual trials (solid). The long-dashed horizontal line in each graph indicates the mean visual SRT (\pm SEM) for the luminance shown. The figure thus represents 42 audiovisual combinations (7 auditory; 6 visual).

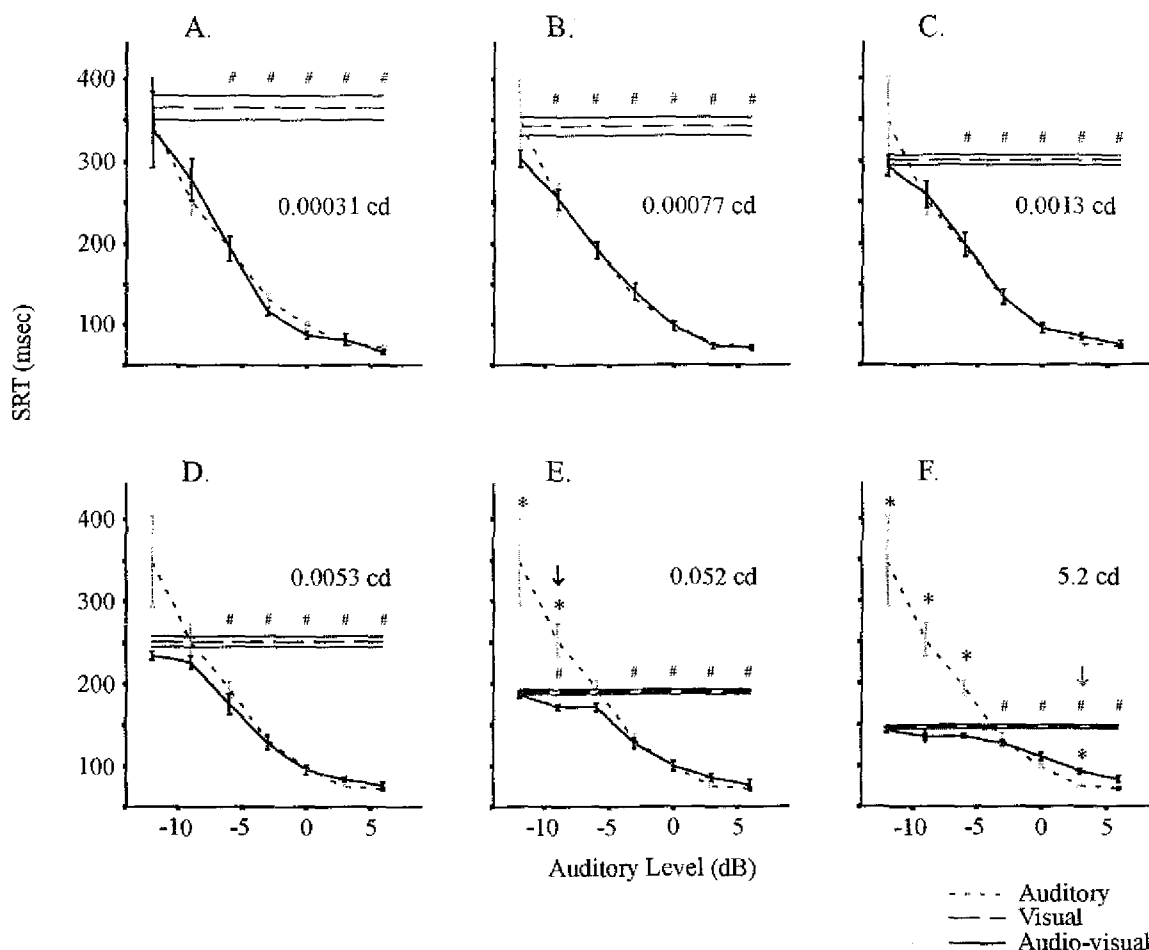


Figure 2-7. Average audiovisual SRTs compared to unisensory values for bird 'N'. One mean visual (long dash) and all seven mean auditory (short dash) SRTs (\pm SEM) are shown for each plot as a reference for the audiovisual conditions (solid). Asterisks (*) mark significant differences between auditory and audiovisual mean SRTs. Pound signs (#) mark significant differences between visual and audiovisual mean SRTs, (student's *t* test, Bonferroni corrections, $p < 0.05$). Significant differences between unisensory and audiovisual marginal means (ANCOVA, target eccentricity as a covariate) were identical to those reported for simple means (pair wise comparison, Bonferroni corrections, $p < 0.05$). Arrows mark audiovisual conditions in which the mean was significantly shorter (black) or longer (grey) than the shortest unisensory SRT.

When the visual stimulus was dim, as in Figure 2-7A, most auditory SRTs were shorter, and the audiovisual SRTs closely approximated the auditory SRTs. With brighter visual stimuli, the audiovisual SRTs followed the visual SRT values at low SPL combinations and the auditory values at higher SPL combinations (Figure 2-7D – F). There was one condition for

each bird in which the mean audiovisual SRT was significantly shorter than the shortest unisensory SRT (black arrow, Figure 2-7E). All of these results are *qualitatively* consistent with the race model. On the other hand, one stimulus combination in each bird (example bird 'N' grey arrow, Figure 2-7F, bird 'J' same stimulus combination, data not shown) resulted in a mean audiovisual SRT that was significantly longer than the shortest unisensory SRT. The following distribution analysis addresses whether these differences in mean SRT significantly violate race-model predictions.

Figure 2-8 shows cumulative probability functions for audiovisual trials at a single low level SPL (-9 dB) and six luminances (A – F) for bird 'J'. Under these conditions, the distribution of auditory and visual SRTs often overlapped. For conditions in which the visual stimulus was fairly dim (Figure 2-8A – D) the effectiveness of either unisensory stimulus alone was less than optimal in evoking a quick, reliable response. Thus, these combinations *theoretically provided the best opportunity for observing positive race-model violations*. The insets plot the difference between the upper boundary of race model predictions (Miller boundary: see Methods) and observed cumulative probability functions. Upward excursions in the inset graphs indicate positive distribution deviations from the upper boundary of the race model. The statistical significances of the excursions were computed with the KS-test (see Methods). As exemplified in this set of conditions, all observed SRTs were consistent with the predictions of the race model, and no significant positive race violations were found for bird 'J'. There was only one audiovisual combination in one bird (-9 dB and 5.2 cd, bird 'N') in which the SRT was significantly shorter than race-model predictions, and thus qualified as a positive violation. This condition is shown in Figure 2-9B (Note: the inset

difference graph for this figure is computed using the Grice boundary, but '#' marks a significant positive violation, as compared to the Miller boundary).

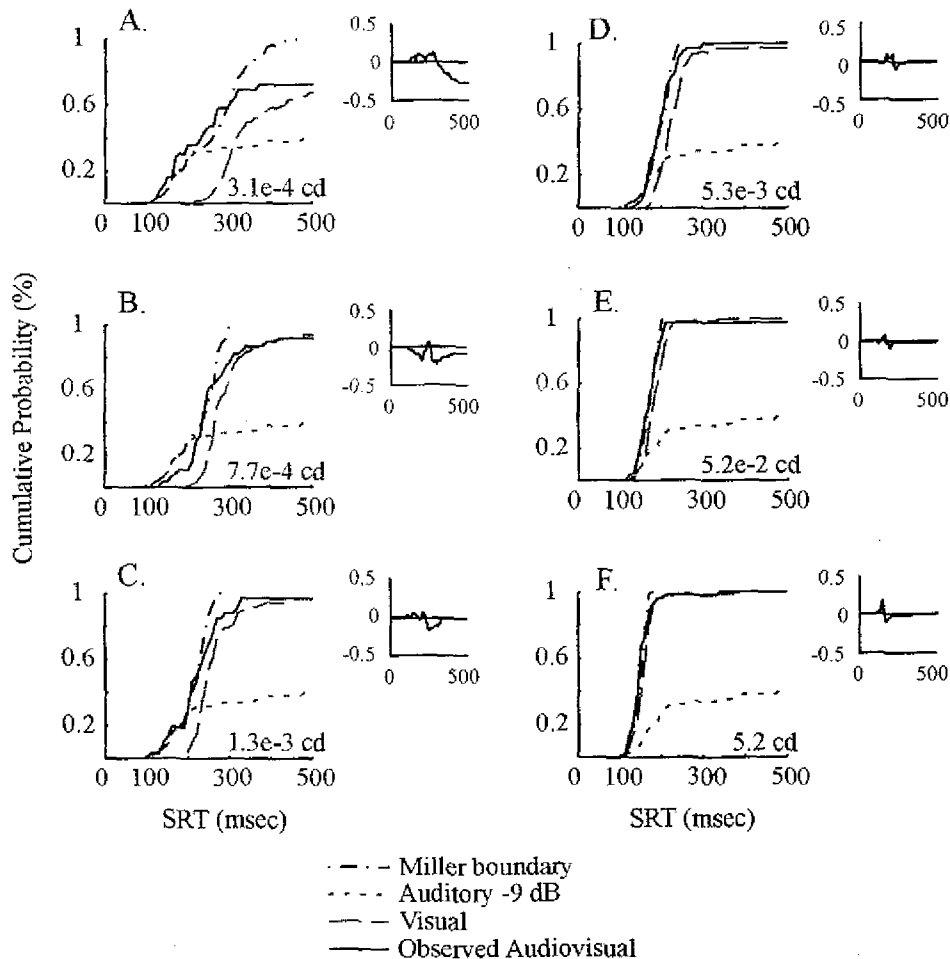


Figure 2-8. Cumulative audiovisual SRT distributions with a quiet sound.

Auditory (short dash), visual (long dash), observed audiovisual (solid) and predicted Miller boundary (dot-dash) SRT cumulative distributions are plotted for each visual stimulus (A-F), paired with the -9 dB sound. Insets show the Miller boundary minus the observed audiovisual cumulative probability for each SRT value. Even though unisensory cumulative distributions greatly overlapped, the observed audiovisual distribution was not significantly enhanced beyond the upper boundary of any race-model prediction, (KS-test, Bonferroni correction, $p > 0.05$ data shown: bird 'J'; $p > 0.05$ data not shown, bird 'N').

Negative violations, where the observed SRTs were significantly longer than race-model predictions, were somewhat more frequent and found with brighter (5.2 and 0.052 cd) and louder (-3 through 6 dB) stimulus combinations.

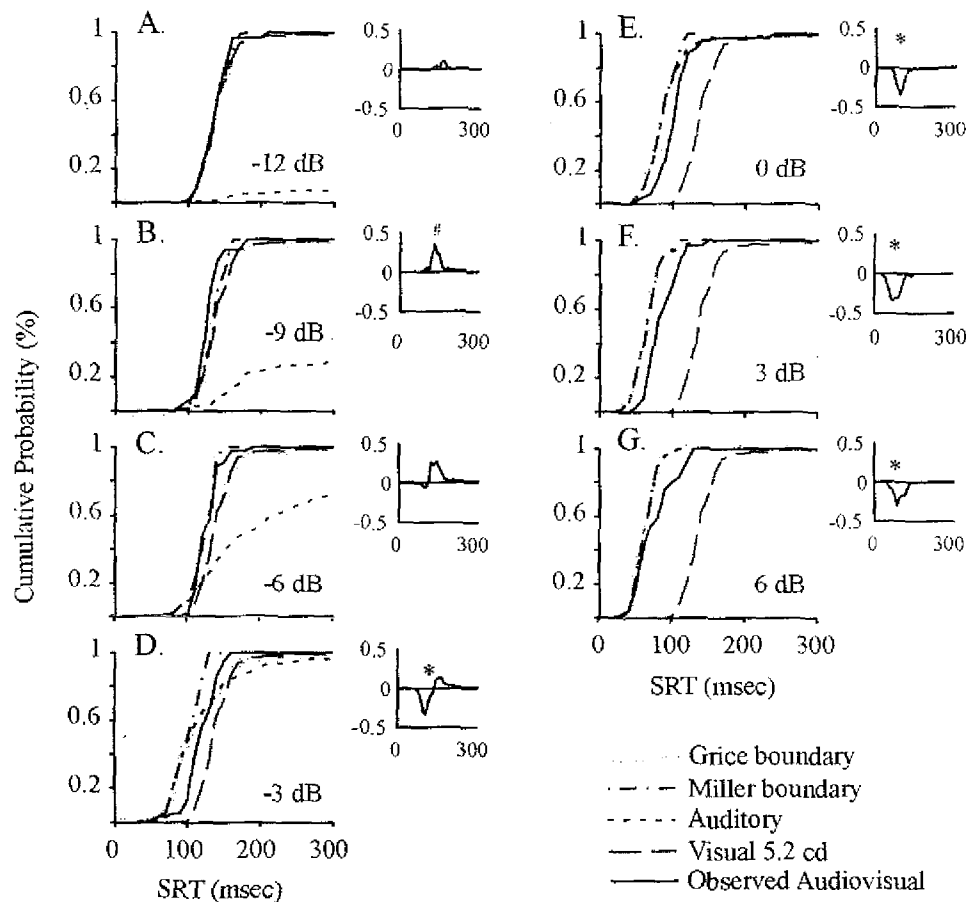


Figure 2-9. Cumulative audiovisual SRT distributions with a bright visual component. SRT distributions are plotted as cumulative probability plots, and insets show the predicted lower boundary (Grice boundary) of the race model minus the observed cumulative probability for each SRT value. Asterisks (*) mark significant negative violations (KS-test, Bonferroni corrected $p < 0.05$, data shown bird 'N'). One condition (B) yielded a significant positive violation (#), as compared to the upper boundary of the race model (Miller boundary).

The condition marked by the grey arrow in Figure 2-7F is an example of a negative violation severe enough to significantly affect mean values. There were multiple conditions, however, that showed significant negative violations in the distribution analysis. Figure 2-9 plots SRT cumulative probability distributions for a single high-intensity visual stimulus (5.2 cd) and seven SPLs (A – G). As shown in Figures 2-9D – G, four of the loud/bright combinations resulted in distributions that fell to the right of the Grice boundary (Data shown for bird 'N', bird 'J' gave similar results.)

Saccade Error in Audiovisual Trials

The race model is based on the competition between two separate streams of sensory information. Does the modality that “wins” the race determine all characteristics of the ensuing saccade, or does the presence of the stream that “lost” have an influence? Specifically, do audiovisual saccades with auditory-like SRTs also have auditory-like errors? In this section, I compare audiovisual errors with unisensory error in the context of race-model predictions. Error and SRT are then considered simultaneously for one exemplary audiovisual combination.

Figure 2-10 plots marginal mean error against SPL for all 42 stimulus-combinations. The conventions are the same as those of Figure 2-7. The insets show the slope of the regression between total error and target eccentricity (see Figure 2-5). The stimulus combinations shown in Figure 2-10A – C typically yielded audiovisual SRTs similar to auditory SRTs (see Figure 2-7A – C). Yet, even in these cases where the auditory stream presumably “won” the race to trigger a saccade, the audiovisual errors approached visual values. This would not have been the case were the winning modality to control all aspects of the saccade.

When brighter stimuli (Figure 2-10E – F) were paired with sounds that were less than -5 dB SPL, visual SRTs were shorter than the auditory SRTs (see Figure 2-7E – F). For these combinations, the audiovisual errors were, again, closer to the visual errors:

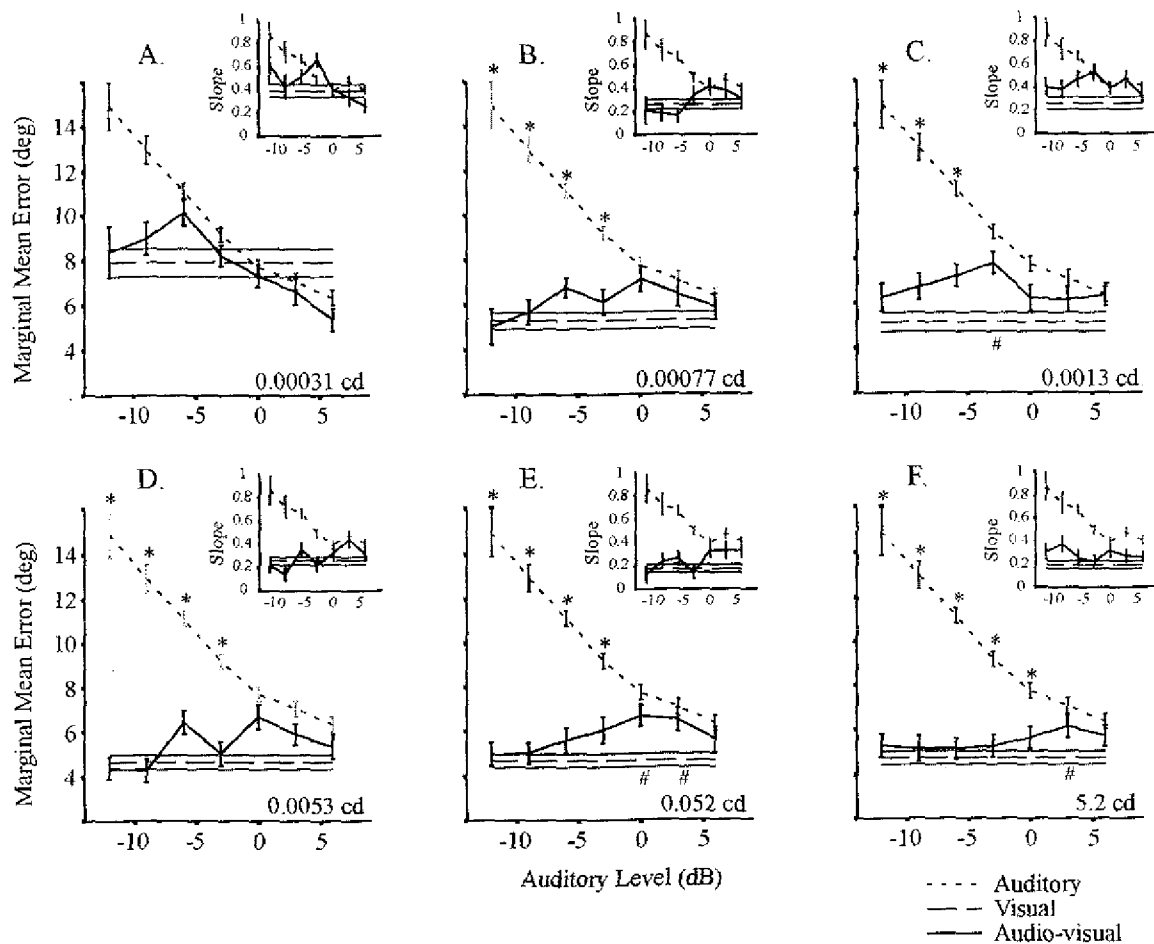


Figure 2-10. Comparison of errors from auditory, visual, and audiovisual trials. Audiovisual marginal mean errors (\pm SEM) follow the more accurate unisensory error, but are never more accurate. All conventions follow those of Figure 7, which was an equivalent plot of SRT. Asterisks (*) and pound signs (#) denote conditions where the audiovisual error is significantly different from the auditory or visual error, respectively (t -test, Bonferroni corrected, $p < 0.05$). Data from bird 'N'.

In bird 'N', whose results are shown, the audiovisual error was slightly but significantly larger than the visual error in 4 of the 42 combinations tested ("#" in Figure 2-10C, E, F). In bird 'J', the audiovisual errors were indistinguishable from visual errors in all conditions. As seen with SRT, audiovisual accuracy was never better than that seen in the most accurate unisensory condition (t -test, Bonferroni corrected, $p > 0.05$, birds 'N' and 'J').

The results above are inconsistent with the idea that the winning modality in the race model controls *all* aspects of the saccade. To test this more directly, I compiled a theoretical distribution of audiovisual saccades for each stimulus combination with a Monte Carlo simulation based on the race model. This simulation utilized the basic race-model assumptions: 1) the two modalities raced to the saccade generator, 2) the winner of the race controlled both the SRT *and* the error of the theoretical audiovisual saccade. For simplicity, I assumed independence of SRTs between the two unisensory streams.

For each multisensory combination, one randomly selected auditory saccade SRT was compared to one randomly selected visual saccade SRT. The characteristics of the saccade with the shortest SRT (including error and SRT) were then pooled in a theoretical audiovisual distribution. If the two unisensory SRTs were equal, a modality was randomly selected and the corresponding saccade SRT and error was added to the predicted audiovisual distribution. Both auditory and visual saccades were then replaced in their original unisensory pools and the selection process was repeated 100 times. The SRT and error of the resulting pool of 100 “race-winning” auditory and visual saccades were analyzed as a theoretical audiovisual distribution.

As expected, the Monte Carlo simulation resulted in audiovisual SRTs and marginal mean errors that reflected the winning modality. For many audiovisual combinations, the errors predicted by the simulation were significantly greater than those observed in the actual audiovisual trials (7/42 bird ‘N’, 9/42 bird ‘J’, *t*-test, Bonferroni corrected, $p < 0.05$).

An example of an audiovisual condition in which the observed error was less than that predicted by the Monte Carlo simulation is illustrated in Figure 2-11 (-3 dB; $5.3e-3$ cd, birds 'N' and 'J' combined). These stimuli were of moderate effectiveness; Unisensory trials with these stimulus-intensities individually evoked a response more than 90% of the time, however both auditory and visual mean SRTs were significantly longer than those seen at the highest intensities (see Figure 2-4). Figures 2-11A and B show the distance traveled in azimuth and elevation as a function of target eccentricity in azimuth and elevation, respectively. In azimuth (A), the unisensory and audiovisual trials had similar errors (deviation from the unity line), although the auditory trials had slightly more scatter at the larger target eccentricities than did the visual trials. In elevation (B), the auditory trials had appreciably more scatter than the visual trials, and notably, the errors in the audiovisual trials were more like those in visual trials.

Total audiovisual errors also closely followed the visual error trends (Figure 2-11C), with the slope and marginal mean error of audiovisual trials being significantly different from auditory but not visual values (pair wise comparison, Bonferroni corrected, $p < 0.05$, birds 'N' and 'J'). However, KS-tests of the SRT cumulative probability distributions, shown in Figure 2-11D, indicate that the audiovisual SRT was significantly different from the visual distribution, but not the auditory (Bonferroni corrected, $p < 0.05$, birds 'N' and 'J'). Figure 2-11E is a bivariate summary of these findings, plotting each head turn's SRT and total error.

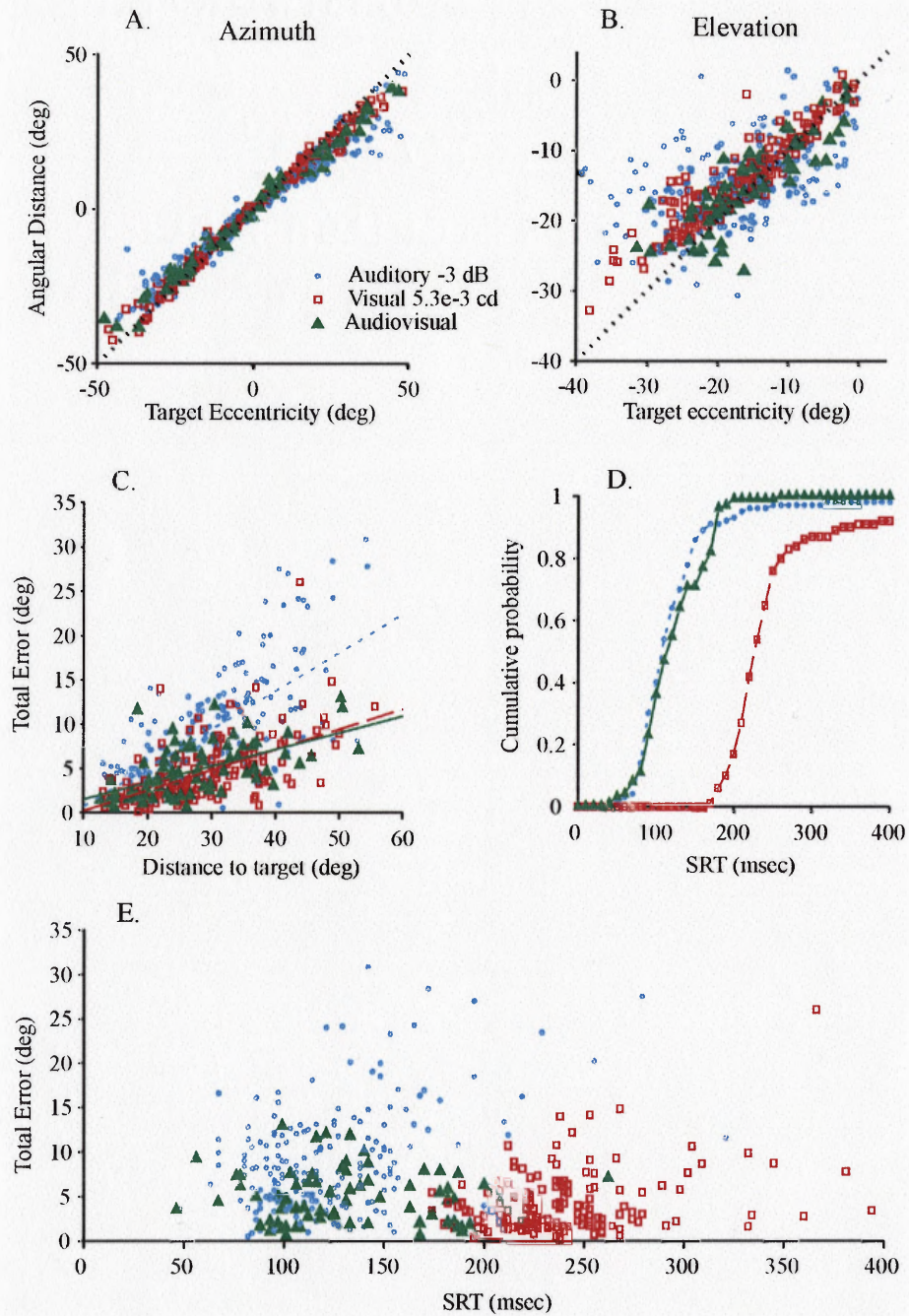


Figure 2-11. Multivariate analysis of saccadic behavior for one stimulus pair.

A&B) Azimuthal and elevational components of error are plotted as in Figure 6. Audiovisual error followed visual trends (audiovisual slope and marginal means not significantly different from visual, $p > 0.05$). C) Total error is plotted as in Figure 5. Again, audiovisual error followed visual trends (audiovisual slope and marginal means not significantly different from visual, $p > 0.05$). D) Cumulative distribution comparisons of SRT show that audiovisual SRT follows auditory trends. E) Bivariate plot shows audiovisual response has auditory speed with visual accuracy. Data pooled from both birds 'N' and 'J'.

The plot suggests that the audiovisual distribution has a shorter SRT than the visual distribution and a smaller error than the auditory distribution. The bivariate distribution of the audiovisual trials was significantly different from both the auditory and visual distributions (2-dimensional KS-test, $p < 0.025$, birds 'N' and 'J').

Figure 2-12 is a summary figure showing mean SRT and error across all unisensory and multisensory conditions for each bird ('N': panels A-F; 'J': panels G-L). In each panel, the data are normalized to the mean visual values (V-norm, grey circle, dashed lines), at the specified luminous intensity. This normalization was done to simplify comparisons across modalities and stimulus-strengths. The same unisensory auditory data (grey diamonds) are reproduced in every panel for comparison to audiovisual data (black squares), and the SPLs are labeled in panels F and L.

Examination of the unisensory auditory data in Figure 2-12 shows that as SPL decreased (grey diamonds, left to right), both errors and SRTs increased. When visual stimuli were added to the sounds at the same SPLs (black squares), however, the errors remained similar to those obtained in visual trials. The vertical disparity between the grey diamonds and corresponding black squares represents the improvement in accuracy afforded by the addition of a visual stimulus. This improvement was more apparent in trials with low-level auditory stimuli than in those with high-level auditory stimuli. For example, in each panel, the greatest vertical separation between black squares and grey diamonds occurs at the right of the plot. This is consistent with the principle of inverse effectiveness: The benefit of

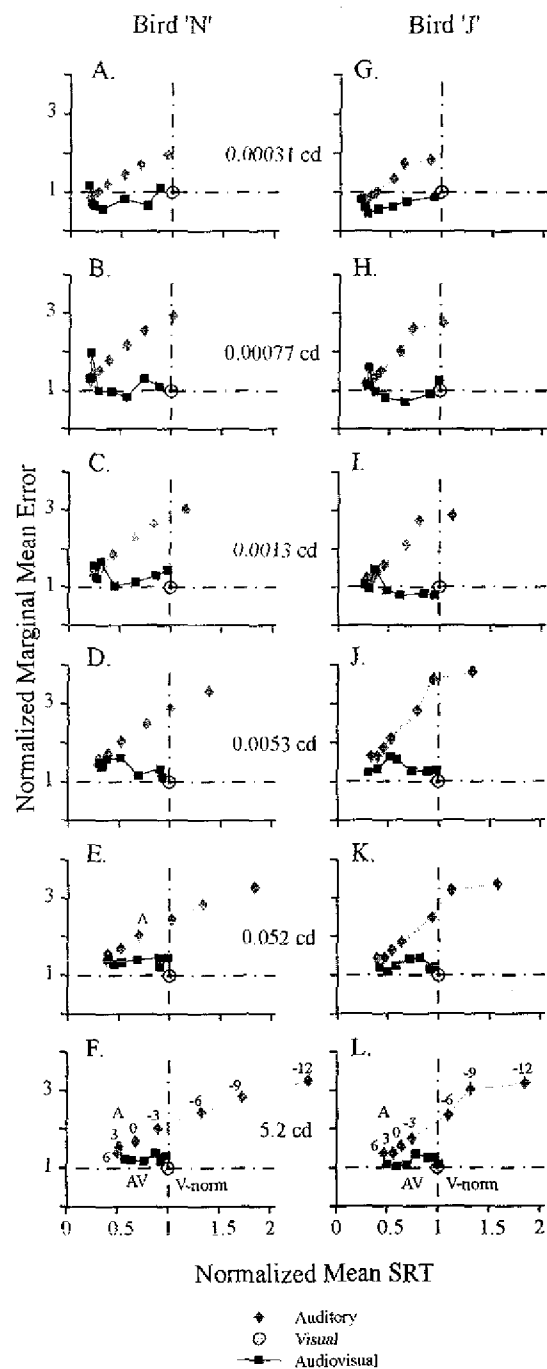


Figure 2-12. Summary of mean SRT and error for birds 'N' (left) and 'J' (right). SRT and error values obtained in auditory (grey diamond) and audiovisual (black square) trials are shown normalized to each visual condition (V-norm; grey circle; dashed lines). The same unisensory auditory data are reproduced in each panel for comparison to audiovisual data, and the SPLs are shown above the grey diamonds in panels F and L. In general, the black squares are below the grey diamonds and to the left of V-norm. Thus, multisensory responses were typically initiated faster than visual responses and more accurate than auditory responses across stimulus pairs.

adding a second modality (in this case, vision) is most apparent when the auditory component is near threshold.

SRTs also benefited from the inclusion of a second modality in a way that is consistent with the principle of inverse effectiveness. In each panel of Figure 2-12, the horizontal disparity between the black squares and V-norm represents the improvement in SRT afforded by the addition of an auditory stimulus. Across the auditory conditions, this improvement in SRT is more evident in low-intensity visual combinations (A-C, G-I) than high-intensity ones (F, L). For instance, panels B and H represent conditions with lower-intensity stimuli, and each show four audiovisual combinations with SRTs less than half the value of the visual SRT (0.5 on x-axis). Audiovisual combinations incorporating the highest-intensity stimuli, as shown in panels F and L, have SRTs that cluster to the right of 0.5 on the x-axis. Therefore, the shortening of SRT afforded by adding an auditory stimulus to a visual stimulus is most apparent when the visual stimulus is weaker.

Finally, figure 2-12 demonstrates that when a very strong stimulus is paired with a very weak one, the audiovisual SRTs and errors *both* take on the values of the stronger modality. This can be seen for high luminance, low SPL combinations in panels F and L, by comparing the right-most black squares to their unisensory counterparts. In these cases, the audiovisual responses resembled the visual responses in both SRT and error. Conversely, when the unisensory strengths in the audiovisual combinations were reversed, as in the left-most conditions in panels A and G, audiovisual responses resembled the auditory in both SRT and

error. When *both* modalities in an audiovisual combination were strong, as in the left-most squares in panels F and L, the audiovisual error can not be classified as being more like auditory or visual because the owls localized higher SPL sounds with near-visual accuracy. However, the audiovisual SRTs did seem to be more like the shorter auditory SRTs. Therefore, any combination incorporating strong unisensory stimuli produce very little evidence of audiovisual integration.

Across stimulus conditions, the data exemplify the general definition of the inverse effectiveness principle; a best-of-both-worlds audiovisual integration is *most* apparent when stimuli are near behavioral threshold. When such stimuli are used, it becomes clear that the owl's saccade system uses information from both streams to optimize orienting behavior.

Discussion

Events in nature often have auditory as well as visual components, and given that their signal strengths may vary widely, I examined how a behavioral response might benefit from various combinations of SPLs and luminous intensities. Initial characterization of the barn owl's saccades to unisensory targets showed that when unisensory signal strength increased, response percentage increased, while SRT and error decreased. Neither the SRT nor the accuracy of an audiovisual saccade was significantly facilitated beyond the level predicted by the race model. However, multisensory saccades had reaction times typical of the auditory saccades while maintaining the high accuracy characteristic of visual saccades. I also found that for stimulus combinations employing bright stimuli, SRTs were longer than the race model's predictions. These observations argue against the race model as being the sole

mechanism of audiovisual integration and suggest a convergence of the auditory and visual modalities within the saccadic sensory/motor pathway.

Probability of Response

Audiovisual response probability was indistinguishable from the predicted probability of responding to vision, audition, or both. To my knowledge, most published psychophysical studies of multimodal integration tend to employ stimuli at amplitudes well above behavioral threshold, thus ensuring detection of nearly all of the stimuli, regardless of modality. By contrast, the behavioral studies of Stein and colleagues did test peri-threshold stimuli, and is therefore relevant. They showed that cats, trained to approach a visual stimulus and press a bar directly beneath the source, were more likely to orient correctly towards the light when a sound accompanied it (Stein et al., 1988; Stein et al., 1989). The percentage of correct trials was, furthermore, enhanced beyond statistical predictions (Stein et al., 1988; Stein et al., 1989; Wilkinson et al., 1996; Burnett et al., 2004). Although the current study included stimuli of comparable magnitudes, it is not directly comparable with previous studies of the cat because Stein and colleagues required not only a reaction to the stimuli, but also that the response be to the correct location. By contrast, I considered all saccade-like movements (within criteria; see Methods), regardless of accuracy and precision. The question remains, whether the feline detection of stimuli *regardless* of correctness would be enhanced to levels beyond that predicted by a reaction to either modality alone or to both.

Saccade Reaction Time

Because saccades to audiovisual targets usually had shorter, auditory-like SRTs, the present results are consistent with a number of studies suggesting that a visual SRT will be facilitated by the addition of an auditory component (Engelken and Stevens, 1989; Hughes et al., 1994; Nozawa et al., 1994; Frens et al., 1995; Corneil and Munoz, 1996; Harrington and Peck, 1998; Hughes et al., 1998; Colonius and Arndt, 2001; Corneil et al., 2002; Arndt and Colonius, 2003; Kirchner and Colonius, 2004). Performance meeting this broad definition of enhancement is also predicted by the race model, and the present results did not show substantial evidence of SRT facilitation beyond race-model predictions, in contrast to previous results (Hughes et al., 1994; Nozawa et al., 1994; Harrington and Peck, 1998; Hughes et al., 1998; Arndt and Colonius, 2003).

Instead, the lack of SRT facilitation beyond race model predictions in the owl is consistent with a recent study on multisensory integration in human saccadic behavior (Corneil et al., 2002), wherein SRTs to simultaneous, spatially-aligned audiovisual stimuli were not enhanced beyond race-model predictions. There are three main similarities between the study by Corneil and colleagues and the present study that may help explain the lack of SRT facilitation beyond race predictions. First, both studies implemented a “divided attention” task, in which subjects are instructed or trained to localize targets of one modality while ignoring the other. In focused attention tasks, by contrast, subjects are instructed to ignore one modality and must therefore identify the modality of the target before initiating a saccade. Because my divided attention task did not require the modality identification step, unisensory responses and thus predictions made by the race model were quicker, making the

observation of positive race violations less likely (Corneil and Munoz, 1996). Second, the difficulty of the task was set at a high level in both sets of studies by presenting stimuli with a low signal to noise ratio. Third, for any given trial, the stimulus-strength, location, and modality were randomly selected, and thus unpredictable. Previous studies have linked response facilitation to *a priori* knowledge of stimulus characteristics (Mordkoff and Yantis, 1991; Schwarz, 1996; Kirchner and Colonius, 2004).

SRTs are also known to be affected by the temporal alignment of the auditory and visual components in a multisensory target (Corneil et al., 2002; Diederich and Colonius, 2004). Corneil and colleagues (2002) demonstrated positive violations of the race model when a light led the sound by 100 ms. Because SRTs in unisensory auditory trials of that study were on average 100 ms earlier than those in visual trials, presenting the light 100 ms earlier than the sound would presumably align the visual and auditory *information streams* within the saccade-control circuitry. Electrophysiological evidence from the mammalian superior colliculus suggests that this is the condition in which positive violations of the race model are most likely to occur (Meredith et al., 1987). Although the visual and auditory streams were similarly aligned in my study by ranging SPL and luminous intensity (Figures 2-4, 2-7), I did not see significant positive race model violations. It is worth noting, however, that aligning the unisensory components in this way during multisensory trials changes not only the processing time, but the strength of the inputs. Future studies in which the temporal alignment of auditory and visual information is achieved by asynchronous presentation of the stimuli, independently of stimulus strength, may yet reveal positive race model violations in the owl's saccade control system.

Negative Race Model Violations in Saccade Reaction Time

SRTs that were longer than predicted by the race model were more frequent in my study than in earlier studies of humans. These negative violations were consistently observed when the strongest visual stimulus was utilized in the audiovisual combination (Figure 2-9). Interestingly, when these bright visual stimuli were presented alone, the SRTs were still longer than those from auditory-alone trials, and therefore, the visual information stream would not have been expected to win a race to saccade initiation in an audiovisual trial. Nevertheless, a significant proportion of SRTs in both birds had vision-like values that skewed the overall multisensory distributions to the right of the Grice boundary. Data suggesting negative violations have also been reported in focused attention tasks where human subjects are instructed to generate ocular saccades to stimuli of one modality while ignoring another (Hughes et al., 1994). Negative violations can not be explained solely by a race between separate sensory streams (Grice et al., 1984; Hughes et al., 1994).

Negative violations suggest that under certain circumstances, the process of multimodal integration incorporate factors other than timing, such as the strength or precision with which the stimulus is represented within the nervous system. When the two information streams arrive within a time window, the “stronger” or more “focal” representation may control saccade initiation, even though it was not the first to arrive at the saccadic trigger site. It is premature, however, to propose the neural mechanism by which the quality of the neural representation affects SRTs. The role of stimulus salience would be better tested by manipulating stimulus onset asynchrony independently of amplitude.

Negative violations may also be due to a perceived spatial misalignment of the auditory and visual sources. Previous studies have shown that SRTs to targets in the presence of a spatially-misaligned second stimulus are significantly longer than in trials with only one stimulus (Corneil et al., 2002; McSorley et al., 2005; Spitzer and Takahashi, 2006). In my study, the LEDs and speakers were co-localized. However, in trials where a quiet sound was presented alone, I saw that the owls' saccades typically fell short of the target (Figures 2-5, 2-6). No such hypometria was observed for visual-alone trials. This difference in saccade metrics in auditory and visual trials may suggest that the neural representation of a quieter sound is shifted toward the midline relative to the visual representation. In other words, although the LEDs and speakers were aligned in space, the visual and auditory neural images may have been misaligned, effectively forcing the owl into a spatially-misaligned condition, which is known to lengthen SRTs.

The lengthening of SRT due to misalignment may stem from a network of local inhibitory neurons in the superior colliculus (Munoz and Istvan, 1998; Olivier et al., 1999; Honda, 2005). In general, these studies have shown that the presence of a misaligned distracter decreases buildup of activity in bursting superior colliculus neurons prior to saccade initiation (Olivier et al., 1999), and also lengthens SRTs beyond what is seen for either modality alone (Munoz and Istvan, 1998). The negative violations seen here, however, are due to a skew of the distribution toward values more common in visual trials, rather than a shift of the mean toward these longer values. Whether or not lateral inhibition may still play

a role in the negative violations observed with this paradigm is empirical, and should be addressed in future electrophysiological studies.

Finally, prior experience may also have played a part in inducing the negative violation. The present study incorporated an accuracy criterion for food rewards. Although owls are renown for their ability to localize sounds, their ability to localize a visual stimulus across intensities is better still, as measured by error. When auditory stimuli were quieter, there was a lesser probability of reward, even with my relaxed criteria in these cases (see Methods). Therefore it is possible that when a visual stimulus was added, the bird weighed the visual input more heavily because it allowed the bird to perform to criterion and receive rewards more consistently. Thus in some or all of the stimulus combinations, I can not rule out the possibility that the birds may have been incorporating an additional decision step in the planning the saccade.

Error

Audiovisual error at different stimulus-strengths mirrored the more accurate of the two unisensory trends and did not improve beyond the level of the most accurate/precise modality, usually vision (Figure 2-10). Others have attributed this lack of enhancement to a ceiling effect, arguing that vision provides such accurate localization that the addition of auditory information cannot further improve performance (Welch and Warren, 1986; Corneil et al., 2002; Hairston et al., 2003). The present study, however, extended the stimulus range to include visual trials with intensities so low, that errors were significantly larger than the asymptotic value (Figure 2-5), i.e., accuracy was below the hypothetical

ceiling, leaving room for improvement. Adding an auditory component to these dim stimuli however, did not improve accuracy beyond the level of the better modality (Figure 2-10A). An absolute accuracy ceiling, therefore, can not entirely explain the lack of accuracy/precision enhancement beyond unisensory levels in the owl.

I also compared observed audiovisual errors with the prediction that the winning modality of the race to saccade initiation determined other features of the ensuing saccade, i.e. accuracy. Interestingly, when an auditory stimulus evoking characteristically shorter SRTs was paired with a visual stimulus evoking characteristically more accurate saccades, the ensuing audiovisual saccade was generally initiated as early as the auditory saccades and terminated as accurately as the visual. This observation is inconsistent with the idea that when an object is audible and visible, the owl acts on *either* the visual or auditory component to generate a saccade, as is the basic assumption of the race model. The present results support Cornil et al. (2002) in suggesting that an audiovisual saccade is often neither purely auditory nor purely visual (whichever won the race) but an optimal combination of both. Furthermore, this effect was seen primarily for stimulus combinations in which neither unisensory component was loud or bright. This is consistent with the inverse effectiveness rule of multisensory integration. Thus, the two modalities do not function independently in the multisensory trial, but rather, are ultimately integrated to enhance the behavioral response.

Neural Mechanisms

Studies of the mammalian superior colliculus have typically shown that auditory and visual stimuli co-localized at the center of a neuron's spatial receptive field evoke responses that are

greater than the response to either modality alone (King and Palmer, 1985; Meredith and Stein, 1986a; Stein and Meredith, 1993; Frens and Van Opstal, 1998; Wallace et al., 1998; Populin and Yin, 2002; Stanford et al., 2005). Although there is some controversy whether the audiovisual response is greater than the sum of the responses to the visual and auditory components (Populin and Yin, 2002), this increased neuronal response parallels the improved performance in the detection/orientation studies cited above (Stein et al., 1988; Stein et al., 1989). Stein and colleagues have pointed out that in both neuronal and behavioral responses, multisensory advantages are greatest when the stimuli are weak, i.e., the principle of inverse effectiveness applies. Moreover, the proper alignment of the stimuli in time and space are required to achieve the improved performance in behavior and the increased response in the neurons.

The recent study of Bell and colleagues (2005), which examined spike timing in addition to spike rate in the superior colliculus of the behaving monkey, sheds light on the improvement of SRTs observed with audiovisual stimuli. They showed a significant correlation between the latency of the first spike evoked by weak stimuli and mean SRTs. Importantly, the addition of a second modality decreased both the first-spike latency and SRT. It is difficult, however, to relate these findings to the current study. The paradigm implemented by Bell and colleagues (2005) was a focused attention task in which the visual signal was always the target and the auditory signal was either a distracter or an enhancer, depending on spatial orientation. The auditory stimulus was never presented alone as a target. Therefore, the decrease in SRT could only be compared to the visual-alone trials and not to race model predictions. Likewise, the decrease in first spike latency was again relative to visual-alone

trials, even though a number of the cells pooled into the analysis were responsive to the auditory stimulus. The present behavioral results predict that the latency of the first spike evoked by AV stimuli will be similar to those evoked by the auditory component alone, if indeed, first spike latencies of tectal neurons determine SRTs.

The observation that the owl's saccades have the combination of speed and accuracy characteristic of the auditory and visual modalities respectively suggest an avenue of exploration of neural mechanism in addition to spike rate and latency. This approach would examine the hypothesis that the first modality triggers the saccade, and the other refines the target location sometime before saccade termination (Cornel et al., 2002; Van Opstal and Munoz, 2004). In the initial phase of this "updating" model of sensory convergence, separate auditory and visual sensory responses race to a common saccade generation site, such as the mammalian superior colliculus or avian optic tectum. Then, if the auditory sensory representation brings the saccade generation mechanism to threshold before the visual, the SRT assumes an auditory-like value. Once the saccade is triggered, the target location may be updated and refined by newly arriving visual information. The updating model thus stated, however, does not address negative violations, which may require assumptions about perceived location as well as the incorporation of themes from decision theory, such as prior experience and stimulus salience (Carpenter and Williams, 1995).

Human psychophysical experiments have shown that when two targets are presented in sequence, I was able to update the trajectory of a saccade even after the response has begun (Vliegen et al., 2004). Behavioral experiments on the precedence effect demonstrate that

such mid-course modification of trajectories is also possible in owls. The precedence effect comprises a cluster of phenomenon related to spatial hearing in echoic environments in which the location of a sound coming directly from the source dominates perception, and an echo, which arrives later, is poorly localized if at all, for review (Litovsky et al., 1999). Owls turn their heads toward the leading source when the delay between the leading and lagging sounds is less than about 10 ms, but begin to localize the lagging sound when the delay is more than 20 ms (Keller and Takahashi, 1996; Spitzer and Takahashi, 2006). In many of the trials at delays > 20 ms, owls first turned their heads toward one source and then abruptly changed their head trajectory to localize the later source. Evidence for updating was also observed in the course of the experiments conducted in the present study: In those trials where stimuli were inadvertently presented during a “spontaneous” head turn (66% of excluded trials, see Methods), the owls often made appropriate mid-course corrections. The owl is therefore neither blind nor deaf during a saccadic head turn, and the sensory information arriving later can affect the final gaze position.

Neurophysiological studies also suggest that an updating model of sensory convergence is plausible. In single multisensory units in the intermediate and deep layers of the barn owl's optic tectum, the auditory first-spike latency is approximately 10-20 msec, whereas the first-spike latency to a light flash is approximately 50-80 msec (Knudsen, 1982; unpublished observations). Although the visual and auditory receptive fields generally overlap in space, the visual receptive field is considerably finer than the auditory receptive field (Knudsen, 1982, unpublished observations). The timing and spatial resolution of responses in the optic tectum are therefore consistent with the idea that audition often triggers the saccade while

later arriving visual cues refine it. To further study the neuronal correlates of the updating model, it will be necessary to observe the complete sensory receptive field as it evolves over the course of an audiovisual stimulus. This may be possible in the barn owl, wherein computer-synthesized visual stimuli and virtual auditory space techniques now allow the rapid assessment of spatial receptive fields (Keller et al., 1998).

CHAPTER III

TIMING AUDITORY AND VISUAL STIMULI FOR OPTIMAL INTEGRATION

Introduction

This chapter was taken from a co-authored manuscript being prepared for publication. Data was collected in collaboration with Hannah Dold and Pamela Johnston. I contributed to 100% of the experimental design, 30% of the data collection, 60% of the data analysis, and 95% of the writing.

In the previous chapter, I showed that barn owls are able to capitalize on the speed of the auditory system and the accuracy of the visual system when localizing simultaneous, co-localized stimuli. What happens to this best-of-both combination when the auditory and visual stimuli are slightly misaligned in time?

Previous studies in humans have shown that reaction times to asynchronously presented audiovisual stimuli are significantly shortened when the auditory stimulus is presented first (Corneil et al., 2002; Diederich and Colonius, 2004). This decrease in SRT is beyond the statistical prediction derived from Raab's 1962 race model hypothesis, and although it does not rule out the race model as part of the underlying phenomenon, it suggests that the race model is not sufficient to explain behavior in asynchronous conditions. Instead, such findings support the idea of neural convergence of the two sensory streams prior to saccade

initiation (see Chapter II for a complete discussion of Raab's Race Model vs. theories of convergence). Although I did not find a significant violation of Raab's race model in Chapter II, there remained the possibility that the temporal alignment of the auditory and visual inputs was not optimal, giving rise to the need for a stimulus onset asynchrony (SOA) to align internal representations in time and produce a shorter-than-race model SRT. Indeed, when stimuli were presented simultaneously in the Corneil et. al. study (2002), they also observed SRTs consistent with race model predictions. However, when the stimuli were offset, with the visual component leading by 100 ms, audiovisual SRTs were significantly shorter than race predictions. Therefore this chapter describes a study designed to follow up findings from Chapter II in the owl, testing a large number of stimulus onset asynchronies (SOAs) and emphasizing those SOAs that compensate for unisensory processing-time differences.

In an orientation response, the goal is often not how fast to respond but rather where in space that response should be directed. For the subjects in this study, the reward criterion was solely based on accuracy, and therefore, it is likely that the importance of spatial location in the response was great. Chapter II showed that the barn owl's visual sense is significantly more spatially acute than the auditory when near-threshold stimuli are used. In fact, one might argue that the bird utilizes only the visual modality and ignores the auditory modality when determining the actual location of a near-threshold, simultaneously presented, audiovisual stimulus. Such a strategy has been suggested to underlie the ventriloquism effect (Bertelson and Radeau, 1981; Radeau and Bertelson, 1987; Bertelson and De Gelder, 2004). Ventriloquism is an example of seemingly complete visual dominance in spatial perception

when auditory and visual cues are spatially misaligned. However, many recent studies suggest visual dominance in localization is not a winner-take all effect, but rather is built on the idea that the spatial certainty of the visual modality far outweighs that of the other senses (Ernst and Banks, 2002; Battaglia et al., 2003; Alais and Burr, 2004; Bertelson and De Gelder, 2004; Heron et al., 2004). In fact, an estimate of the amount of visual dominance may be made by examining the difference in variances between the unisensory responses. This measure is thought to reflect the certainty of the internal representations and thus may be used to assign importance to each of the two inputs. A prediction of multisensory localization may thus be made by taking the weighted sum of the unisensory estimates. In essence, this is the maximum likelihood estimation of the multisensory stimulus.

Battaglia et al. (2003) recently tested the validity of these two models in human audiovisual behavior, visual capture vs. maximum likelihood estimate, and found that although subjects did not completely ignore the auditory modality, as would be the case in visual capture, the perception was weighted in favor of the visual cue more than the maximum likelihood model would allow. Other studies however, found that by varying the reliability of the visual stimulus, they could bias the response in ways that were consistent with the maximum likelihood estimate (Ernst and Banks, 2002; Alais and Burr, 2004; Heron et al., 2004). In the Ernst and Banks study, the only condition in which the visual modality captured the perception was that in which the reliability of the visual signal far outweighed the reliability of the other modality.

The goal of this study was to determine the utility of the later arriving modality in conditions of induced sensory asynchrony. I found that contrary to earlier studies in humans (Cornell et al., 2002; Diederich and Colonius, 2004), there was no significant advantage gained in reaction time by adding a second modality, regardless of the temporal alignment. Instead, results were generally consistent with the findings reported in Chapter II, showing that audiovisual saccades were initiated by the earliest arriving modality and that the second modality could influence saccade trajectory even in conditions of induced temporal misalignment (Whitchurch and Takahashi, 2006). This study highlights the ability of the owl to integrate target information from the greatly delayed modality into a smooth saccade. However in the most extreme SOA conditions, a proportion of the saccades were no longer smooth single saccades, but rather had two peaks in their velocity profiles. The proportion of trials with this second velocity peak was predictable, given the timing of the stimuli and the underlying unisensory SRT and saccade duration distributions. From this I estimate a time window of integration of multiple inputs into a smooth saccade trajectory, beyond which additional information can not be used to update an ongoing saccade and rather must be used to initiate a second step in the saccadic behavior.

Methods

All experiments were carried out under protocols approved by the University of Oregon Institutional Animal Care and Use Committee and in compliance with *The Guide for the Care and Use of Laboratory Animals*, (Institute of Laboratory Animal Resources (U.S.) and NetLibrary Inc., 1997).

Apparatus

The equipment used here is identical to that used in Chapter II. Briefly, experiments were conducted in a double-wall, sound-isolating anechoic chamber (Industrial Acoustics Company, IAC). The chamber's ambient noise between 2 and 10 kHz was below 18 dB SPL_A. Stimulus synthesis, data acquisition, and data analysis were all carried out with Matlab 6.5 (The Mathworks). Because the barn owls' eyes are virtually immobile in the head (du Lac and Knudsen, 1990), gaze was estimated by head-aim, sampled at 468 Hz from a custom-built magnetic search coil system (Rommel Labs, Ashland, MA). The head coil was mounted on a post that was cemented to the skull prior to training. To account for any day-to-day changes in the imposed magnetic field, the head coil system was calibrated over a 100° range in azimuth and elevation at the beginning of each session. Trials were conducted while the bird was tethered to a perch located near the center of the isolated, completely darkened, anechoic chamber. Food rewards were presented from a remotely controlled feeder located in front of the bird, at the level of the perch. Each session, seven targets were pseudo-randomly positioned on either a "T" shaped or "inverted-V" shaped frame, separated by at least 7°. The frame was alternated between the "T" and "inverted-V" orientations on a session by session basis. Target locations ranged from -30° to 30° in azimuth and 0° to -30° in elevation, with 0° azimuth and 0° elevation being at the intersection of the owl's midsagittal plane and eye-level.

Sensory Stimuli

Thirty distinct broad-band noises (2-12 kHz) were digitally pre-synthesized, ramped on and off by a 5 ms cosine envelope, and sampled at 30 kHz (Tucker-Davis Technologies Power

DAC PD1). The same sounds used in Chapter II were used here, and because they were reproduced identically, these noises are considered “frozen”. The analog signal was attenuated (PA4, Tucker Davis Technologies), amplified (HB6, Tucker-Davis Technologies), and randomly presented through one of seven lightweight speakers (Peerless, 5.08 cm cone tweeters). All noises used in audio or audiovisual trials were presented for 100 ms at -3 dB SPL_A. Suspended in front of each speaker, single red (635 nm) light emitting diodes (LEDs) were illuminated for 100 ms in each trial in the otherwise completely darkened, light-sealed chamber. The two luminous intensities that were probed in Chapter II were used here in visual and audiovisual trials – one relatively bright and one relatively dim (5.2 and 0.0053 cd, respectively). Multisensory trials were always spatially aligned. Stimulus onset asynchronies (SOAs) ranged from -50 to +300 ms in bright LED combinations and -150 to +200 ms in dim LED combinations. Negative values denote conditions in which the auditory stimulus was presented first. Each audiovisual combination had 11 possible SOAs. Thus there were 25 possible stimulus conditions: auditory alone, bright visual alone, dim visual alone, 11 SOAs for the dim AV condition, 11 SOAs for the bright AV condition.

Behavioral Paradigm

Two adult barn owls (birds ‘N’ and ‘J’) were hand-reared and trained to make head saccades to auditory, visual, and simultaneously presented audiovisual stimuli. Prior to this experiment, the birds had approximately 2 years of experience localizing auditory, visual, and simultaneous, spatially aligned, audiovisual stimuli for food rewards (including those experiments described in Chapter II). Each bird was subjected to 88 test sessions utilizing auditory, visual, and asynchronous, spatially aligned audiovisual stimuli. Thirty percent of

the trials within any given session were unisensory and 70% were audiovisual. Trials varied randomly in stimulus-strength, location, and modality. Audiovisual stimuli alternated randomly between the bright and dim combinations. All 25 possible conditions were randomly presented and equally sampled in each session.

As in Chapter II, animals were maintained at 85% free-feeding body weight by administering 40 individual food rewards in each session. The session ended when all rewards were dispensed. Saccades were rewarded when they terminated within 5° of the target. Analysis included both rewarded and unrewarded trials. Throughout the session, the birds were remotely monitored with an infrared camera, and trials were manually triggered when the spontaneous head turns ceased within the calibrated region of space. Head-aim was recorded for 3 seconds, a time-span comprising a 200 ms pre-stimulus window, the 100 ms stimulus, and a 2,700 ms post-stimulus window. The inter-stimulus interval varied from 15 to 90 sec.

A total of 8933 trials were recorded between the two birds. Thirty-six percent of these (3234 trials) were excluded from analysis for the following reasons: 1) The head moved more than 2° in azimuth or elevation in the 200 ms pre-stimulus window, or before 40 ms post-stimulus onset (18%). 2) The head was outside the calibrated region of the coil system at any point during the 3 second recording (6%). 3) Targets were located above the initial head-aim and/or either too close or too far away (12%); In training and preliminary testing, both birds showed a bias against responding to stimuli closer than 12° and farther than 55° away, as well as those targets presented in the upper visual hemi-field. The following

analyses were conducted on the remaining 5699 trials that were accepted, and they consisted of 98-155 trials per condition for each bird.

Data Analyses

Saccade Reaction Time (SRT)

SRT was defined relative to stimulus onset as the first point at which the instantaneous head speed continuously exceeded two standard deviations of the average speed measured in the 200 ms pre-stimulus window. As shown in Chapter II, SRT was not found to correlate significantly with the initial distance to target. Therefore SRT was pooled within a given stimulus condition, regardless of the initial distance to the target. SRT distributions were analyzed for normality with the Lilliefors hypothesis test and found to be significantly different from the null, $p > 0.05$. However, mean SRT was reported as a measure of centrality for these distributions so as to maintain a consistent presentation with Chapter II of this dissertation, as well as to compare with other studies in the field. It is known that SRT distributions are generally not normal (Carpenter and Williams, 1995), yet SRT means are consistently reported in the field of audiovisual integration in saccade generation (Colonius and Diederich, 2004). The observed audiovisual SRT *distributions* were compared to a predicted set of distributions derived from the basic principles of the race model. This model assumes that the two competing modalities race to a common saccade generation site and that the first modality to reach this site triggers the saccade (Raab, 1962).

Predicted audiovisual SRT distributions were generated with a boot-strapping algorithm that randomly selected one observed auditory and one observed visual SRT value from the

measured unisensory SRT distributions. The absolute value of the stimulus onset delay was then added to the later presented modality so as to incorporate the induced SOA in the audiovisual prediction. The minimum of the unisensory SRT values was taken into a new distribution, the predicted audiovisual distribution. This sampling process was repeated approximately 100 times, a number that corresponded with the number of observed trials for each audiovisual condition, and thus a predicted distribution of audiovisual SRTs created. One thousand of these predicted distributions were generated, thereby creating a mean prediction, with 95% confidence intervals. The observed audiovisual SRT distribution was compared to this prediction.

Localization Error

The end point of a saccade was designated when the head speed dropped below $5^\circ/\text{sec}$. Across saccades, this criterion for head speed accurately delineated the end position of the saccade. Errors in azimuth and elevation were measured, respectively, as the horizontal and vertical distance between the saccade endpoint and the target. Total error was computed by vector summation of the horizontal and vertical components.

Saccade error is known to depend on target eccentricity in the barn owl (Whitchurch and Takahashi, 2006). In the current study, total target eccentricity (vector sum of the azimuthal and elevational components) was a continuous variable, with targets ranging from approximately 5° to 60° away. In elevation however, the targets ranged from 0° to 35° away. From the analysis in Chapter II, I found that the elevational component of error was best

suited to delineate the differences between stimulus conditions. Therefore, the elevation component of error was used in the majority of analyses in this chapter.

To summarize error across target eccentricities in the elevation, individual trials were first binned with a sliding 5 degree window and the root mean squared error (RMS-error) and standard deviation within this window was computed and plotted. A 5 degree window was chosen because this was the largest window showing no significant correlation between target eccentricity and saccade amplitude. This binned RMS-error was averaged across target eccentricities to compute a measure of general accuracy for each condition (modality and SOA). Similarly, the standard deviations of the binned data were averaged across the same range as a general measure of precision.

In previous studies, the maximum likelihood estimate of multisensory integration has been predicted using Bayes rules (Ernst and Banks, 2002; Alais and Burr, 2004). However, a principal assumption that accompanies such predictions is that the unisensory estimates are non-biased in nature. I showed previously (Chapter II) that the barn owl's spatial estimates of both auditory and visual stimulus locations are indeed biased, as measured by the consistent under-shoot present especially in conditions where the stimulus certainty is low. Therefore a more general estimate of maximum likelihood in the form of Fisher information is used. Fisher information (FI) is a statistical index used to report in terms of variance how one measured variable changes with respect to another. In this case, I am measuring how saccade amplitude changes with target eccentricity. Because the birds undershoot targets that are further away, there is a systematic difference in the relationship between the two

variables. Thus the FI statistic takes into account the slope of the regression as well as the variance of the correlation in the following equation:

$$FI = \frac{(dy/dx)^2}{\sigma_y^2} \quad \text{Eqn. 3-1}$$

The numerator is the square of the local slope of the regression line between saccade amplitude and target eccentricity in elevation, and the denominator is the local variance of the saccade amplitude. The maximum likelihood estimate of multisensory integration using FI predicts that the FI of the multisensory response should be equivalent to sum of the auditory and visual FIs (Papoulis, 1991; Ma et al., in press). Local values for FI were computed in 9 degree sliding bins across target eccentricity instead of an overall FI statistic for each condition, because error was not found to be linear with elevational target eccentricity. Audiovisual FIs were then compared to the sum of the unisensory FIs.

Saccade Kinematics

Normal saccades typically have a single velocity peak (Chapter II). Figure 3-1 illustrates a smooth, single saccade taken from the 0 ms SOA combination of a bright visual stimulus and a -3 dB auditory stimulus presented to bird 'N'. Red dots in Figure 3-1A indicate head position post stimulus onset but prior to the latency point. Green dots show head position taken prior to stimulus onset and after the saccade end point, whereas blue dots show the actual saccade. The processing of the saccade trajectory began by smoothing both azimuth and elevation with a 10 ms boxcar convolution. The vector sum of the azimuthal and elevational components of head aim was computed, smoothed with a 10 ms boxcar convolution, and subtracted from the target location to give the overall position as a

function of time (Figure 3-1B, green trace). The derivative of position with respect to time was smoothed with a 20 ms boxcar convolution and plotted as instantaneous speed (Figure 3-1B, blue trace). The derivative of instantaneous speed with respect to time was smoothed with a 20 ms boxcar convolution and plotted as acceleration (Figure 3-1B, red trace).

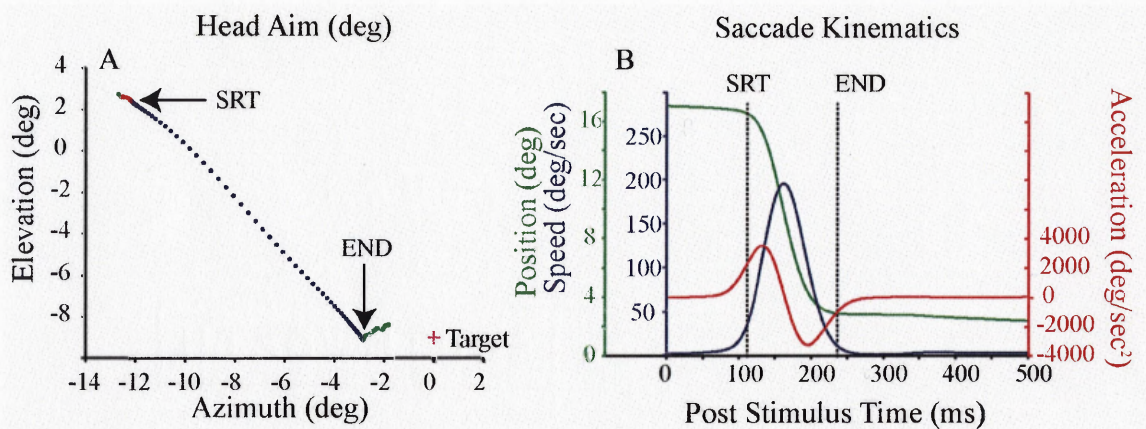


Figure 3-1. Example of a normal saccade.

Saccades were generally characterized by a single peak in the speed of the head movement (blue trace, B). The saccade in A is an example of this, showing the azimuthal and elevational location of head aim through time (468 Hz sampling rate). The saccade begins at the top left corner of the panel and proceeds toward the target (pink cross). The saccade was made by bird 'N' in a bright audiovisual trial with an SOA of 0 ms.

Because the induced temporal difference in auditory and visual presentation ranged to extreme values, I measured the probability that a saccade would be deviant from this normal speed/acceleration profile. Each saccade was tested for a second acceleration phase prior to the end point. The presence of this second acceleration marked the presence of a second peak in the velocity trace, and is thus denoted a “double-step” saccade.

Simulation of Double-Step Probability in Audiovisual Trials

The simulation to predict what percentage of trials in audiovisual conditions might be double-stepped in nature is based on the following assumptions: 1) Processing of the auditory and visual components localization progresses in a parallel fashion. 2) Auditory and visual processing times in multisensory conditions are represented by unisensory trials. 3) A later arriving stimulus is used in some fashion within the range of SOAs tests here. 4) There is a stage in saccade generation after which any additional information can not be integrated smoothly into the current saccade but must rather be used to initiate a new one.

This simulation is based on race model principles and derived from the SRT simulation, above. Initially, a distribution of multisensory SRTs was generated by a race between the two senses, taking into consideration the imposed SOA (see above). Linked to this distribution however, was a distribution of saccade durations, sampled equally from target eccentricities. The SRT of the losing modality was then compared to the total time required to produce a saccade to the first modality alone (SRT + saccade duration). If the SRT of the losing modality was longer than the time required to initiate and complete a saccade in response to the first modality presented, a double-step was predicted. As in the simulation above, this process was repeated 1000 times and averaged.

SRT was initially set as the predicted time point marking the end of processing for the second modality; however it is reasonable to assume that this is an overestimate. SRT is known to incorporate a sensory processing time, motor processing time, and the time required to engage the muscles in an overtly observable movement. Therefore it is intuitive

that predictions utilizing SRT as the end-point for processing give a much higher probability of double-steps than what was observed. I consequently repeated the simulation, subtracting discreet amounts of time (ω) from the second modality SRT to estimate its relevant processing time. The time window (ω) that yielded updating probabilities most like those observed, as measured by the lowest root-mean-squared-difference across SOAs, is reported for each bird in each stimulus combination.

Results

I analyzed 5699 head turns in response to auditory, visual, and audiovisual stimuli presented in the lower frontal hemi-field. Auditory SPL was kept at a constant -3 dB and visual luminous intensities were either relatively bright (5.2 cd) or relatively dim (0.0053 cd). These stimulus strengths are known to evoke saccades >80% of the time in these birds (see Chapter II). The auditory trials are expected to have significant errors and longer SRTs than what would be seen in louder conditions. Thus saccadic behavior has room for improvement with the addition of a visual component. Similarly, the dimmer visual stimulus is known to evoke saccades with a significantly longer SRT than the brighter visual stimulus. Audiovisual stimuli were spatially co-localized and SOAs ranged from -150 to 200 ms in the bright conditions and from -50 to 300 ms in the dim conditions. Data describing saccades from two subjects, birds 'N' and 'J', were individually analyzed and summarized in the figures below.

SRT

When auditory and visual stimuli were temporally misaligned in audiovisual trials, SRTs were generally equivalent to unisensory values for the modality presented first. Figure 3-2 shows the mean SRT for all conditions in both birds. Panels A and B summarize data for higher intensity visual combinations, whereas panels C and D show SRT data for lower intensity visual combinations for both birds.

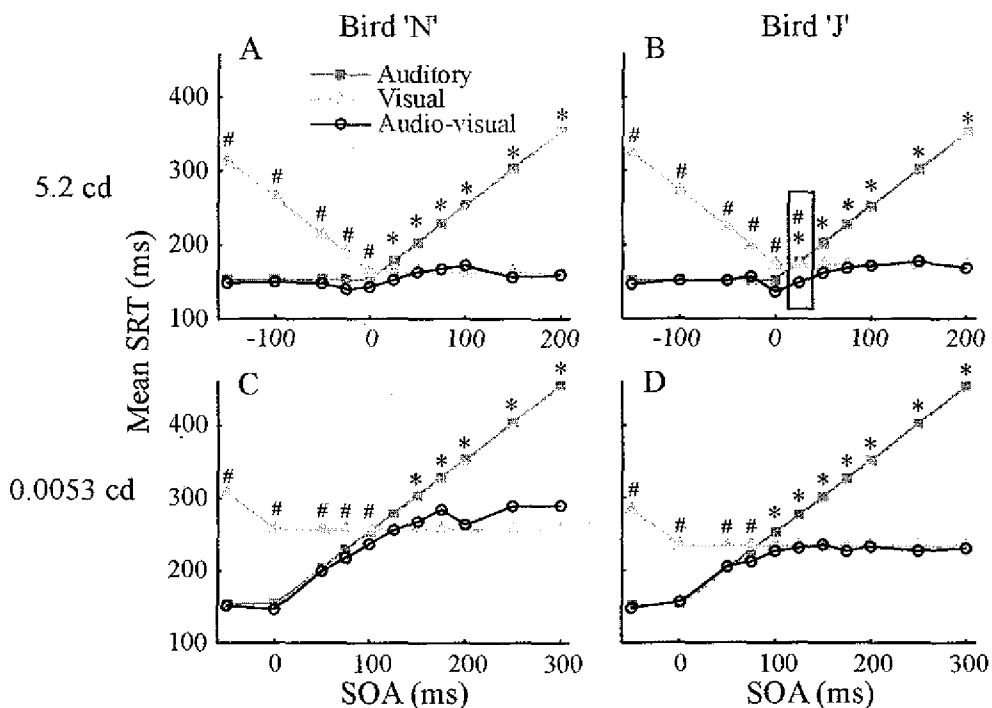


Figure 3-2. The effects of SOA on SRT.

Audiovisual SRTs followed the earliest expected unisensory SRT. Mean SRTs for each observed audiovisual condition are shown (black circles) relative to predicted auditory (grey triangle) and visual (grey squares) mean SRTs. Observed unisensory SRTs are plotted at 0 SOA. In nonzero SOA conditions, the absolute value of the SOA was added to the delayed unisensory SRTs to predict unisensory values along the abscissa for comparison purposes. Significant differences between audiovisual and predicted auditory SRTs are shown by the asterisks, whereas the pound symbols mark significant differences between audiovisual and predicted visual SRTs. One SOA condition resulted in a mean audiovisual SRT that was shorter than either predicted unisensory SRT (bird 'J', 25 ms SOA, boxed).

In each panel, the unisensory SRTs are plotted relative to the timing of their presentation for comparison purposes. For example, auditory SRT values (grey squares) are repeatedly

plotted for all negative SOAs because in these cases, the auditory modality was presented first. It follows then that the visual SRT is summed with the absolute value of SOA and plotted (grey triangles) for these negative SOAs. Because visual stimuli came first in positive SOA conditions, the absolute value of the timing delay is added to the auditory SRT and plotted. As shown in Chapter II, the visual alone SRT (grey triangle at 0 SOA) is longer for the dimmer than the brighter conditions.

Asterisks denote conditions where audiovisual SRT is significantly different than the predicted auditory alone SRT at that SOA (*, *t*-test, $\alpha=0.05$, Bonferroni correction).

Similarly, significant differences between visual and audiovisual SRTs are marked with a pound sign (#, *t*-test, $\alpha=0.05$, Bonferroni correction). Only one audiovisual combination for one bird resulted in a mean SRT that was significantly different from both unisensory predictions (boxed condition, panel B, SOA=25).

The results summarized in Figure 3-2 are generally consistent with race model expectations (see Chapter II). To test this more directly, a series of audiovisual SRT distributions were predicted using race model assumptions and a boot-strapping algorithm (see Methods).

Each theoretical SRT distribution was generated from the observed unisensory distributions, and for conditions in which SOA was not equal to zero, the absolute value of SOA was added to the SRTs of the modality presented later. Figures 3-3 and 3-4 show in grey the 95% confidence intervals of the predicted audiovisual SRT distributions for each SOA condition.

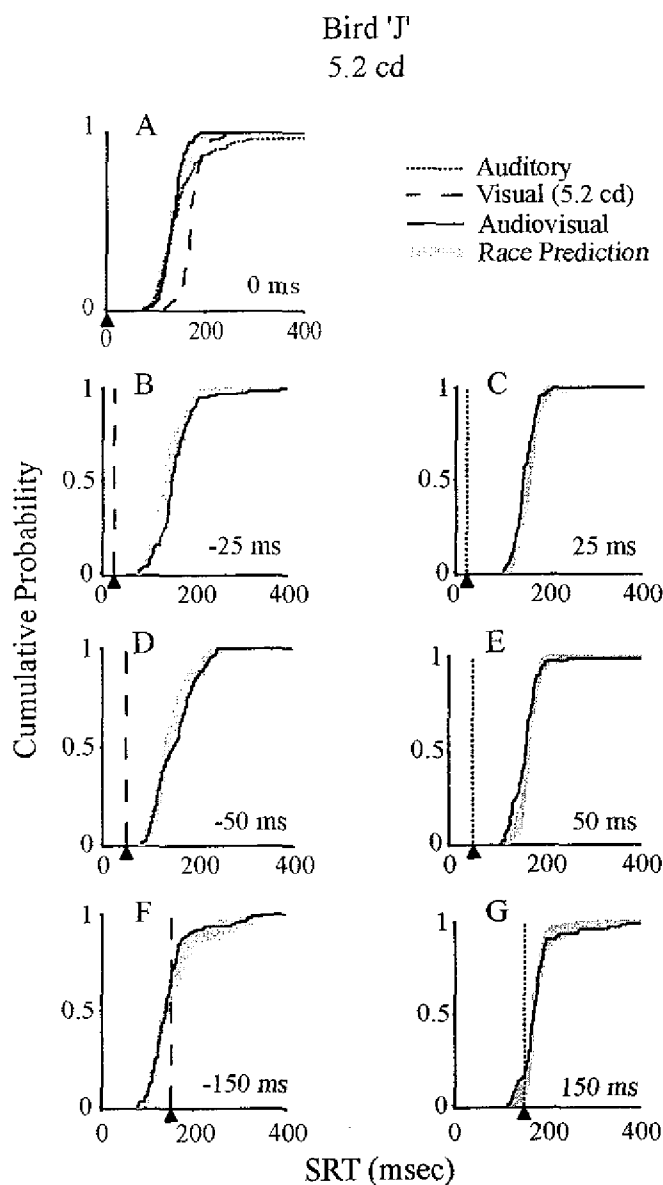


Figure 3-3. Observed SRT distributions vs. race model predictions – bright conditions. Audiovisual SRT distributions (solid black line) did not substantially deviate from race model predictions (grey, 95% confidence intervals). Panel C shows the audiovisual distribution outlined in Figure 1B, and although it is near the border of the 95% confidence intervals, it does not substantially deviate from the race model prediction. The onset of the second modality is shown by dotted or dashed vertical line (dotted – auditory, dashed – visual). Cumulative distributions for the auditory and visual alone conditions are plotted for comparison in panel A (dotted and dashed, respectively). The data shown is from bird 'J', bright combinations (5.2 cd), and is essentially identical to observations from bird 'N', data not shown. All auditory stimuli were presented at -3 dB SPL.

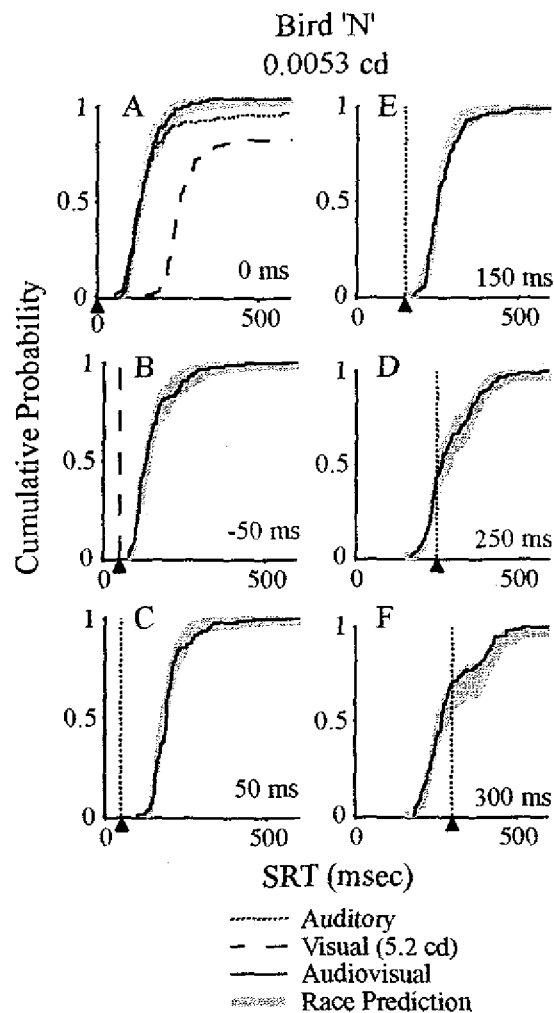


Figure 3-4. Observed SRT distributions vs. race model predictions – dim conditions. Labeling conventions are identical to those from Figure 2. Audiovisual SRT distributions did not substantially deviate from race model predictions for dim stimulus combinations for bird 'N' (data shown) or bird 'J' (data not shown).

The observed audiovisual SRT distributions are overlaid in black. All SRT distributions are shown as cumulative probability distributions, as in Chapter II. Vertical bars show the onset of the delayed modality (dashed – visual, dotted – auditory). Figure 3-3A shows observed SRT distributions for the unisensory and 0 ms SOA audiovisual conditions for comparison. In general, the distributions did not substantially deviate from race model predictions. Moreover, the distribution of the one condition in which the audiovisual mean SRT was

significantly different from both the unisensory means (bird 'J', 25 ms – SOA) was also very near race model predictions.

Similarly, Figure 3-4 illustrates that neither positive nor negative race model violations were evident in the data from dim conditions. These findings support those described in Chapter II of this thesis and do not support the theory of sensory summation and convergence prior to saccade initiation.

To illustrate the proposed relative contributions to audiovisual SRT from each modality, the probability of winning the race to saccade initiation is plotted for each SOA condition in Figure 3-5 (auditory – open bar, visual – solid bar). Note that the probability that the audition or vision would win the race is roughly equal at about 25 ms SOA in the bright combination (panels A & B), and about 100 ms SOA in the dim combination (panels C & D). Therefore, the SOAs presented in the dim conditions were offset by 100 ms so as to include an approximately equal number of conditions of each modality winning the race.

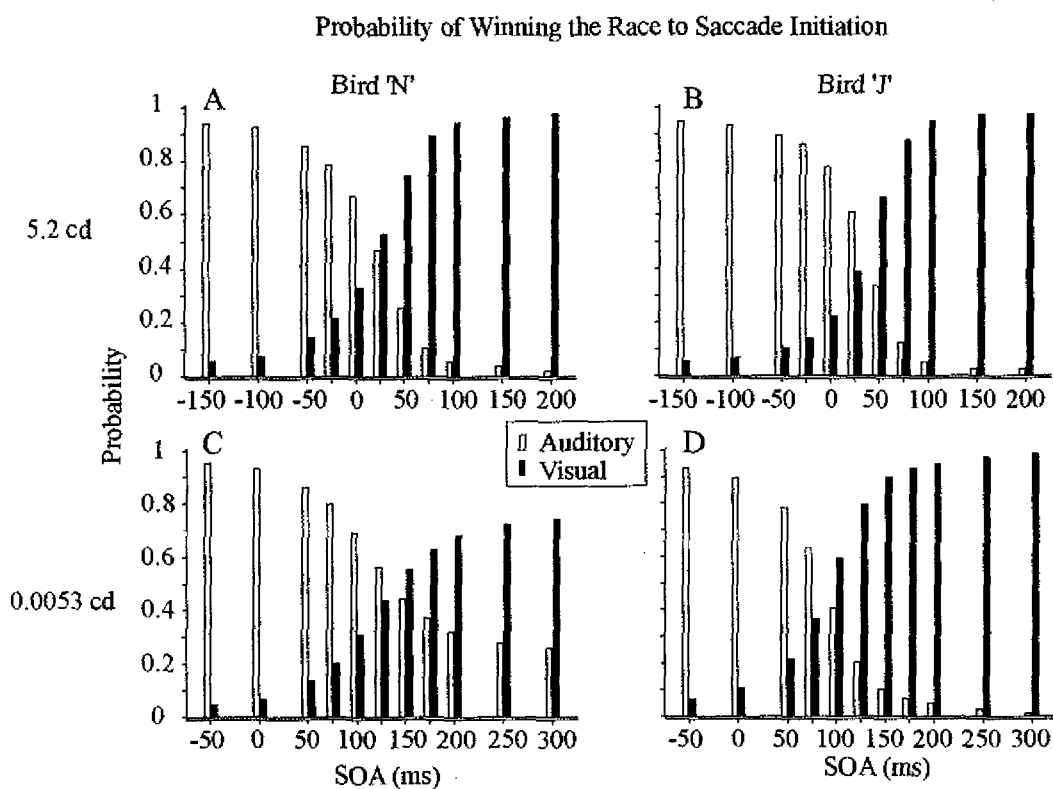


Figure 3-5. Probability of winning the race to saccade initiation – Simulation results. The race model predicts that when the auditory modality is delayed 25 ms relative to the bright visual modality (5.2 cd), there is approximately a 50% chance that either modality will win the race (A & B). A delay of 125 and 100 ms must be imposed on the auditory modality to obtain an equal probability across the two modalities in dim conditions (C & D).

Error

As shown in Chapter II, the barn owls' accuracy depended on the initial distance to the target, or target eccentricity. Figure 3-6 plots the saccade amplitude against target eccentricity. The unity line (dashed) shows the expected saccade amplitude, were the owls performing without error. Although the distance traveled in azimuth, elevation, or overall (panels B, C, and A, respectively) was correlated with the initial distance to the target (eccentricity), the distance off the unity line increased at greater target eccentricity. As shown in this figure, the difference in performance between the auditory (blue) and visual (red) modalities is most apparent for trials with greater target eccentricities. Furthermore,

the elevational component of error (Figure 3-6C) shows this difference to the greatest extent. Thus elevational component of error will be the focus of the remaining analyses.

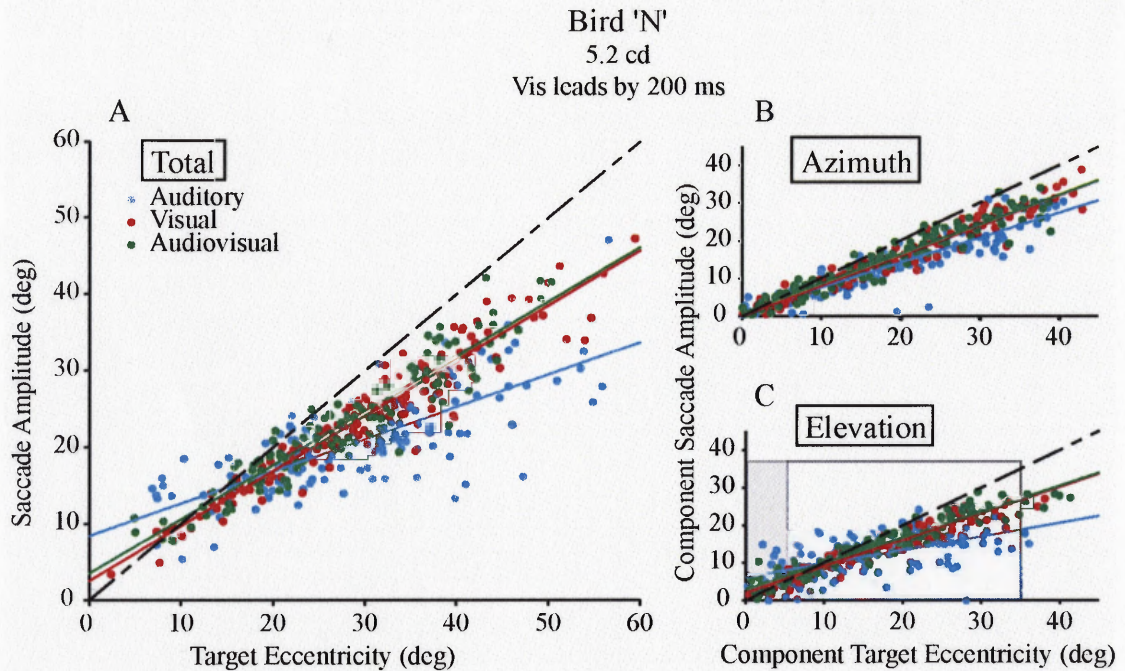


Figure 3-6. Saccade amplitude as a function of target eccentricity. Subjects were more likely to undershoot more eccentric targets (A). This trend was identifiable in both the azimuth and elevation components of auditory, visual, and audiovisual trials (B & C). The elevational component showed the most separation between behavior in response to auditory or visual stimuli (C, blue and red, respectively). The shaded region of the box in C illustrates the sliding window used in the analysis for Figure 6. Data from bird 'N', 5.2 cd, SOA – Visual leads by 200 ms.

Figure 3-6 shows that audiovisual responses (green) from the 200 ms SOA condition follow most closely the visual responses (red). Because there is no appreciable difference between the visual and audiovisual saccade amplitude vs. target eccentricity relationship, it appears that neither the SRT (Figure 3-2A) nor error is affected by the later presentation of the auditory modality in this 200 ms SOA condition in bird 'N'.

As shown in Figure 3-6C, fewer saccades were obtained for targets at greater eccentricities.

Therefore, my data analysis will focus on those saccades that were directed to targets that

were between 0° and 35° away in elevation (Figure 3-6C, box). Within this region, I used a sliding 5 degree window to bin data into 30 bins, as shown in Figure 3-7. In panel A of this figure, the root-mean-squared-distance between the target and the end position for that bin (RMS-error in elevation) is plotted against the initial distance to target (elevational eccentricity). In panel B, the standard deviation of the error is plotted. RMS-error is a measure of accuracy, whereas standard deviation is a measure of precision. Although accuracy depends on the distance to target, this dependence is not necessarily linear for all conditions, across all eccentricities. To estimate of the overall accuracy and precision across eccentricities for a given condition, the average RMS error and average STD are estimated from the binned data shown in Figure 3-7.

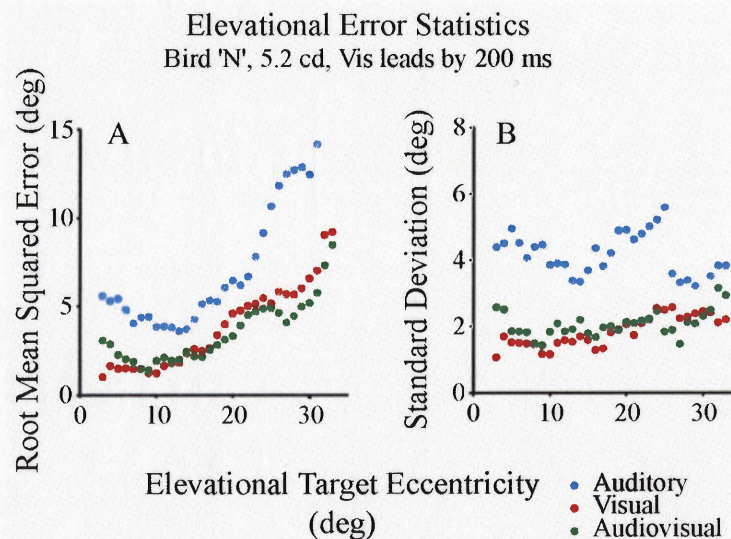


Figure 3-7. Elevation error statistics.

As a measure of accuracy, the root-mean-squared-error in elevation was computed for trials within a 5 degree sliding window along target eccentricity (3 to 35 degrees elevational eccentricity – see region in box of figure 5C). In general, RMS-error increased with target eccentricity (A), and subjects were less accurate in auditory trials than visual trials. Precision was reported as the standard deviation of error reported in Figure 5C, similarly binned (B). As expected from this SOA (visual leads by 200 ms), both the accuracy and precision of the audiovisual condition followed the visual trend for all target eccentricities.

Figure 3-8 shows the general accuracy of saccades to auditory, visual and audiovisual targets. Bird 'N' showed accuracy trends that were not quite as good as visual, but clearly enhanced relative to auditory trends (Figure 3-8A, C). Bird 'J' responded to audiovisual trials with more accuracy than even visual trials (Figure 3-8B, D).

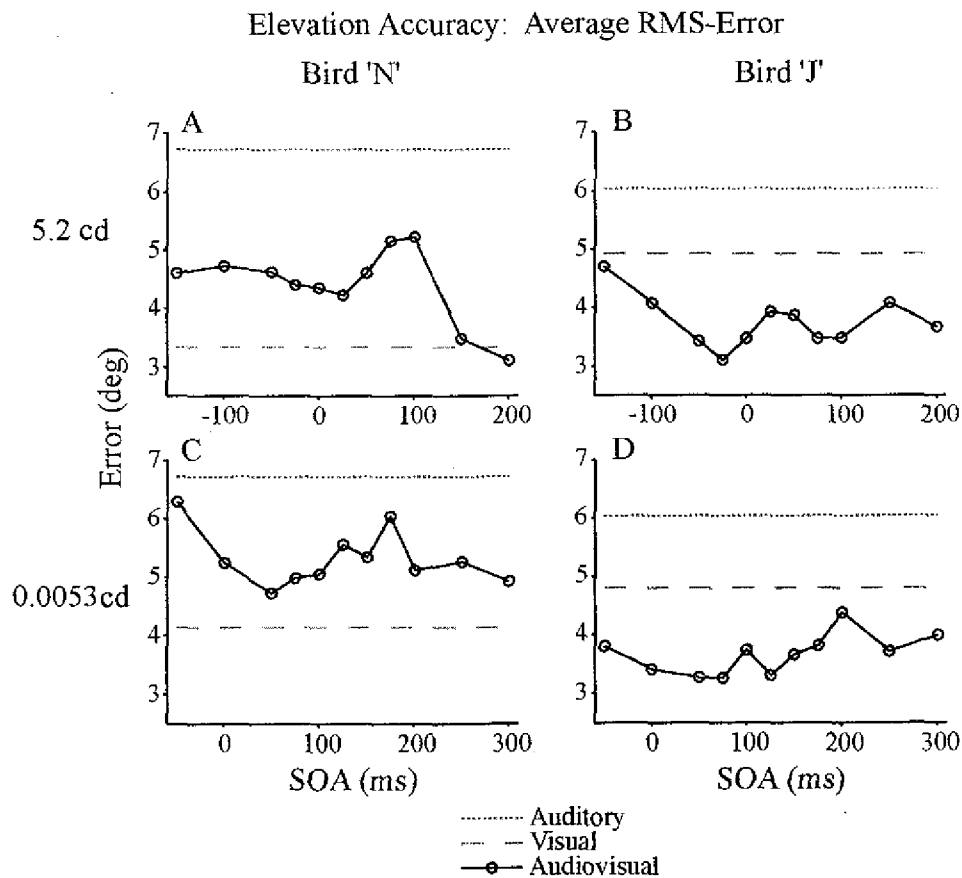


Figure 3-8. Summary of accuracy.

Accuracy in elevation for audiovisual trials (black circles) did not vary systematically with SOA. RMS-error was computed for bins across target eccentricity (see Figure 6) and averaged across bins to obtain overall estimates of accuracy for each SOA and unisensory condition (auditory – dotted, visual – dashed, audiovisual – solid). In general, bird 'N' (A & D) showed accuracies somewhere between what was seen for the either the auditory or visual conditions. Conversely, bird 'J' was more accurate in audiovisual trials than either unisensory modality.

Figure 3-9 shows that general precision for audiovisual conditions were again most similar to visual precision. The one exception, shown in Figure 3-9C, was the dim condition in bird 'N'. Here, the precision generally intermediate between the two unisensory values. Figures 3-8 and 3-9 thus illustrate that in general, neither accuracy nor precision depended on SOA.

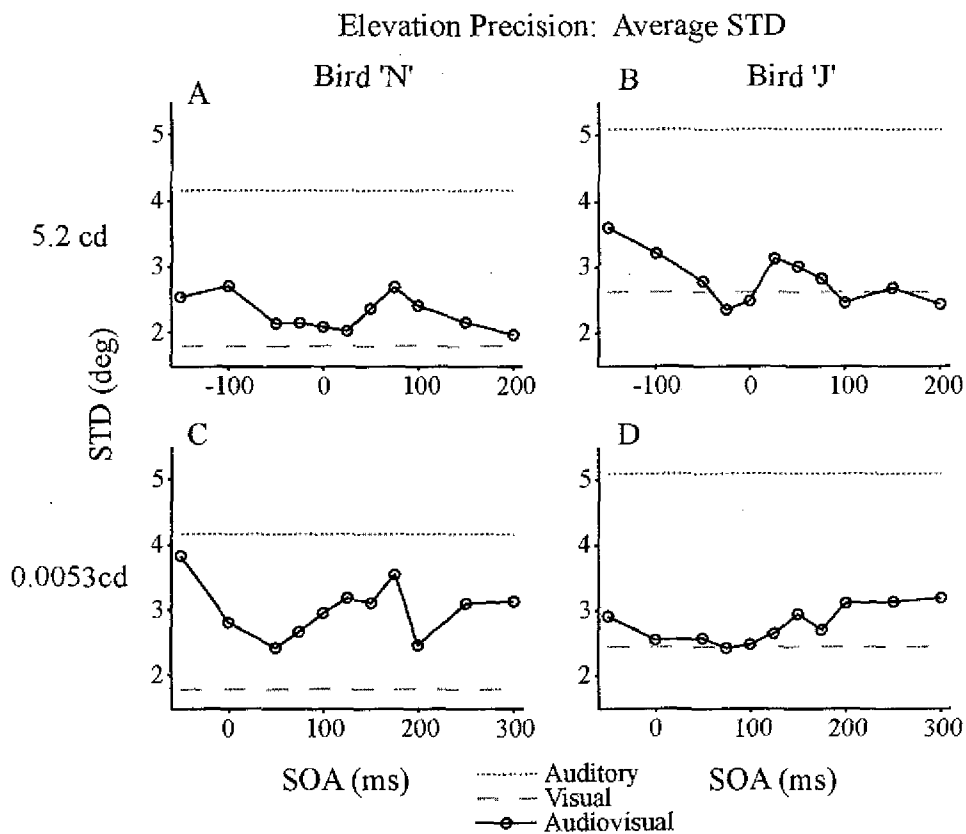


Figure 3-9. Summary of precision.

Precision in elevation for audiovisual trials also did not vary systematically with SOA (labeling conventions identical to Figure 3-7). The standard deviation of error in audiovisual trials, averaged across eccentricity bins, followed very closely that seen for the visual trials in both the bright and dim combinations for bird 'J' (B, D), and the bright condition for bird 'N' (A). Precision in the dim condition for bird 'N' (C) was not as good as the visual alone condition, but did not reach auditory levels for most SOAs.

The maximum likelihood estimate of audiovisual integration can be predicted by taking the simple sum of the unisensory FI's (see Methods above). In Figure 3-10A, the local FI's are plotted relative to target eccentricity for auditory (dotted blue) and visual (dashed red) conditions. The dotted line is the linear sum of the unisensory FI-statistics, and thus the MLE-prediction of audiovisual FI as a function of target eccentricity. Panels B-H show the observed audiovisual FI-statistics for each distance to target (solid green) as compared to the predicted FI-statistics (dotted, re-plotted in each panel). SOA for the observed audiovisual trials is noted in the center of each panel. Across target eccentricities, the -150 SOA condition (Figure 3-10B) resulted in FI-statistics most different from the predicted FI-statistics.

Figure 3-11 summarizes the difference between the predicted FI and the observed FI for each bird and for each condition by plotting the RMS-difference between the observed and predicted FI across target eccentricities. In the bright conditions (Figure 3-11A), the data from bird 'J' more closely fits the predicted trends across SOAs. However, in the dim conditions (Figure 3-11B), both birds show a larger difference between observed and predicted FI.

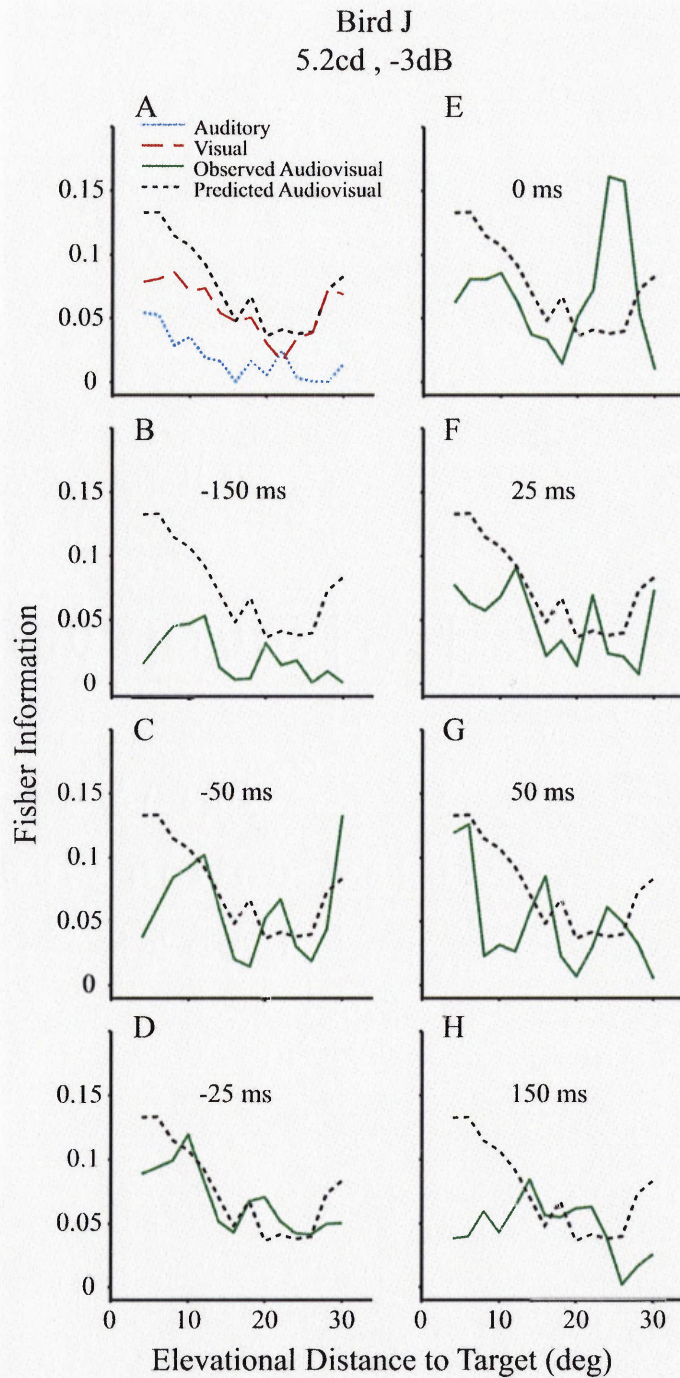


Figure 3-10. Fisher information – Examples from bird ‘J’.

Fisher information for the elevational component was computed for auditory (blue), visual (red), and audiovisual (green) trials within 9 degree sliding windows across target eccentricity. Most SOAs tested (noted by the text in each panel) showed audiovisual fisher information close to that observed in visual alone trials. However the condition in which the auditory led by the greatest amount (A) showed fisher information more like what was observed for auditory alone trials.

MLE Prediction vs. Observed Fisher Information

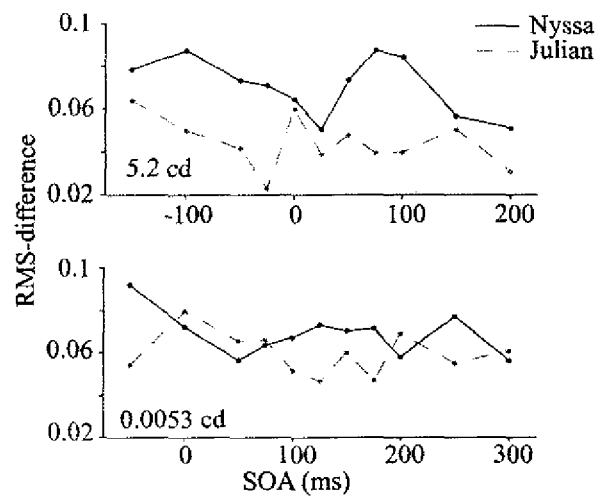


Figure 3-11. Summary of MLE fit to observed data.

Saccade Kinematics

Saccades were analyzed for the presence of a second peak in speed (see Methods). Figure 3-12A shows an example head saccade that was found to have a second acceleration phase, and thus a second peak in speed prior to the end of the saccade. The saccade shown in Figure 3-12A begins near the top right corner of the graph and moves downward and leftward toward the target (red cross). The kinematic details of this saccade are shown in Figure 3-12B.

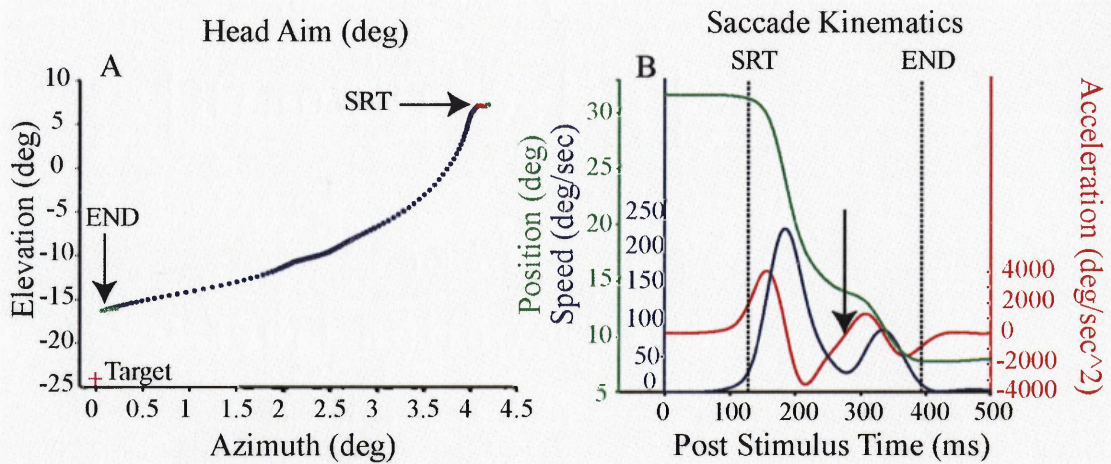


Figure 3-12. Example of double-step saccade.

In some fraction of trials, saccades were characterized by a second peak in the speed of the head movement (blue trace, B). The saccade in A is an example of this, showing the azimuthal and elevational location of head aim through time (468 Hz sampling rate). The saccade begins at the top right corner of the panel and proceeds toward the target (pink cross). This saccade was identified in an algorithm designed to detect those saccades with a substantial acceleration after a substantial deceleration (black arrow, B). The saccade was made by bird 'N' in a bright audiovisual trial with an SOA of -150 ms, auditory leads.

The black arrow marks the point at which the second acceleration began (Figure 3-12B, acceleration – red trace). For reference, the head position relative to the target is plotted in green and the instantaneous speed is plotted in blue. SRT and saccade end point are marked with the vertical dotted black lines. For simplicity, those saccades identified to have a second velocity peak prior to the official endpoint will be termed “double-step” saccades. For both birds, the probability of observing a double-step saccade was higher with larger SOAs (Figure 3-13).

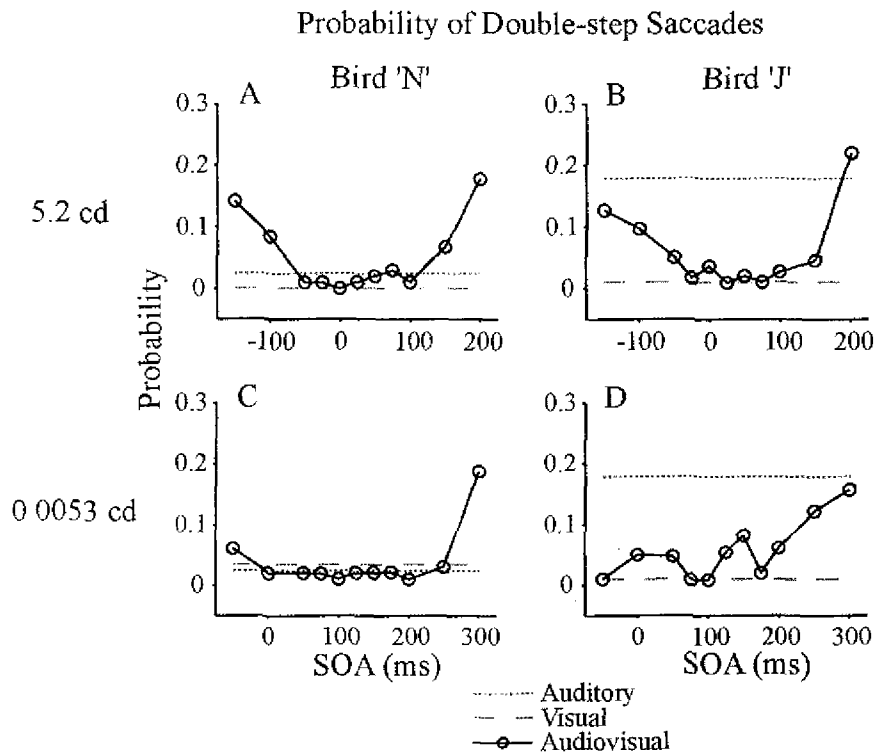


Figure 3-13. Observed probability of double-step saccades. Across subjects and intensity combinations, those trials with more extreme SOAs showed an increase in the observed probability of a double-step saccade. Double-step saccades were identified with an acceleration criterion, as described in Figure 12.

I hypothesize that double-step saccades are more likely with extreme SOAs because it is at these time delays that saccades initiated by the first modality may have progressed far enough so as to render the delayed modality ineffective in the current saccade. Because the second modality provides further information about the target however, the bird utilizes the information in the form of a second step. Figure 3-14 illustrates this hypothesis. The onset of stimulus one and stimulus two ($S1_{ON}$ and $S2_{ON}$) are shown to be separated by some extreme SOA value. At some point after $S1_{ON}$, the processing of the first stimulus is complete and a saccade is triggered ($S1_{SRT}$).

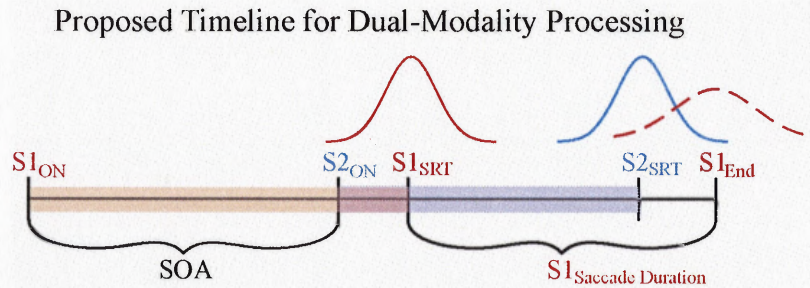


Figure 3-14. Timeline for parallel processing of two target signals.

In conditions of extreme SOA, there arose a chance that the onset of a response to a delayed modality ($S2_{SRT}$) would not occur until after the completion of a response to the modality presented first ($S1_{End}$). The delayed modality SRT was used as a measurable endpoint to the processing time required for that signal (possible processing time – shaded bars between ‘On’ and ‘SRT’). Given the observed distribution of unisensory SRTs (solid schematic Gaussians above time line), the known SOA, and the distribution of saccade durations (dashed schematic Gaussian), I predicted in what percentage of audiovisual trials the response might reflect this later processing of the delayed modality.

Theoretically, the maximum possible saccadic processing time can be estimated as the time between stimulus onset and reaction time (highlighted in the broad bars). Of course, processing must be complete well before the actual SRT, so SRT provides a starting point for the simulation and a theoretical upper limit. The process leading up to the saccadic response is a probabilistic process, and therefore a distribution of observed SRTs is depicted above the timeline for each stimulus (solid Gaussian curves). The time between the beginning and end of the saccade, or saccade duration, is also a probabilistic event and has its own associated distribution (dashed Gaussian curves). If there existed a non-zero probability that the onset of a response to a delayed modality ($S2_{SRT}$) would occur after the predicted completion of a response to the first modality ($S1_{End}$), the bird might be given cause for a second corrective or reactive saccade. This would correspond to a double-step saccade. Conversely, in the less extreme SOA conditions, the two modalities would

conceivably both be processed prior to either saccade initiation or saccade completion, and therefore both modalities could be integrated into one smooth saccade.

Using SRT as the measure for processing times, the probability of a double-step saccade was estimated from the underlying unisensory SRT and duration distributions. The race model provided the framework for the designation of which modality was S1 and which was S2, as described in the Methods. Figure 3-15 shows the results of this simulation.

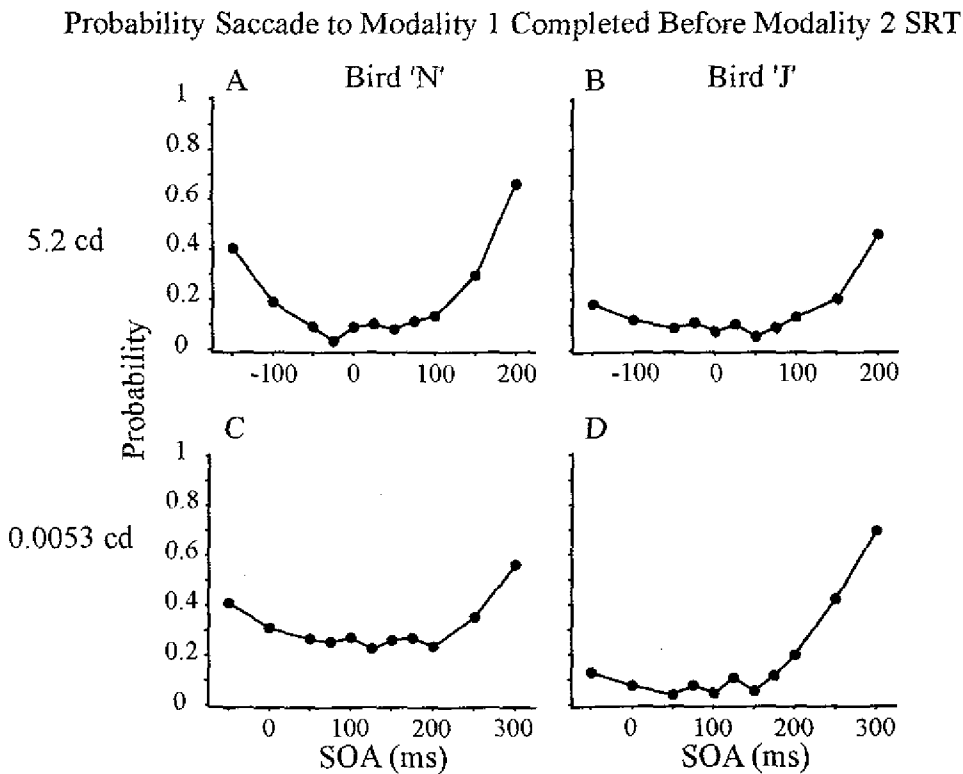


Figure 3-15. Delayed modality processing may occur after the completion of the saccade. With the hypothesized timeline shown in Figure 10, I predict the probability that the second stimulus will not be processed in time to affect the saccade initiated to the initially presented modality. For the extreme SOA conditions, this suggests that in some percentage of audiovisual trials, the auditory and visual stimuli will not be fused into a single smooth saccade.

In comparison to the observed probability of double-step saccades in audiovisual trials (Figure 3-13), the simulation was, in general, a severe overestimation for the larger SOA conditions. However, the general shapes of the observed and predicted curves were similar.

Because SRT was known to be an overestimation for the processing time of the second modality, I sought to optimize the simulation by decreasing the processing time of the second modality in increments of 10 ms. Thus a constant (ω) was introduced into the simulation. This constant was subtracted from the randomly chosen SRT value for the delayed modality and the remaining time was assumed to be this modality's processing time. If $S1_{\text{SRT}} + S1_{\text{Saccade duration}}$ was shorter than $S2_{\text{SRT}} - \omega$, the saccade was predicted to be a double-step saccade. The values of ω ranged from 0 to 90 ms and the predictions were compared to the observed probabilities. The root-mean-squared difference between the observed and predicted modalities across SOA were computed for each value of ω and plotted in Figure 3-16. The shortest ω -value that minimized RMS-difference between the observed and predicted trends were taken as the optimal, and ranged from 30 to 80 ms (open circles on each panel in Figure 3-16).

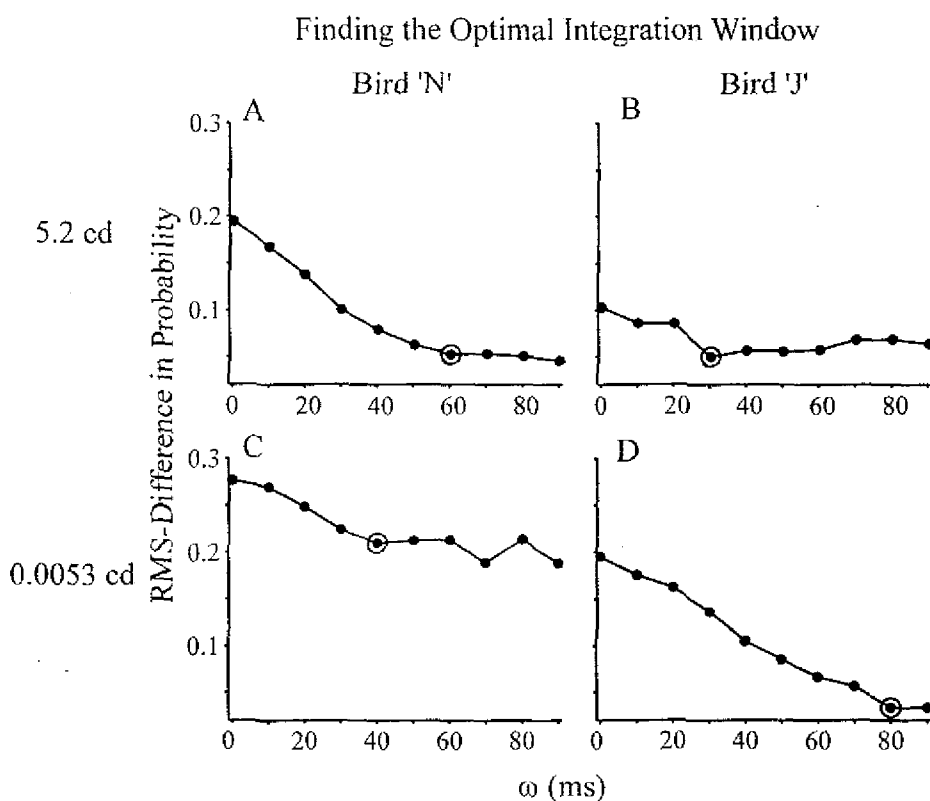


Figure 3-16. Finding the optimal integration window.

Given that SRT is the point at which a measurable response is present, it is possible that the point marking the end of signal processing for a given modality happens prior to SRT. The simulation outlined in Figures 10 and 11 was repeated while parametrically decreasing the proposed processing time of the delayed modality relative to SRT in increments of 10 ms (ω - units on abscissa). The root-mean-squared-difference between the predicted and observed probability of a double-step saccade across SOAs was then plotted for each bird and intensity combination as a function of ω . The window (ω) corresponding to earliest best-fit between the predicted and observed probability of double-step saccades is outlined in each panel (open circle).

The results from the simulation corresponding to the optimal ω -value are shown relative to the observed probabilities in Figure 3-17. For each condition and each bird, this simulation, based only on the unisensory timelines, predicted values near to the observed probability of double-step saccades in multisensory trials. One possible exception was the dim condition for bird 'N' for which, regardless of ω -values, the simulation continually overestimated the probability of a double-step saccade.

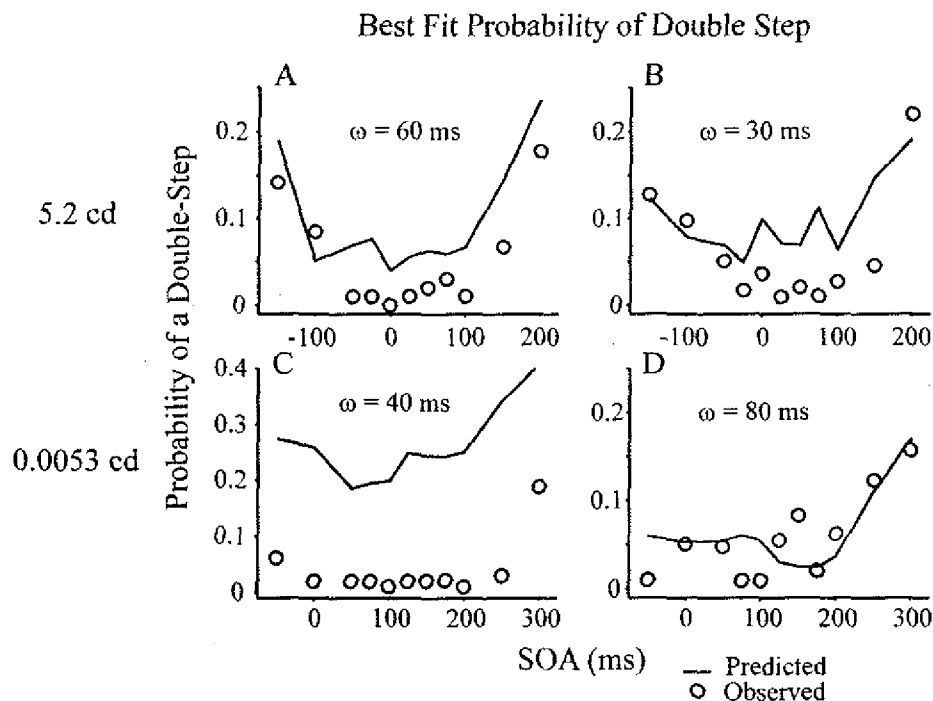


Figure 3-17. The best-fit prediction of double-step probability.

The observed probabilities of double-step saccades were best approximated with a simulation assuming the processing of the second modality occurs 30-80 ms prior to saccade initiation. Best-fit ω values are shown for each intensity combination for each subject. Of note was the failure of the simulation to more closely approximate observed responses in the dim condition for bird 'N' (Panel C). Although the general shape of the prediction followed observed trends across SOA for this condition, the overall predictions were high relative to observed values.

Discussion

As I observed with synchronous audiovisual stimuli (Chapter II), saccades to asynchronously presented audiovisual targets were initiated as predicted by the race model. Similarly, error followed the more accurate and precise modality for the majority of asynchronous conditions. This was somewhat of a surprise, given the wide range of SOAs used in this study. From the unisensory SRT and duration distributions, I initially predicted that the most extreme SOA conditions would be more unisensory in nature. However, I found that the birds were able to incorporate a later arriving second stimulus into a smooth saccade in all but the most extreme SOA cases. In these extreme SOA conditions, the birds responded

with double-step saccades, suggesting a time window for integration of the second modality into a single, smooth saccade. This time window began with the SRT of the first stimulus ($S1_{SRT}$) and theoretically ended with the end of the saccade ($S1_{END}$). If the predicted $S2_{SRT}$, shortened by w (30 – 80 ms), fell in this window, the two modalities were integrated into a smooth saccade. Conversely, if it fell outside of this window, a double-step saccade was likely. In either case, I conclude that the birds utilized both modalities by whatever means necessary to optimize the accuracy of the final head aim and obtain the reward.

SRT

In Chapter II, I reported the presence of negative race model violations: Significant portions of the SRT distribution were longer than the Grice boundary of the race model. Differences in the method of analysis may explain the lack of negative race model violations seen in the SOA study described here. Examination of Figures 3-3 and 3-4 show that there were some observed audiovisual distributions with distinctly jagged cumulative distribution functions that would deviate from the Grice boundary, similar to those seen in the negative violation cases reported in Chapter II. However, when a boot-strap simulation was done to estimate 1000 possible distributions from the underlying unisensory SRTs, I found that a number of predicted distributions similarly violated the Grice boundary. It follows that when I compared the observed audiovisual distributions to the 95% confidence intervals of the predicted distributions, there was no substantial deviations in either the positive or negative directions. This illustrates the power of the simulation used in this SOA study. Were there more trials in each condition, it is possible that the theoretical method of comparing the observed audiovisual SRTs against the Grice and Miller Boundaries would be more similar to the boot-strapping simulation results.

Previous studies in mammals have shown a significant decrease in SRT beyond race model predictions when stimuli were offset in time by the difference in unisensory processing time (Corneil et al., 2002; Colonius and Diederich, 2004). In the current study, I did not see this positive violation of the race model: Observed audiovisual SRT distributions were not significantly shifted to the left of the 95% confidence intervals produced by the race model simulation. I can not explain this difference in results, and therefore look forward to more behavioral and physiological experiments to probe the issue further.

Error

Previous studies of multisensory integration in spatial perception have shown that a maximum likelihood estimation adequately describes the integration of the sensory information (Ernst and Banks, 2002; Alais and Burr, 2004). In the current study however, the barn owls use the visual modality almost exclusively and do not combine the unisensory fisher information in a way that is consistent with the maximum likelihood estimation. This observation may be explained by the fact that the spatial certainty of the visual modality was not sufficiently decreased so as to release visual capture of the perception. It is known that if the certainty of the visual modality is high enough, little to no weight is given to the accessory modality in localizing the stimulus (Battaglia et al., 2003; Heron et al., 2004). Future behavioral studies will incorporate a spatially diffuse, yet easily detectable visual stimulus to more effectively probe this possibility.

Another difference between this study and those mentioned above is that I presented the auditory and visual components of the audiovisual stimulus from the same location in space. Instead of inducing a spatial discrepancy within the stimulus, I relied on the bird's natural tendency to undershoot eccentric auditory targets more than visual ones as a means by which to separate out a predicted saccade error. This caveat may have rendered the predicted audiovisual error statistically equivalent to the visual alone error. Future behavioral experiments will include audiovisual trials that incorporate a minimal spatial separation.

The Significance of the Double-Step Saccade

It is remarkable that the best-of-both modalities phenomenon reported in Chapter II was observed even when the auditory and visual components were presented with delays as long as 150 ms. Thus, at almost every SOA in which audition was expected to win the race (See Figure 3-5), accuracy and precision were better than those in unisensory auditory trials (Figure 3-8, Figure 3-9).

Saccade trajectories differed for trials with short and long SOAs despite the fact that the end results – fast, accurate saccades – were the same. At the short SOAs, the saccades accelerated, reached peak velocity, and then decelerated so that plots of velocity vs time had a single peak. This is consistent with a scenario in which visual information, although arriving later, still arrives early enough to be incorporated into a single smooth motion characteristic of unisensory saccades. At the long, negative SOAs, it became possible that the visual information arrived after a critical time point in saccade control and the double-

stepped saccades may indicate the only temporal option in incorporating the later arriving visual signal.

It is important to note in this context that auditory saccades were often hypometric. The first deceleration may have occurred when the owl had concluded that its saccade accurately localized the sound, which may have been represented hypometrically within the OT. The re-acceleration may have occurred when the owl had “noticed” the LED, which though co-localized with the speaker, may have been represented in the OT without the hypometria. Thus the double-step saccade may be an indication of those trials in which the bird perceived two distinct targets, giving rise to two distinct velocity peaks. Alternatively, the bird may have perceived two distinct targets in SOA conditions much closer to zero but was able to incorporate the visual target into a smooth saccade. In this case, the double-step saccade would indicate those conditions in which there was a temporal impossibility of incorporating a second stimulus in an ongoing saccade. From these experiments, there is no way to determine whether the double-step saccade marked instances where bird perceived two stimuli or instances where the bird was just not physically able to incorporate the later stimulus into a smooth saccade. Future experiments utilizing paradigms in which the bird is trained to respond differently to a single target vs. a double target would be useful in teasing out more details in this regard.

What Happens if Audition Loses the Race

In simplest terms, the data from Chapter II suggests that audition triggers the saccade and vision refines it, so that audiovisual saccades are both fast and accurate. SOAs can be

inserted, however, that cause audition to lose the race. One might expect then that saccade errors would reflect the later-arriving modality, audition. Interestingly, however, I did not observe audition like errors in trials with large vision-leading SOAs (Fig. 3-8; Fig. 3-9.) In bird 'N', audiovisual errors were less than auditory, but greater than visual, and in bird 'J', audiovisual errors were in fact less than even vision. Thus, the potentially-degrading effect of audition was mitigated in bird 'N'. In bird 'J', by contrast, auditory information may have actually helped to further refine the target's position. Such results compel us to investigate the neural representation of audiovisual stimuli as it evolves over the course of the stimulus having different SOAs.

How was it possible that the later arriving auditory signal could improve the accuracy of the audiovisual response in bird 'J'? Perhaps the answer lies with the timeline of stimulus processing (Figure 3-13). In many of these large SOAs, it is probable that some portion of the 100 msec sound was presented while the head was in motion, mid-saccade, toward the target. Thus for some of the trials, birds were receiving dynamic auditory cues. Previous reports of auditory error have only been from trials with near-zero pre-stimulus head motion (this chapter and Chapter II). It is possible however, that motion enhances accuracy in localizing auditory targets. Furthermore, the degree of hypometria associated with auditory trials depends on the target eccentricity, with closer targets evoking more accurate saccades (Figure 3-7). If the head was in the process of moving closer to the target when the sound was perceived, any error introduced from the auditory signal would be less than what might be expected from a sound placed at the original target eccentricity.

The Double-Step Model

By incorporating the ω -value, my simulation closely predicted the probability of observing a double-stepped saccade in audiovisual trials, with the exception of the dim conditions in bird 'N'. Bird 'N' was generally less likely to respond to the dim visual stimulus, suggesting that it may not have seen the visual stimulus in a proportion of the audiovisual trials (See Figure 3-4A, dashed distribution). The simulation to predict the probability of a double-step saccade incorporated the overall percent response to each unisensory stimulus. However, the fact that the bird was operating on a level so near threshold may have affected the behavior beyond what was predictable by the model.

The best-fit ω -value from the simulation is of particular interest in the context of sensory-motor integration in saccade generation. Its relevance in the simulation is two-fold. First, as initially posed, the ω -value incorporates the idea that the observable SRT is an overestimation of the time required for the processing of the second modality to an extent capable of affecting the saccade. The ω -value might also incorporate however, the idea that the observable saccade duration is an overestimate of the window of integration in the production a smooth single saccade. Introducing the ω -value into the simulation can be interpreted as accounting for both of these overestimations. For example, the simulation predicts the probability of observing a double-step saccade by using the following rule:

If: $S1_{SRT} + S1_{SaccadeDuration} < SOA + S2_{SRT} - \omega$,
 Then: The saccade is predicted to be a double-step saccade.

However, this equation may be rewritten:

$$\text{If: } S1_{SRT} + S1_{\text{SaccadeDuration}} - \omega_1 < SOA + S2_{SRT} - \omega_2,$$

Then: The saccade is predicted to be a double-step saccade.

Thus ω , as used in the first equation, is simply the difference between ω_1 and ω_2 , values that compensate for the overestimation in the temporal integration window and the overestimation in the S2 processing time, respectively. The present observation can not, therefore, provide specific insight into the exact processing time required for the second modality ($SRT - \omega_2$), nor can it tell exactly the time point that an ongoing saccade can no longer be smoothly altered ($S1_{SRT} + S1_{\text{SaccadeDuration}} - \omega_1$). I conclude only that ω_2 is greater than ω_1 , and that the overall temporal alignment of the two stimuli must fall within a certain time window to ensure smooth integration into a single saccade.

CHAPTER IV

NEUROPHYSIOLOGICAL CORRELATES OF BEHAVIORAL CHANGES WITH
SOUND PRESSURE LEVEL**Introduction**

As shown in Chapter II, saccadic error and reaction time in the barn owl increases when softer sounds are used as targets. Additionally, the probability of initiating a saccade at all (response probability) decreases as quieter stimuli are presented. Why are these near-threshold stimuli less efficient in evoking a response? What is it about the internal sensory representation that changes when quieter sounds are used? I asked this question by measuring single unit responses to relatively loud and quiet sounds in the barn owl's midbrain.

The auditory space map, located at the level of the external nucleus of the inferior colliculus (ICx, Chapter I), is the first point along the auditory pathway at which space is represented. Single units from this nucleus have been extensively studied in an effort to ascertain the fundamental neuronal building blocks of the internal representation of auditory space (Moiseff and Konishi, 1981, 1983; Takahashi, 1989; Carr and Konishi, 1990; Wagner, 1993; Zheng and Knudsen, 2001; Euston and Takahashi, 2002; Bala et al., 2003; Spezio and Takahashi, 2003; Takahashi et al., 2003). The ICx projects directly to the optic tectum, a nucleus known to be central in orientation control and saccade generation. Thus the ICx is an optimal location to record sensory responses to both louder and quieter sounds, with the

explicit purpose of deciphering a connection between behavior and internal sensory representations.

This chapter includes the following important points: In the Methods section (Data analysis), I first present a detailed account of how SPL was estimated for near-threshold sounds presented in virtual auditory space (VAS). Next, I explain how first-spike latencies and cellular thresholds were detected above spontaneous firing rates. In the results section, cellular thresholds for the population of recorded cells are summarized and compared to behavioral saccadic thresholds and detection thresholds. Whereas saccadic threshold was discussed thoroughly in Chapter II, the presentation of detection thresholds is new, and includes original data I collected as a rotation student in Dr. Terry Takahashi's lab, in collaboration with Dr. Avinash Bala. Thus in addition to the main electrophysiological experiment, the methods section briefly outlines the behavioral paradigm and data analysis from this rotation project. Next, first-spike latencies in response to loud and quiet stimuli are compared. Finally, I present spatial response profiles taken at various SPLs to illustrate the distinct and predictable change in the spatial response profile as sounds approach cellular threshold.

Methods

All experiments were carried out under protocols approved by the University of Oregon Institutional Animal Care and Use Committee and in compliance with *The Guide for the Care and Use of Laboratory Animals*, (Institute of Laboratory Animal Resources (U.S.) and NetLibrary Inc., 1997).

Apparatus and Stimuli

Stimulus synthesis, data acquisition, and data analysis were all carried out with Matlab 6.5 (The Mathworks). Experiments were conducted in an anechoic chamber, similar to the one explained in Chapters II and III. The bird was suspended in an immobilization swing, while the head was fixed firmly in place to a stereotaxic device by means of a surgically-implanted head-plate on the skull (see Chapter II Methods). This kept the bird's head immobile throughout the course of an experiment. Miniature Sony headphones (MDR-E272) were fitted with 1.4 cm long Delrin cones that fit snugly within the outer canal, and these ear-phones were used to present sounds in VAS. VAS was created by convolving the auditory signals with location specific head-related impulse responses (HRIRs). In those cases where noises were presented in the free field, a lightweight speaker (Peerless, 5.08 cm cone tweeters) was mounted on a movable hoop approximately 0.9 m from the bird's head.

Individualized HRIRs were obtained in an initial experiment session, as discussed in Keller, et. al., 1998. Briefly, sounds were presented from the frontal hemifield at a constant angular separation of 5° (double polar coordinates) and recorded with miniature Knowles microphones (EM4046) fitted with a 3.2 cm long probe-tube (Clay Adams PE160, ID 1.14 mm, OD 1.57 mm). Recorded spectrums were then converted to the time domain, and the energy within a 3.6 ms sliding window was calculated and processed. The filtering effects of the loudspeaker, microphones and probetubes were then removed (Keller et al., 1998). All sounds in VAS were filtered through these impulse responses designed to mimic the natural filters of the head and ruff (Chapter I).

Auditory stimuli consisted of broad band noises (2-12 kHz), 100 ms in length, varying in spatial location as well as sound pressure level (SPL). These sounds were the same frozen sounds used in the behavioral experiments described in Chapters II and III. Spike times were recorded for 400 ms, an epoch comprising a 100 ms pre-stimulus recording period, a 100 ms period during which the sound was presented, and a 200ms post-stimulus recording period.

For calibration purposes, SPLs of sounds presented from free-field speakers were measured with a ½ inch microphone (Brüel & Kjaer model 1760) and sound-level meter (Brüel & Kjaer model 2235). Sounds were also recorded inside the ear canal, 2 mm in front of the eardrum, in both the free-field and VAS with miniature Knowles microphones (see previous paragraph for specifications).

PDR Paradigm

Two barn owls were acclimated to lie still in a stereotaxic device (identical to the one described above) without the use of anesthetics. Sounds (100 ms, 3-11 kHz) were presented from speakers placed either directly in front, 45° above, or 45° below the bird, all at 0° azimuth (midline). The birds' pupillary dilation responses (PDR) were measured, as in Bala and Takahashi, 2000. Briefly, an infrared light-emitting diode (LED; F5D1QT; emission peak at 880 nm; QT Optoelectronics, Sunnyvale, Calif.) and an infrared detector diode (QSC114; absorption maximum at 880 nm; QT Optoelectronics, Sunnyvale, Calif.) were positioned 2.5 and 5 mm, respectively, from the cornea. The output from the detector was amplified and digitized at 1 kHz. When a stimulus was detected, the bird's pupil

automatically dilated relative to the level of dilation just prior to stimulus onset, causing a net positive change in the detector voltage (Figure 4-1, solid line). If the bird did not detect the stimulus, a continuation of the natural oscillation of pupil dilation was recorded (dotted line). Across many trials, no-response recordings averaged out to baseline (dashed line).

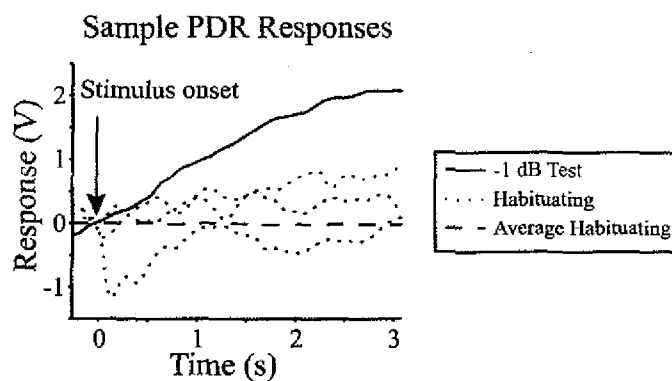


Figure 4-1. Raw voltage traces from a PDR experiment. The instantaneous voltage was recorded in PDR discrimination experiments and plotted in time. Three examples of dilation responses to habituating stimuli (well below threshold) are shown in dotted lines. The average across hundreds of habituating stimuli (dashed line) is compared to the voltage trace in response to one test stimulus presented at -1 dB SPL_A. The significant increase in voltage marks an instance in which the bird noticeably detected the stimulus.

Electrophysiological Paradigm

Data was collected from 8 captive-bred barn owls (*Tyto alba*), anesthetized with periodic intramuscular injections of ketamine (KetaVed, Vedco; 0.1 ml at 100 mg/ml) and valium (diazepam; 0.08 ml at 5 mg/ml), as needed (approximately every 3 hours). The birds' heart-rate and temperature were monitored throughout the course of the experiments, as well as visually monitored with an infrared camera. The brain was accessed through previously fitted recording wells, located bilaterally above the forebrain just dorsal to the ICx. Cells were isolated by moving a 250 μm -diameter tungsten electrode (Frederick Haer, Brunswick, ME) in the dorsal-ventral dimension in increments of 5-15 μm , while searching for stimulus-

locked neuronal responses. From this approach, the ICx was approximately 13 to 15 mm below the dura of the forebrain.

Systematic tests began once a clear threshold level was obtained on the spikes evoked by the search stimulus, and the spike was of constant shape and size across repetitions. The first test measured spike rates in response to sounds presented well above the dynamic range (approximately 24 dB) in a checkerboard pattern across space, separated by $5\text{-}10^\circ$ (double-polar coordinates). This test determined the best area from which to present future sounds. Each location was presented twice, and the location that responded with the greatest average spike rate across the 100 ms stimulus period was termed the cell's best location.

The next test presented noises from the best location that ranged in SPL from approximately 24 to -12 dB. Each SPL was presented 10-20 times. The data from this test was used to determine the cell's first-spike latency and rate-level function (see Data Analysis below).

The last test consisted of a randomized presentation of pre-synthesized sounds from locations in and around the best area, at SPLs in the dynamic range of the cell. All conditions were repeated at least 5 times, and most often 20 times.

Data Analysis

Calibration SPL in VAS

The calculation of near-threshold SPLs in VAS required three steps. First, sounds with equal, but unknown original SPLs were attenuated with the same equipment used to

attenuate the virtual sounds and presented through a speaker at audible levels to be recorded by the B&K sound-level meter (see Apparatus and stimuli, above). Figure 4-2 shows the recorded SPLs (ordinate) corresponding to 11 different attenuations from 4 spatial locations (directly in front of the microphone, 30° up, 30° down, and 30° to the right). The same attenuations at each location resulted in equal SPL readings, and there was a direct relationship between attenuation and measured SPL (fitted solid line). This relationship was extrapolated down to the SPL region used in the experiments to follow, and I predicted that a 76 dB attenuation of the original sound would result in a measured SPL of -9 dB (large diamond), 0 dB being human threshold. Although this calculation assumes a perfectly linear attenuator across a wide range, the assumption is necessary, given the limited sensitivity of the SPL-meter: The sound floor of this meter in this environment was approximately 30 dB SPL, in practice.

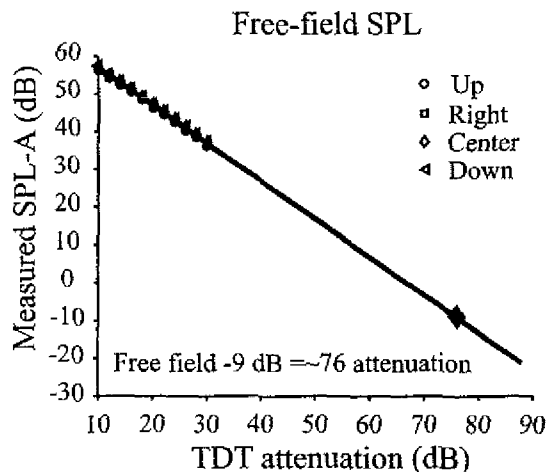


Figure 4-2. Free-field extrapolation of SPL. A-weighted SPL measurements were taken from sounds presented within the SPL range of the microphone from four locations (30° up, down, right and left). SPLs were extrapolated to levels below human threshold (0 dB), showing that an attenuation of 76 dB from maximum was equivalent to -9 dB SPL_A.

Once an absolute SPL was assigned to sounds presented from a speaker at a given attenuation, I set out to relate this to the SPL at the eardrum in response to a sound presented through headphones in VAS. To do this, I used miniature Knowles microphones, inserted into the ear canal and affixed approximately 2 mm in front of the eardrum (see Apparatus and stimuli, above). Sounds were presented from a speaker positioned at one of 5 locations (30° up, 30° down, 30° left, 30° right, and directly in front of the bird). Recorded voltages were filtered offline with a high pass filter (corner frequency = 2 kHz) and the RMS voltage was recorded at each ear, converted into dB and averaged binaurally. As shown in Figure 4-3A, the average binaural RMS voltage changed linearly with attenuation in the region significantly above the background noise. The average x-intercept was extrapolated out, using the loudest sounds that were in the linear range of the microphone and found to be 79 dB.

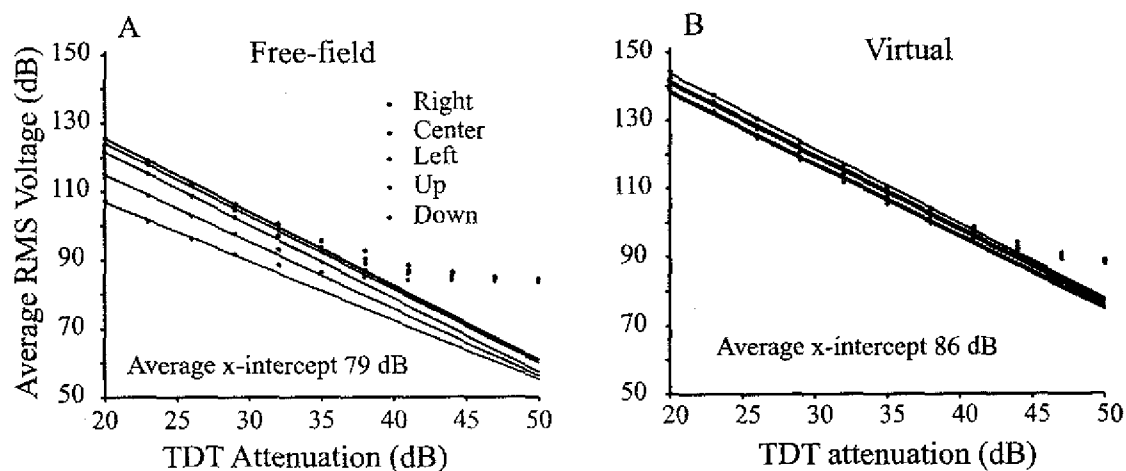


Figure 4-3. Measurements of free-field vs. virtual sounds from within the ear canal. Sounds of varying attenuations were presented from a speaker placed at 5 locations (legend, A) and in virtual space, through earphones (B). The average x-intercept for the free-field measurements was 79 dB, whereas the average x-intercept for virtual sounds was 86 dB. Therefore, a sound presented from earphones is on average 7 dB louder at the ear drum than the same sound (with the same mechanical attenuation) presented from a speaker in the open environment.

The same measurements were taken in response to virtual sounds. Earphones were inserted in the outer ear canal and sounds were presented to the bird from the virtual equivalents of the 5 free-field locations used previously. Figure 4-3B shows the recorded voltage dependency on attenuation to also be linear above the sound floor. The average x-intercept in the virtual condition was 86 dB, 7 dB greater than the free-field calculation. Thus from these calculations, I estimate that virtual sounds are approximately 7 dB louder at the eardrum than equally attenuated free-field sounds.

To further test this 7 dB relationship between free-field and VAS SPLs, I used a biological relative SPL meter, an auditory neuron from the barn owl's space map. Spike rates to various sound pressure levels (rate-level responses) were measured in response to both free-field and virtual stimuli in a single ICx neuron. Figure 4-4A shows the raw data (open circles) and fitted sigmoids (solid curve), \pm 95% CI (dashed curves) describing how spike rate changed with SPL. The half way point between saturation (plateau on the left side of the graph) and no response (plateau on the right side of the graph) is marked with the vertical dashed line at an attenuation of 57 dB in the free-field condition (Figure 4-4A). Figure 4-4B shows analogous data from sounds presented in virtual space at the same location to the same cell. The 50% point for the virtual condition was found to be at an attenuation of 64 dB. As in the microphone measurements in Figure 4-3, the difference between the virtual and free-field conditions was 7 dB, supporting the idea that virtual sounds are approximately 7 dB louder than equally attenuated free-field sounds. Given my initial measurements of SPL in free-field conditions (Figure 4-2), and the 7 dB difference between free-field and virtual sounds of the same attenuations (Figures 4-3 and 4-4), I

estimated the absolute SPLs of the sounds in space represented by the virtual sounds, and placed this SPL scale (grey) below the attenuation scale (black) on the abscissas of figures 4-4A and 4-4B. This direct relationship between attenuation and SPL is utilized in the remainder of the analyses described in this chapter.

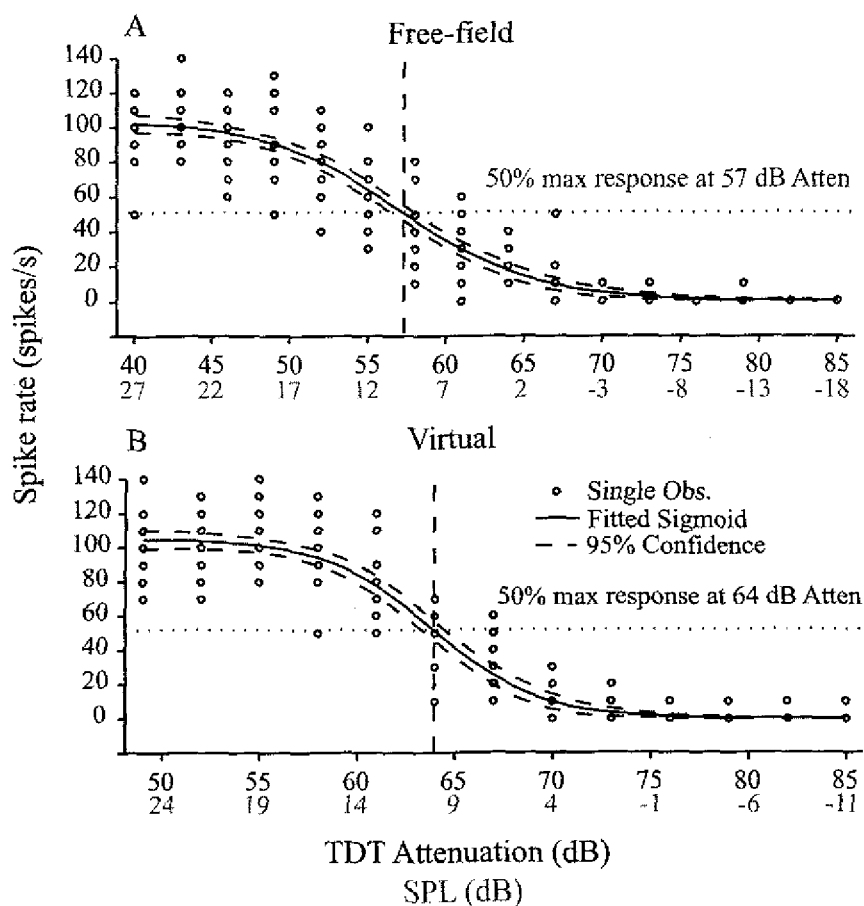


Figure 4-4. Neuronal comparisons of free-field vs. virtual sounds.

Sounds of varying attenuations were presented from a speaker placed at an ICx neuron's best location (A) and from the same location in virtual auditory space (B). Each circle represents the spike rate for one trial. The data were fitted with a sigmoid, and the attenuation at which the response was 50% of the maximum response, as compared to spontaneous, is marked for each condition (vertical dashed line). The 7 dB difference in the virtual and free-field 50%-points support the observation that virtual sounds are approximately 7 dB louder than free-field sounds (see Figure 4-3). From the extrapolation illustrated in Figure 4-2, the absolute SPL (dB) is reported on the abscissa in grey.

First-Spike Latency

Spike times were recorded relative to stimulus onset. A 100 msec pre-stimulus period was included in the recording to obtain spontaneous firing rates. The first spike latency (or simply “latency”) of each cell was assigned based on the time required, post-stimulus onset, for the average spike rate (10 ms sliding window) to exceed the average spontaneous rate. Latencies were calculated based on responses to relatively loud sounds (>20 dB SPL).

The inset in Figure 4-5 is a raster plot illustrating a typical ICx neuron’s response to sounds of various SPLs (ordinate) presented at that unit’s best location. Each horizontal bar (alternating grey and white) highlights the 20 repetitions of the same SPL. SPLs ranged from relatively loud (34 dB bottom) to relatively quiet (-6 dB, top). The stimulus onset and offset are marked with solid black vertical lines, and labeled “on” and “off” respectively. The vertical dark grey bar illustrates the sliding 10 ms window that was used to bin the spikes and calculate the local spike rate, shown in the main plot of Figure 4-5 (34 dB SPL condition from the inset). The local spike rate in each time bin was averaged across repetitions and compared to the pre-stimulus spike rate. The pre-stimulus spike rate was used as a measure of the cell’s spontaneous spike rate. The cell’s latency (Figure 4-5, arrow) was calculated as the first time, post-stimulus onset, at which the local spike rate rose above 3 STD spontaneous rate (Figure 4-5, dot-dashed line).

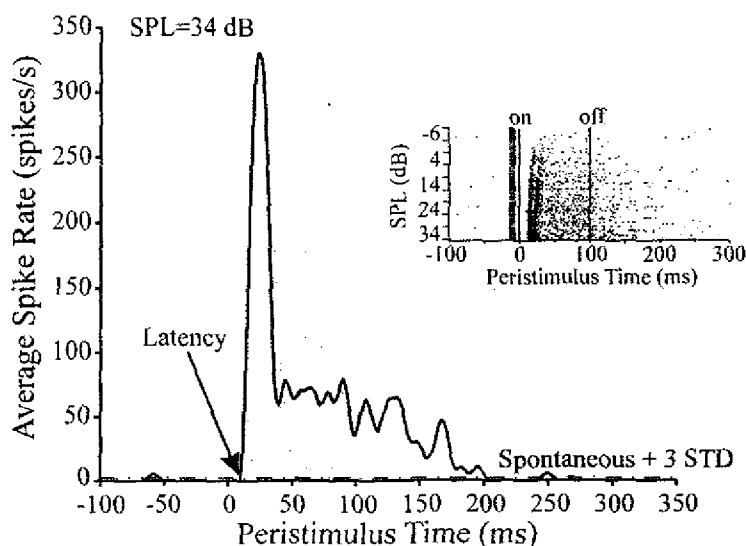


Figure 4-5. ICx neurons response to sound presented from the best area.

This peristimulus time histogram shows the average spike rate of a typical ICx neuron in response to a relatively loud sound (34 dB). The inset illustrates the spikes obtained from 20 repetitions of a pseudo-random noise presented at 9 different SPLs (bottom to top, loudest to softest). Stimulus duration is shown by the space between the vertical solid lines in the inset (100 ms). A sliding 10 ms averaging window (vertical grey bar – inset) was used to find the first point post-stimulus onset (arrow – latency) that the average spike rate rose 3 STD above the pre-stimulus spike rate (horizontal dot-dash line – main figure). This latency was used as a starting point for a 100 ms window (grey area – main figure), across which the spike rate was computed for each rep.

Cellular Threshold

Rate-level relationships were generated by comparing overall spike-rates in response to sounds. The grey shaded region in the main plot of Figure 4-5 illustrates the time period used to calculate the overall spike rate of the cell in response to a given stimulus. This epoch is the size of the stimulus (100 ms) and begins with the cell's characteristic latency, measured for the loudest presented sound. The overall spike rates for each SPL (same cell as Figure 4-5 inset) are plotted in Figure 4-6 (dots). These points represent the average spike rate across all 20 repetitions. A sigmoid function was fit to the data, and the point at which this fitted line crossed 3 STD of the spontaneous rate indicated cellular threshold (-7 dB for the cell in

Figure 4-6). The cellular threshold was calculated in this way for all 121 cells subjected to this rate-level test.

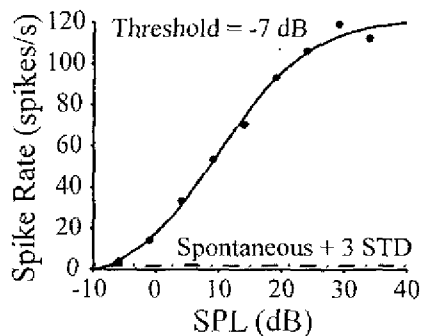


Figure 4-6. Rate-level plot. Average spike rates were computed across a 100 ms window, beginning with the cell's characteristic latency at the loudest sound presented (Figure 4-5, main plot, grey area). Data across SPL and repetitions were fit with a sigmoid. The point in the sigmoid that first exceeded the spontaneous rate + 3 standard deviation, was labeled the cell's threshold (-7 dB for this cell).

Pupillary Dilation Response (PDR)

The bulk of the stimuli used here were well below threshold (estimated approximately -42 dB, relative to human threshold) and were termed habituating stimuli. Examples of these trials are shown in Figure 4-1 (dotted lines). Test stimuli ranged near behavioral threshold from -25 to +6 dB. Inter-stimulus-intervals were set to 5 s. The response to each stimulus was summarized by integrating the area between the voltage trace and baseline (Figure 4-1, between solid and dashed lines) for the first 2 seconds after stimulus onset. Figure 4-7B shows examples of these integrated responses. Grey points represent habituating trials. Black and red symbols represent data from test trials. The absolute SPLs of the test trials are labeled on the abscissa of Figure 4-7A. These data represent one session, from one bird. All stimuli in this session were presented 45° below the bird. The overall z-score for each test SPL value is plotted in Figure 4-7A. Overall, sounds presented at -5 and -1 dB were statistically distinguishable from the habituating response (z-score, red cross and red square, respectively).

Standard separation (D) was used to as a non-parametric index of discrimination, describing the relationship between the habituating stimulus and the test stimulus. Figure 4-9 utilizes the D-statistic to summarize data from two locations (center, and 45° above center). The D-statistic was computed as follows:

$$D = \left| (\mu_h - \mu_t) / \sqrt{(\sigma_h \sigma_t)} \right| \quad \text{Eqn. 4.1}$$

Where μ_h and μ_t refer to the mean magnitudes of the PDR to habituating and test stimuli, and σ_h and σ_t are the respective STDs. D-values greater-than, or equal-to 1 mark test SPLs that were statistically discriminable from the habituating stimuli. Thus the first test SPL (as SPL increases from well below threshold) to evoke a D-value of 1 or greater was noted as the detection threshold.

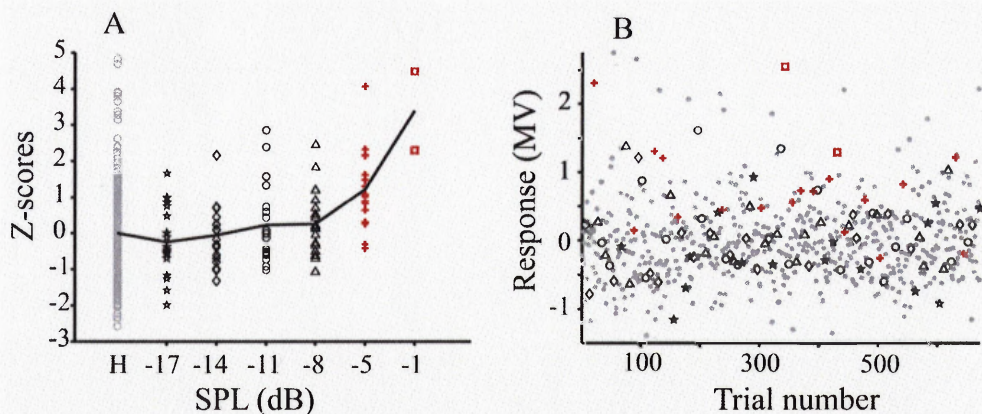


Figure 4-7. Integrated PDR responses to stimuli of varying SPL.

Individual voltage traces (illustrated in Figure 4-1 in response to habituating and test stimuli) were integrated across time from stimulus onset to 2 s post-stimulus onset (B). Red symbols represent trials from test conditions that were generally distinguishable from the habituating trials (grey). Black symbols represent trials from test conditions that were not distinguishable. SPLs for the test stimuli are noted in the abscissa of A. Habituating trials were presented well below threshold, approximately -42 dB.

Results

Cellular Thresholds

I isolated 149 cells in 8 birds, and tested rate-level relationships in 121 of these (Figure 4-6).

From this analysis, each cell's SPL threshold was obtained. The summary of these values across the population of recorded cells is shown in the histogram in Figure 4-8. Of these cells, 41% had cellular thresholds below -5 dB. This approximates the SPLs at which the probability of responding with a saccade to an auditory stimulus decreases below 100% (Figures 2-3, 2-4).

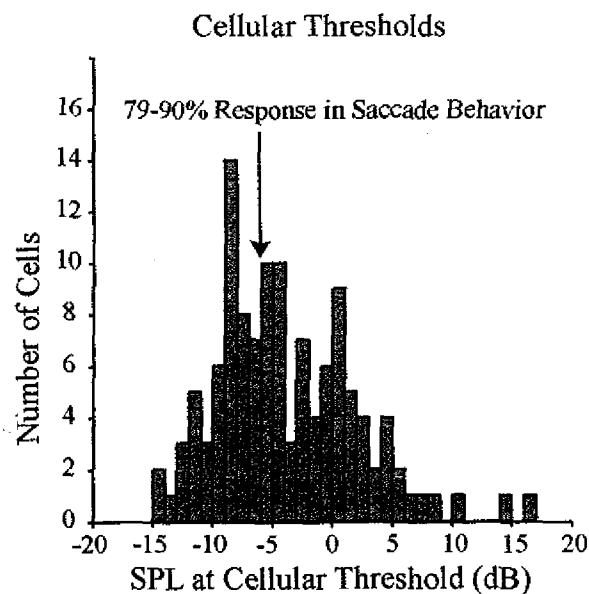


Figure 4-8. Population summary of cellular threshold.

Cellular SPL thresholds were computed, as shown in Figure 4-6, for 121 ICx cells. 41% of cells had thresholds below -5 dB SPL_A. Interestingly, saccadic behavior (Figures 2-3, 2-4) shows that the probability of responding to a sound presented at -6 dB is just below 80% for bird 'N', data shown, and around 90% for bird 'J'.

This also corresponds to the behavioral detection thresholds measured in awake, restrained birds performing the PDR paradigm (see Methods). Figure 4-9 shows the discrimination value (variation of z-score, see Methods) for each presented test stimulus, averaged across

birds and sessions. Detection thresholds to stimuli presented directly in front of the birds were near -11 dB (D-value of 1). Interestingly, the detection threshold observed for stimuli presented above the bird were approximately 9 dB louder than that observed for trials directly in front of the bird (Figure 4-9, solid line, -2 dB SPL). This is discussed in more detail in the Discussion section below.

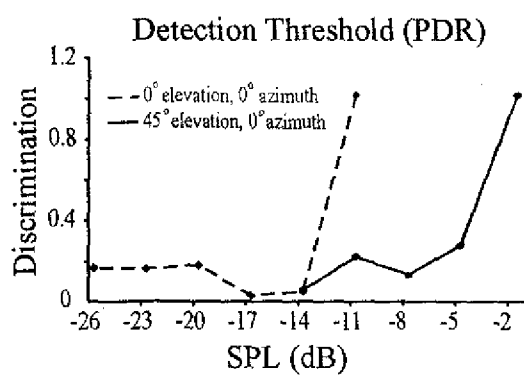


Figure 4-9. Behavioral discrimination thresholds. Discrimination measurements illustrate that a test sound from directly in front of the bird (0° azimuth, 0° elevation, dashed) must be at least -11 dB SPL, to be distinguishable from the habituating stimulus. The SPL of a sound placed 45° above the bird however (solid), must be at least -2 dB for the bird to detect it.

First-Spike Latency

Of the 121 cells for which rate-level functions were obtained, 41 were presented sounds at -5 dB and 25 dB, and had cellular thresholds below -6 dB SPL. The first spike latencies of these cells to the louder (25 dB) and softer (-5 dB) stimuli are plotted in the histogram in Figure 4-10. On average, cells responded with a longer latency to the -5 dB sounds (dot-dashed line, mean 30.7 ms), relative to the louder sounds (solid line, 13.7 ms). Although this difference of 17 ms is striking, it can not account for the observed increase in SRT in response to quieter sounds. Figure 2-4A shows mean SRTs in response to auditory stimuli for bird 'N' (bird 'J' auditory SRTs were almost identical, data not shown). Here the mean SRT for louder sounds plateau around 55 ms, whereas SRTs are predicted to be around 150 ms for -5 dB sounds (Figure 2-4A, solid line). Thus the 100 ms difference in behavioral SRT

must reflect the longer processing time required to initiate a saccade to quieter sounds, perhaps in addition to this increase in first-spike latency.

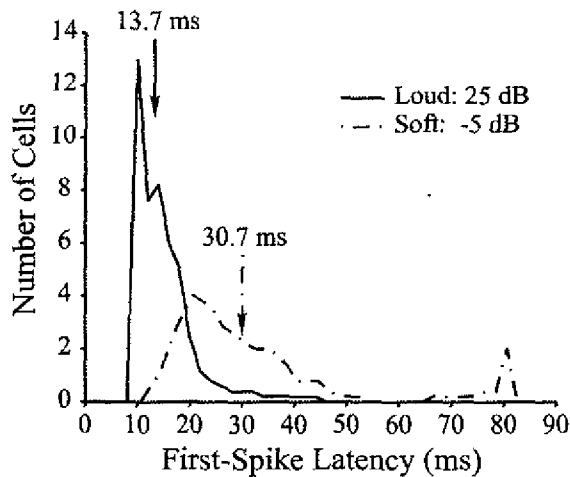


Figure 4-10. First-spike latencies of neurons with low threshold. Neurons with cellular thresholds below -6 dB were pooled, and the first-spike latency of a loud stimulus (25 dB, solid) was compared to that of a soft stimulus (-5 dB, dot-dashed). On average, loud stimuli evoked a response 17 ms faster than soft stimuli. The data was obtained from 41 neurons.

Spatial Response Profiles

Average spike rates across the duration of the stimulus (offset by the first-spike latency, see Methods) were used to describe the spatial response profiles of 76 neurons at different SPLs. Figure 4-11 is an example of the measured spatial response profiles of a single ICx unit to 3 distinct SPLs in the dynamic range of the cell. Only a subset of frontal locations were tested (outlined in the black square), comprising locations in and around the cell's best area, as defined by the initial coarse space test (see Methods). The spatial resolution of this sampling was 5°, in a checkerboard pattern. The missing data points within this area were interpolated by taking the average of the surrounding measured data points. Spike rate color coded, relative to the peak response for that SPL (red, absolute values noted above the graphs) and the spontaneous rate (blue, 20 spikes/s). A grey scale was used to illustrate spike rates below the spontaneous rate. The peak response in space is marked with a black cross.

The peak of the cell's spatial response profile in Figure 4-11 moved centrally as the stimuli decreased in SPL. This was observed, qualitatively, in approximately 40% of the neurons. In a larger proportion of cells, the spatial response profile was substantially skewed toward the center as SPL decreased, regardless of whether or not the peak was shifted. These findings were more likely for neurons with more peripheral best locations.

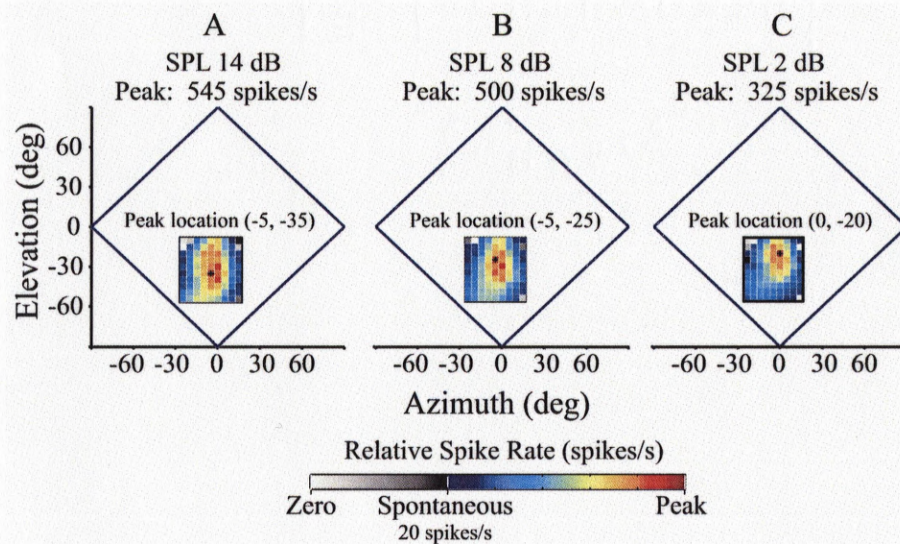


Figure 4-11. Spatial response profiles vary with SPL. ICx neurons responded differentially to sounds presented across space. Spike rates were maximal when the sound was presented from a certain spatial location in virtual space (black cross). This best location was not necessarily constant across SPLs. Responses to 3 SPLs are shown here: 14 dB (A), 8 dB (B), and 2 dB (C). The best area tended to be more central in lower SPL conditions: loud (-5° azimuth, -35° elevation) vs. soft (0° azimuth, -20° elevation). The color-bar shows relative spike rate above and below the spontaneous spike rate (20 spikes/s). The general excitatory receptive field also shifted centrally in quieter conditions (compare green and yellow areas of A through C). This data was taken from a single unit, with 20 repetitions of each SPL at 5° -offset resolution. Missing points were interpolated from recorded data. Only locations from within the black square were presented.

Discussion

Sounds presented at or near behavioral threshold are known to evoke saccades with delayed reaction times and increased errors, as compared to those evoked by louder sounds. What is the neural correlate to this behavioral observation? In this chapter, I summarize three key differences in the sensory representations of near- and well-above threshold sounds at the level of the barn owl's auditory space map (ICx). I noted that cellular thresholds illustrate that as SPL decreases below -5 dB, fewer and fewer neurons respond to the stimulus with an increase in spike rate. Furthermore, first-spike latencies were longer in quieter sounds. Finally, the spatial response profiles of neurons to quiet sounds were spatially shifted relative to those of louder sounds. Although direct correlations may be made between these behavioral and electrophysiological observations, causality remains an answered question.

Linking Cellular to Behavioral Thresholds

Cellular thresholds were generally found to be distributed around -5 dB, SPL_A. This is approximately 10-15 dB quieter than what has previously been reported for auditory nerve fibers in the barn owl (Köppl and Yates, 1999), although Köppl and Yates report absolute-SPL, and not SPL_A. The SPL_A scale utilizes a frequency weighting system based on the spectral properties of the human ear, and thus low frequency components of sounds are essentially removed from the overall SPL measurement. The A-weighted scale is useful in the present study because these stimuli did not include frequencies below 2 kHz, and thus any power in this region of the recorded spectrum was associated with ambient noise. Inclusion of this low frequency noise would lead to higher overall SPL readings. Possible alternative explanations for this discrepancy might include the fact that the ICx is

substantially downstream of the auditory nerve fibers and responses in these cells utilize data pooled from across many frequencies. Thus it is conceivable that the cellular thresholds of ICx neurons reflect this compiling of data and subsequently show substantially lower thresholds than auditory nerve fibers.

Interestingly, the percentage of neurons responding to a sound of a given SPL was tightly correlated to the probability that the animal would detect and/or localize that stimulus. Saccade %-response had a very short dynamic range, going from near 100% response at SPLs of -3 dB to almost no-response at SPLs of -12 dB. This 9 dB range marked the behavioral limits in localizing quiet sounds for the two birds studied in the saccade task. Similarly, two different birds, studied in a completely separate involuntary discrimination task, showed an ability to detect -11 dB sounds when the stimuli were presented directly in front of the bird (a condition, of course, not tested in the saccade task). However, that detection threshold increased to approximately -5 dB for sounds positioned 45° below, and -1 dB for sounds presented 45° above the birds. The relevance of these detection values at different locations will be discussed below, but overall, detection thresholds were nearly identical to saccade thresholds. The fact that fewer cells are likely to respond to SPLs presented between -3 and -12 directly corresponds with the observed behavioral thresholds and suggests there is a direct correlation between the number of simultaneously active cells in the ICx to a given stimulus and the probability that the bird will detect and localize it.

The Relevance of the First-Spike Latency

It has been suggested that a decrease in the first-spike latency of the sensory response in the OT is directly correlated to a decrease in saccade reaction time – SRT (Bell et al., 2005). We know from Chapter IV and 5 that first-spike latencies of neurons in the OT and ICx in response to auditory stimuli are tightly correlated. Therefore I probed the latency differences of ICx neurons in response to loud (25 dB) and quiet (-5 dB) sounds. Latencies to louder stimuli averaged around 13.7 ms, whereas those of the same cell to quieter sounds averaged around 30.7 ms. This increase in latency (17 ms) was substantial but did not equal the difference in SRTs seen in loud vs. quiet trials (~100 ms).

The origin of this increase in latency is thought to be at a level in auditory processing much more peripheral than the ICx or OT. For instance, it has been shown that the timing of the first-spike latency varies systematically with SPL in auditory nerve fibers of the cat (Heil and Irvine, 1997). In a relatively comprehensive review of the first-spike latency in auditory neurons, Heil describes a temporal integration theory that attempts to correlate first-spike latency with:

... the calcium-binding steps to calcium sensors at the domains associated with the individual synapses between the inner hair cell and the afferent fibers ... (Heil, 2004).

Although it is tantalizing to speculate on a cochlear origin of the observed increase in first-spike latency observed here, I conclude only that this difference did not arise at the level of

the mid-brain, and that it is only indirectly correlated with the observed increase in SRT. In addition to the 17 ms neuronal delay imposed on behavioral responses to quiet sounds, the decrease in overall activity in the ICx across time, as shown by the rate-level responses in Figure 4-6 and the spatial response profiles in Figure 4-11, most probably underlies a decrease in the rate of rise to activation threshold and ultimately an increase in the observed SRT, (Carpenter and Williams, 1995).

SPLs at the Eardrum Depend on Stimulus Location

The detection thresholds measured for sounds above and below the bird were 6 and 8 dB higher than that measured directly in front of the bird (Figures 4-7, 4-9). This is not altogether surprising, given the nature of the barn owl's ears and facial ruff (See Chapter I). Figure 4-12 illustrates this point directly.

Subplots A-D in Figure 4-12 show the relative average binaural levels (ABLs) for narrow band sounds, ranging from lower to higher frequencies, presented in the frontal hemi-field. These plots were obtained using the head-related-transfer-functions (HRTFs), derived by transforming the head-related-impulse-responses (HRIRs) measured for each bird (See Methods) from the time domain into the frequency domain. Thus HRTFs are frequency—specific mathematical filters that mimic the filtering properties of the head and facial ruff. Figure 4-12 shows the relationship between average binaural level (ABL) at the eardrum and space for 4 frequencies known to be within the range used by the barn owl in spatial localization.

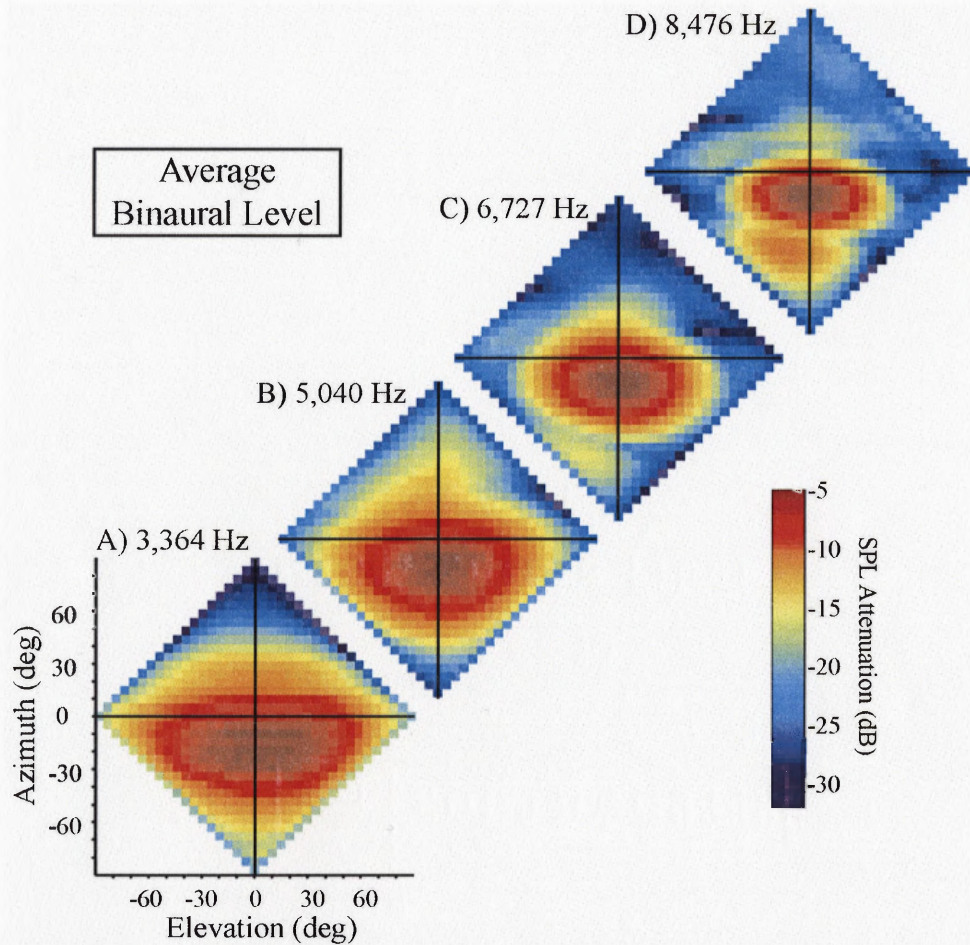


Figure 4-12. HRTFs attenuate frequencies differentially, depending on spatial location. Sounds of certain frequencies are attenuated more when they are presented from the more peripheral locations. The colorbar shows the amount of attenuation, blue being more attenuation and red less attenuation. Thus a sound presented directly in front of the bird is louder at the eardrum than one presented 45° above the bird. Although this attenuation depends on frequency (see subplots), central locations are generally less attenuated than peripheral ones.

The color codes for attenuation, with red signifying less attenuation (sound is louder at the eardrum) and blue signifying more attenuation (sound is quieter at the eardrum). The entire frontal hemifield is represented in these diamond plots, sampled at 5 degrees constant angular separation (double-polar coordinates).

Immediately obvious is the red area encompassing and directly below the center of space (0° azimuth, 0° elevation). Relative to the rest of space, these locations are attenuated less across frequencies. It is important to note the differences in Figures 4-12A through 4-12D because ICx neurons do not equally weight frequency inputs (Euston and Takahashi, 2002; Spezio and Takahashi, 2003; Keller and Takahashi, 2005). For example, many weight frequencies between 6 and 8 kHz more strongly than others, and thus the difference in response between a sound above and a sound directly in front may be greater for this particular cell than what would be seen for a cell that weights lower frequencies more strongly. Across cells however, there exists an overall predicted average decrease in spike rates in response to sound presented peripherally, as compared to the same sound presented centrally, when that sound is within the dynamic range of the cell, or near behavioral threshold.

I propose that this decrease in predicted neuronal activity, as well as the difference in detection threshold, correlate with the dependency of ABL on space. From the HRTF data alone, I predicted the 6 and 8 dB difference in SPL seen in the detection task. Similarly, with the frequency-weighting of any given cell a detailed description of that cell's characteristic rate-level response, it is possible to predict the general shift in spatial response profiles illustrated in Figure 4-11. An important question remains: What does this mean for localization? The key to answering this question lies with the ability to record from multiple isolated units simultaneously, while the animal is behaving. Future electrophysiological experiments of this kind utilizing multi-electrode technology in animals during a localization task will undoubtedly address this issue more directly.

CHAPTER V

AUDIOVISUAL INTEGRATION IN THE OPTIC TECTUM OF THE BARN OWL

Introduction and Background

Chapters II and III illustrate that barn owls are able to utilize both auditory and visual modalities in localization behavior, when doing so increases the chances of success in a task, linked ultimately to a food reward. This strongly supports the theory that the barn owl *integrates these keenly adapted senses in natural behavior*. The experiments in Chapter IV were aimed at correlating specific neuronal response properties with different aspects of auditory localization behavior. These results suggest a link between the nature of the internal sensory representation in time and an animal's ability carry out an effective behavioral response. To address both of the issue of audiovisual integration (Chapters II and III) and sensory-motor transformation (Chapter IV) simultaneously, I begin, in this chapter, to examine single unit responses in an area of the midbrain known to be at the crossroads of both the audiovisual and sensory-motor pathways, the optic tectum (OT).

As described in Chapter I, the OT is a highly laminated midbrain structure, receiving both auditory and visual inputs while making downstream connections with brainstem nuclei responsible for proper motor responses in orientation. A basic topography is conserved in neuronal responses to both modalities across the OT, and this topography is aligned with the OT motor map, thought to drive orientation behavior in a host of animals, including the

barn owl (Hartline et al., 1978; Bastian, 1982; Masino and Knudsen, 1992; Nemeč et al., 2001; Valentine et al., 2002; Burnett et al., 2004).

Audiovisual Integration in the Barn Owl

The barn owl OT has been studied extensively with regard to its role in the visual calibration of auditory space in development (Knudsen and Knudsen, 1983, 1989; Knudsen, 1999; Hyde and Knudsen, 2001; Zheng and Knudsen, 2001; Gutfreund et al., 2002; Hyde and Knudsen, 2002; Bergan et al., 2005). Important conclusions from this body of work include the following: 1.) When the visual modality is skewed by affixing prisms to the barn owl's head during development or even as an active adult, the auditory space map in the ICx shifts in the same direction by a comparable amount. 2.) The OT is proposed to be the source of the instructive visual signal that shapes the auditory space map with visual experience. 3.) There is *reciprocal anatomical connection* between the OT and the ICx, and furthermore, when inhibition was artificially relaxed with injections into the OT, visual responses were seen in ICx cells.

Throughout the barn owl literature however, there is only one paper that addresses audiovisual responses in an isolated OT cell: Knudsen's 1982 paper in the *Journal of Neuroscience*, entitled "Auditory and visual maps of space in the optic tectum of the barn owl". In this paper, Knudsen describes in great detail single unit responses to flashed lights and sound bursts. For instance, these single units had overlapping auditory and visual receptive fields, however the overlap was not found to be 100%; Cells with more peripheral auditory excitatory receptive fields often had visual receptive fields that were shifted centrally

in comparison. Furthermore, auditory first spike latencies in multisensory cells tended to be approximately 50-60 ms earlier than visual latencies, relative to stimulus onset. Finally, it was noted that the more dorsal layers of the OT responded primarily to light, whereas the intermediate and deep layers responded to sound and many times both light and sound. The Knudsen 1982 paper does not, however, discuss responses to simultaneously presented audiovisual stimuli. In a 1995, *Annual Review of Neuroscience* paper, Knudsen and Brainard relate the barn owl optic tectum with the mammalian superior colliculus (SC), stating that the term "tectum" could be applied to "both the superior colliculus of mammalian species and the optic tectum of other classes of vertebrates" (Knudsen and Brainard, 1995). With this caveat, they went on to describe "Visual-Auditory Integration in the Tectum", citing studies primarily carried out in mammalian systems, including the cat, guinea pig, and non-human primate. Thus it is commonly assumed in the literature that findings from mammalian SC regarding the nature of audiovisual integration apply to other species, including the barn owl. However, it has not been directly tested and published, to date.

Audiovisual Integration in the Mammalian SC

Studies of audiovisual integration in the mammalian SC are indeed numerous, for review see (Stein and Meredith, 1993; Calvert et al., 2004). Most of these studies record average spike rate in response to auditory, visual, and audiovisual stimuli presented at one location in space and compare audiovisual rates with unisensory rates, or predictions based on unisensory rates. Throughout the course of these studies, three general principles have been repeatedly reinforced: 1.) The temporal alignment of the auditory and visual components of

audiovisual stimuli will affect the nature of audiovisual integration. Audiovisual stimuli that are misaligned in time such that the neuronal latencies are compensated for (Visual leads by approximately 100 ms, for instance) are more likely to show spike rates that exceed those of either unimodal response, or in other words audiovisual responses are “enhanced” (Meredith et al., 1987). 2.) The spatial alignment of the auditory and visual components of audiovisual stimuli will affect the nature of audiovisual integration. When unimodal stimuli were presented in a temporally ideal fashion from their respective best areas, audiovisual responses were more likely to be enhanced (Meredith and Stein, 1986b, 1996). Because excitatory spatial receptive fields measured in auditory and visual unisensory tests generally overlapped, this translated into the concept that spatially aligned stimuli, or stimuli that were relatively close in space, were more likely to evoke an enhanced response. Conversely, when stimuli were grossly misaligned, spike rates often plummeted to levels below those seen in either unisensory condition. 3.) Audiovisual stimuli that were presented at levels close to neuronal threshold (very quiet or very dim) were more likely to evoke an enhanced spike rate relative to the most effective unisensory modality alone (Wallace et al., 1996). This final principle has been termed the “inverse effectiveness rule”.

These three principles were correlated to early behavioral experiments in the cat (Stein et al., 1988; Stein et al., 1989), and they are reviewed in detail in Chapter II of this dissertation.

Briefly, cats were trained to approach a lever when the light and/or sound associated with that lever was presented. When presented alone, the auditory and visual stimuli were fairly ineffective at evoking a correct response (pressing the correct lever). However, when the stimuli were presented from the same location in the correct temporal arrangement, the cat’s

percent correct increased to levels beyond the mathematical sum of the unisensory responses. An interesting caveat is that trials in which the cat did not respond at all *as well as* trials in which the cat pressed the incorrect lever were considered incorrect responses. Thus an increase in correct response might mean an increase in the probability of detection, or it might mean an increase in orientation accuracy, or both. The enhanced responses observed in the single SC recordings were then correlated with this enhanced probability of correctly orienting toward a weak audiovisual stimulus.

An interesting twist to the story developed with a paper published in 2002 by Populin and Yin, entitled "Bimodal interactions in the superior colliculus of the behaving cat". This paper called into question two ideas that had previously gone relatively unchallenged. The first was the idea that comparing the audiovisual response to the best unisensory response was not truly a measure of audiovisual enhancement. Instead they proposed a comparison between audiovisual responses to the *sum* of the unisensory responses as a measure of enhancement. This definition may be more directly comparable to previous behavioral observations in cats (Stein et al., 1988; Stein et al., 1989). With this reevaluation of the definition of neuronal enhancement, Populin and Yin go on to suggest that true enhancement is only very rarely observed in a behaving animal. Prior to this 2002 publication, very few studies had used unanesthetized animals. Instead, the vast majority of these studies were in animals anesthetized with ketamine and curare. Populin and Yin assert that the relevance of multisensory enhancement, as previously defined, was overestimated and an artifact of the inadequate definition of enhancement, the anesthesia, and the lack of a

summary of just how many cells were truly “enhanced” in proportion to the number of cells from which recordings were taken.

Audiovisual Integration in Awake vs. Anesthetized Preparations

In the studies that do show population data from awake, behaving animals, no more than 25% of the cells recorded in the OT responded in a supra-linear fashion (Peck et al., 1995; Populin and Yin, 2002; Bell et al., 2003). Additionally, there is now evidence that the anesthesia influences the audiovisual interaction at the level of the superior colliculus in the cat (Populin, 2005). Most recently however, a population study in an anesthetized preparation showed that only 23.8% of recorded responses to audiovisual stimuli were significantly supra-linear (Stanford et al., 2005). In fact, the majority of the population (69.4%) was not significantly different from the predicted linear sum. This study was unique in that it incorporated response variance into the classification of additivity, and this statistical method conceivably depressed the number of reported supra-linear responses relative to previous studies in anesthetized preparations. In light of this new statistically sound method of classification and the realization that anesthesia changes the nature of audiovisual integration in these cells, previous studies may support the hypothesis that the majority of responses to audiovisual stimuli can be predicted by the sum of the component auditory and visual response.

Audiovisual Integration and the Sensory-Motor Transformation

Very few studies have attempted to directly link aspects of audiovisual behavior with neuronal sensory responses. One such study, published by Bell, et. al. in the Journal of

Neurophysiology (2005), relates neuronal responses in the OT of awake, behaving, rhesus monkeys with saccade reaction times to visual and audiovisual stimuli. In this study, the subjects were trained to localize visual stimuli regardless of the presence of an auditory component. The auditory components of audiovisual trials were sometimes spatially aligned and sometimes misaligned with the visual stimulus. Aligned audiovisual SRTs were found to be significantly shorter than SRT in visual alone trials. The difference in mean SRT between the two conditions was 5-10 ms. The neuronal correlate of this behavior was two-fold. 1.) In low intensity combinations, the first-spike latency was seen to decrease significantly with the addition of an aligned auditory component, relative to visual alone (decrease of 5-10 ms). However the first-spike latency of an auditory alone stimulus was not measured or reported. Only 9 out of 89 recorded cells responded to the auditory stimulus, as determined by first-spike latencies more characteristic of an auditory stimulus (less than 70 ms). Because auditory first-spike latencies in the SC are generally much shorter than visual however, the mean first-spike latency of a given recorded population would greatly depend on the proportion of cells with sensory responses to auditory stimuli. A previous study cataloguing the proportion of cells in the deep layers of the rhesus SC that responded to both auditory and visual stimuli found that only 37% showed sensory responses to vision alone (Wallace et al., 1996). Given that 89% of the cells recorded in the Bell, 2005 study only had sensory responses to visual stimuli, the mean first-spike latency of the population from this study is not a viable correlate to the decrease in behavioral SRT observed, which presumably arose due to information from a population of cells with a greater percentage responding to auditory stimuli.

The second neuronal correlate of the shortening of mean SRT observed in the Bell 2005 study was a significant increase in pre-motor activity of neurons in high intensity conditions. Contrary to the previously proposed neural mechanism of SRT being linked to first-spike latency, this correlation may actually underlie a phenomenon that has been previously hypothesized to directly influence SRT: the rise to activation discussed in saccade generation models (Carpenter and Williams, 1995).

The present and final data chapter of this dissertation reports preliminary evidence of audiovisual integration in the optic tectum of the barn owl. I tested whether or not a neuron in the barn owl OT responded to audiovisual stimuli presented in its best area in a manner similar to what has been reported for anesthetized mammalian preparations. Although these results do confirm that barn owl OT units respond as predicted from mammalian studies, they do little to address the issue of audiovisual integration in saccade generation. Ideally, future electrophysiological paradigms in awake animals will be designed specifically to test how the entire audiovisual spatial response profile differs from unisensory (auditory and visual) spatial response profiles.

Methods

All experiments were carried out under protocols approved by the University of Oregon Institutional Animal Care and Use Committee and in compliance with *The Guide for the Care and Use of Laboratory Animals*, (Institute of Laboratory Animal Resources (U.S.) and NetLibrary Inc., 1997).

Apparatus and Stimuli

The apparatus and general nature of the acoustical stimuli are identical to those described in Chapter IV. Stimulus synthesis, data acquisition, and data analysis were all carried out with Matlab 6.5 (The Mathworks). Experiments were conducted in an anechoic chamber, similar to the one explained in Chapters II and III. The bird was suspended in an immobilization swing, with the head fixed firmly to a stereotaxic device by means of a previously implanted head-plate on the skull (see Chapter II Methods). Miniature Sony headphones (MDR-E272) were fitted with 1.4 cm long Deltin cones that fit snugly within the outer canal, and these ear-phones were used to present sounds from virtual space. Virtual space was created by convolving the auditory signals with location specific head-related impulse responses (HRIRs). Sounds were attenuated in the manner described in Chapter IV, so all attenuations are reported as SPL here. Auditory stimuli consisted of broad band noises (2-12 kHz), 200 ms in length. Unlike previous chapters, the fine structure of each auditory stimulus was unique, and generated during the course of each experiment. Spike times recorded in the 100 ms preceding the onset of the stimuli were used for spontaneous rate calculations.

Visual stimuli consisted of red LED flashes, 200 ms in duration. As in Chapters II and III, the luminous intensity was 5.2 cd. The eyelid contralateral to OT being recorded was held open by taping thread to it and affixing the thread to the stereotaxic device. This did not hinder the occasional "blink" of the third eyelid, the nictitating membrane, which kept the eye moist throughout the course of the experiment. Audiovisual stimuli were presented simultaneously, each in their respective receptive fields. Generally auditory and visual stimuli were spatially aligned.

Electrophysiological Paradigm and Data Analysis

Single units were isolated in a manner identical to that described in Chapter IV, with the use of 100 ms auditory search stimuli. Once a unit was isolated, the best auditory area was mapped out and the attenuation was set well above neuronal threshold, generally at 17 dB SPL_A. Visual responses were verified with a quick visual alone-test, to ensure that the neuron was responsive to both modalities. Auditory and visual stimuli were presented separately and combined from the best area, at SPLs and luminous intensities well above neuronal threshold. Auditory, visual, and audiovisual stimulus was presented sequentially. Thus modality was not randomized. Spike rasters and first-spike latencies similar to those shown in Figure 4-5 were computed as described in Chapter IV. Average spike rate, as reported in Figure 5-2, was computed across the duration of the stimulus, beginning at the latency point, and averaging across repetitions. Response enhancement was computed using the following equation (Populin and Yin, 2002):

$$\% \text{ Enhancement} = \frac{R_{AV} - (R_A + R_V)}{R_A + R_V} * 100 \quad \text{Eqn 5.1}$$

R_{AV} , R_A , and R_V refer to the average response rate of observed audiovisual, auditory, and visual trials, respectively.

Results and Discussion

Spike rates of multisensory OT neurons were sometimes higher than responses to either modality alone (Figure 5-1). Instantaneous spike rates were calculated over a sliding 10 ms window from the data in Figure 5-1B and plotted relative to stimulus onset (left side of grey bar, Figure 5-1A). First-spike latencies to auditory stimuli (blue) were generally shorter than

those to visual (red, vertical bars, Figure 5-2B). First-spike latencies to audiovisual stimuli (green) were nearly identical to those in auditory trials. The predicted audiovisual response across time (black) was computed by taking the mathematical sum of the auditory and visual instantaneous spike rate, minus the spontaneous spike rate (horizontal dot-dashed line).

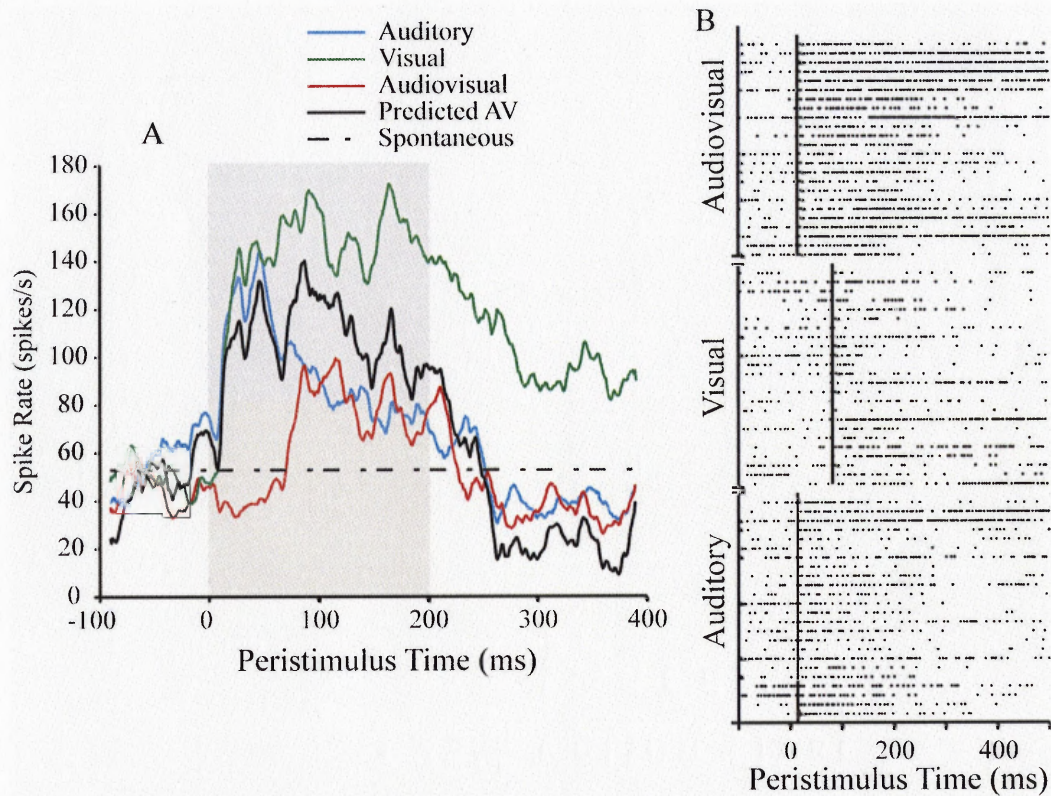


Figure 5-1. OT single unit responses in audiovisual trials.

The cell shown here is one that exemplifies the ability of neurons in the OT to respond with more than the predicted linear sum of the unisensory responses. Rasterplots of the raw data show multiple repetitions of the same stimulus condition. Stimuli were presented at time 0 and lasted 200 ms (grey bar). The horizontal lines in B show the computed first-spike latency for each modality in this cell average.

Across time, the cell in Figure 5-1 responded with a greater spike rate than what would be predicted from the linear sum of the auditory and visual responses. This corresponds with observations reported previously in mammalian anesthetized preparations (Introduction and

background). Not all neurons responded in this manner. In fact, only approximately 30% of recorded neurons showed augmented responses when the two modalities were combined. It is important to note that most cells were presented sounds at relatively high absolute SPLs. The auditory stimuli in the trials shown in Figure 5-1 were presented at 14 dB SPL. Because I did not measure rate-level functions for these cells, it is possible that this SPL was above the saturation level for most cells. However, for this cell, I was able to see a substantial increase in spike rate with the addition of the visual modality.

Figure 5-2 tests the 3rd principle of audiovisual integration, as presented by Stein and colleagues: inverse effectiveness. This principle is built around the observation that enhancement is more likely when the effectiveness of the unisensory stimuli is low. Figure 5-2 shows data for audiovisual combinations in which the auditory SPL was varied, while the visual intensity remained constant. The data for each SPL group was recorded in a single test, consisting of auditory, visual, and audiovisual stimuli presented in the cell's best area. The tests were done in order of highest to lowest SPL. Each modality was presented 10 times in each test. The insets above the bars show the raster plots corresponding to the average spike rates below. In the three loudest conditions (14 dB, 9 dB, and 4 dB), the audiovisual response (green) was greater than the sum of the unisensory responses (black bar), but not to a large extent. When the auditory SPL was dropped to -1 dB, the % enhancement rose dramatically (see Eqn 5-1, Methods, Figure 5-2 above rasters).

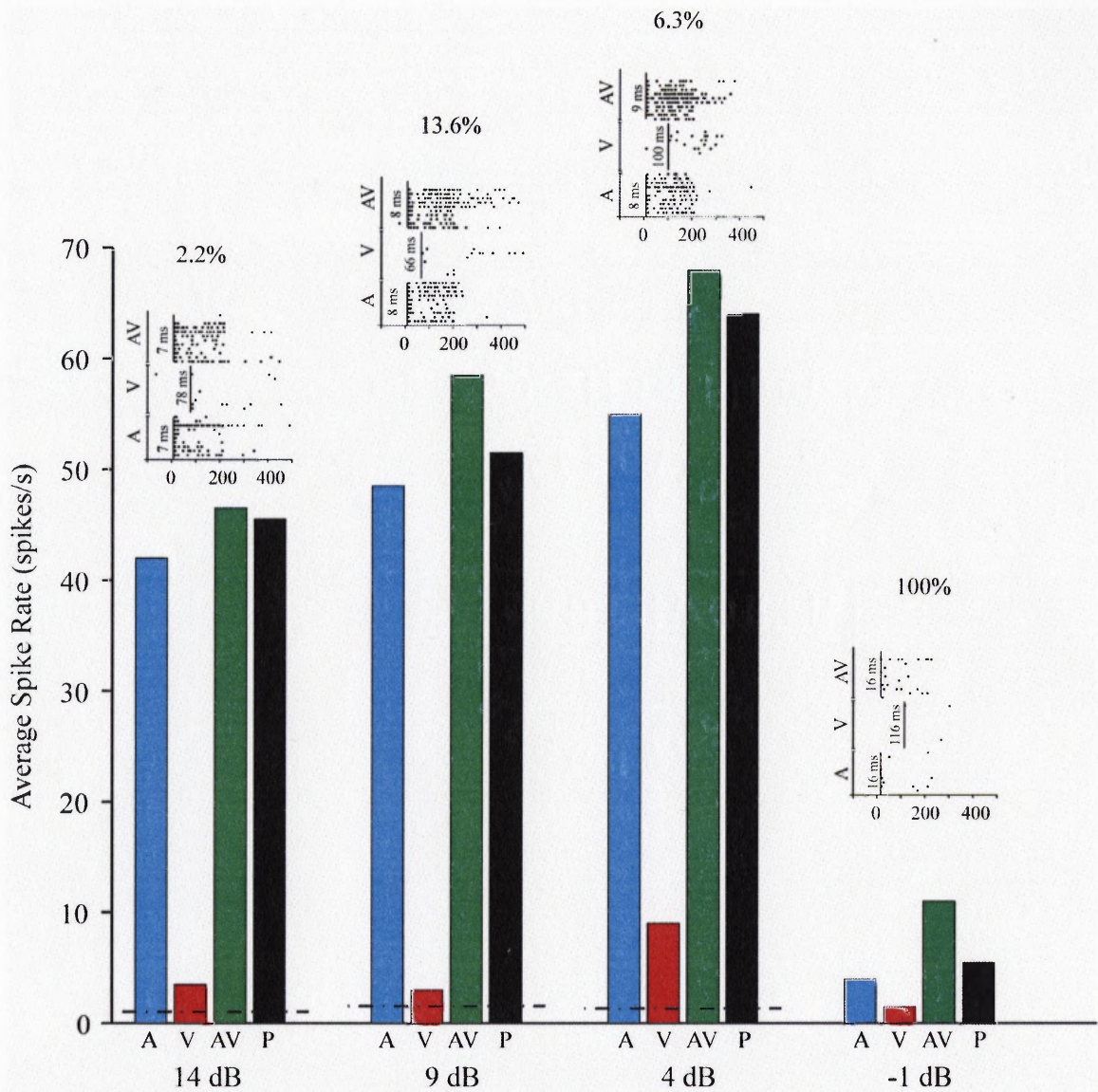


Figure 5-2. OT neurons follow the inverse effectiveness principle.

Audiovisual responses (AV – green) were generally greater than the auditory (A – blue), visual (V – red), and predicted audiovisual (P – black) responses. The data is grouped according to the auditory SPL, noted below the abscissa. The bars represent the average spike rate shown in the insets above. First-spike latencies are shown by the vertical line in the rasters, and the absolute value is reported in each condition at the left of the vertical line. Response enhancement is noted above each SPL group. The greatest enhancement was seen in the trials with the lowest SPL. Spontaneous rate for each group is shown by the dot-dashed line across the bars.

Thus this cell responded in a manner much like what has been seen in mammalian studies: Audiovisual responses were most enhanced when the unisensory components were relatively *ineffective at eliciting a robust response on their own*.

Of note was the observation that the overall responses to auditory and visual stimuli changed over time. Those data shown in Figure 5-2 illustrate how average responses to auditory stimuli actually increased as SPL decreased. I noticed that during long experiments, the birds would often go through periods of time in which they were more aroused, as seen by the positions of the facial feathers and the amount of overall movement. Blink rates also changed. I therefore designed an exceedingly long experiment to test the stability of a neuronal response over time: I measured responses to auditory, visual, and audiovisual stimuli presented at exactly the same SPL and visual intensity, and exactly the same location with an inter-stimulus-interval of 2 seconds. Each stimulus was 200 ms in duration, and repeated 20 times.

Figure 5-3 shows the raster plots from this experiment. In the audiovisual trials, the visual modality led the auditory modality by 80 ms, the approximate difference between the auditory and visual first-spike latencies. This temporal misalignment is marked in Figure 5-3A (Visual onset, dot-dashed vertical line; Auditory onset, solid line). The duration of the stimuli are shown by the grey bar. To show a general consistency of the background spontaneous rate, a blank trial was inserted after every audiovisual trial for every repetition. Thus the order of stimuli presented in each repetition was auditory, visual, audiovisual, and finally blank.

As suspected, the responses across time changed dramatically (Figure 5-3, repetitions 1-20). Trials recorded in the first half of the experiment (repetitions 1-10) showed a much greater response to all three stimulus types than trials recorded in the second half. Interestingly, the spontaneous rate did not fluctuate (Figure 5-3D). This suggests that it wasn't the overall activity of the cell that increased, but rather the activity in response to the stimuli.

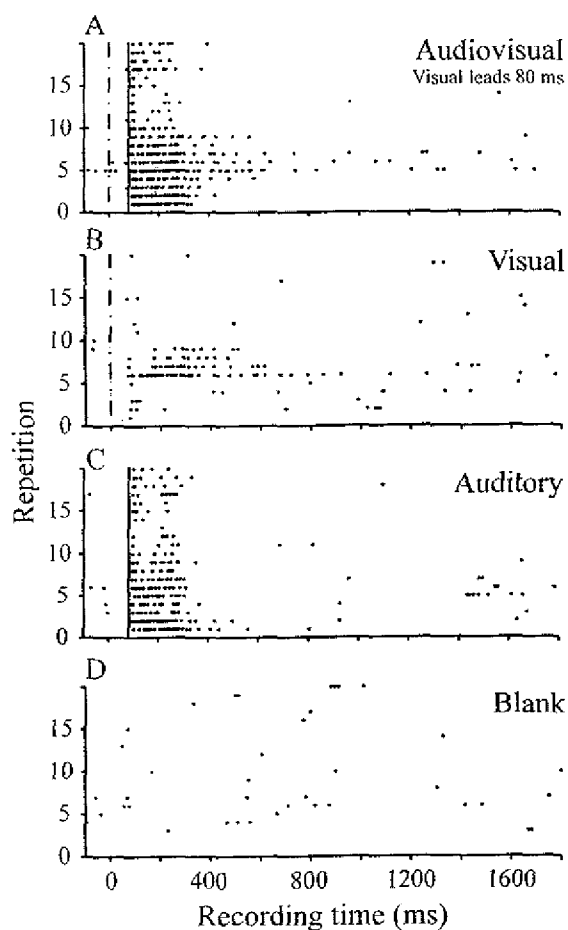


Figure 5-3. OT responses are variable in time.

The data are reported in the order they were collected. Repetition 1 was collected first and repetition 20 was collected last. The modalities were presented in a sequential order: auditory, visual, audiovisual, blank. As the test progressed, the responses to each stimulus type fluctuated. Grey bars show the duration of the stimuli and the vertical lines show the stimulus onset point for the auditory (solid) and visual (dot-dashed) modalities. Visual stimuli led the auditory stimuli by 80 ms in audiovisual trials.

In the temporally offset audiovisual condition illustrated in Figure 5-3, audiovisual responses were generally greater than the sum of the unisensory responses, however I did not see enhancement greater than what was observed for simultaneous audiovisual stimuli in this cell (Figure 5-2). More data from more cells are needed to probe further the importance of *temporal alignment in eliciting an enhanced response in the OT cells of the barn owl.*

Electrophysiological responses in single units from the anesthetized barn owl's optic tectum were not unlike those reported for mammalian subjects. Audiovisual spike rates were *sometimes greater than the predicted sum of the auditory and visual component and sometimes not.* Because the data were so variable across cells and even within a single cell, I did not attempt to report population statistics. Future experiments will be designed to keep the animal's state of arousal at a constant level to avoid such complications.

Ultimately, the population response across the OT must be correlated with behavior, instead of single neuron responses to a single stimulus in space. Previous estimates of the population response to auditory stimuli have been obtained by assuming that a single unit's response to multiple stimulus locations was analogous to activity across a population in response to a single stimulus in space (Keller and Takahashi, 2005). This method is viable for predictive purposes, however the question may only truly be addressed by recording from more than one cell at a time during behavior. Future electrophysiological data *simultaneously gathered from multiple units in behaving barn owls are required to obtain more direct correlations between behavior and the neural computations underlying them.*

CHAPTER VI

CONCLUSIONS

It is truly amazing that the barn owl is able to localize prey in the absence of any visual cues, using their auditory sense alone. My measurements of behavior in auditory alone tasks underline this ability. The visual sense for these birds however, is far from vestigial. With eyes adapted for acuity in low-intensity environments, the owl is extremely accurate and precise in generating a saccade to even the faintest light source. The combination of these senses, especially in noisy environments where signal-to-noise ratios are low, results in behavior that is both quick and accurate, as described in Chapter II. Furthermore, these birds are conceivably able to update their sensory representations during extraordinarily fast movements (peak head speeds in saccades were as fast as $600^\circ/\text{s}$) and convert newly arriving information into a smooth behavioral adjustment, optimizing the chances of obtaining a reward (Chapter III).

The neural computations underlying these observations of behavior remain essentially unknown. However, I found that neuronal responses to quiet stimuli were less prevalent across cells, delayed in time, and noticeably different across space relative to those responses obtained in louder trials (Chapter IV). I also found that cells responding to both auditory and visual stimuli in the OT were sometimes excited beyond what would be expected were the owl to simply add the auditory and visual inputs, much like what has previously been seen in mammalian studies. Future experiments probing the connection between these behavioral and

neurophysiological observations will help to elucidate fundamental components of both sensory-motor and audiovisual integration, and ultimately move us closer to the goal of understanding the neural basis of behavior.

BIBLIOGRAPHY

- Alais D, Burr D (2004) The ventriloquist effect results from near-optimal bimodal integration. *Current Biology* 14:257-262.
- Arai K, Keller EL (2005) A model of the saccade-generating system that accounts for trajectory variations produced by competing visual stimuli. *Biological Cybernetics* 92:21-37.
- Arndt PA, Colonius H (2003) Two stages in crossmodal saccadic integration: evidence from a visual-auditory focused attention task. *Experimental Brain Research* 150:417-426.
- Bala AD, Spitzer MW, Takahashi TT (2003) Prediction of auditory spatial acuity from neural images on the owl's auditory space map. *Nature* 424:771-774.
- Bastian J (1982) Vision and electroreception: integration of sensory information in the optic tectum of the weakly electric fish *Apteronotus albifrons*. *Journal of Comparative Physiology A, Neuroethology, Sensory, Neural, and Behavioral Physiology* 147:287-297.
- Battaglia PW, Jacobs RA, Aslin RN (2003) Bayesian integration of visual and auditory signals for spatial localization. *Journal of the Optical Society of America A-Optics Image Science and Vision* 20:1391-1397.
- Bell AH, Corneil BD, Munoz DP, Meredith MA (2003) Engagement of visual fixation suppresses sensory responsiveness and multisensory integration in the primate superior colliculus. *European Journal of Neuroscience* 18:2867-2873.
- Bell AH, Meredith MA, Van Opstal AJ, Munoz DP (2005) Crossmodal integration in the primate superior colliculus underlying the preparation and initiation of saccadic eye movements. *Journal of Neurophysiology* 93:3659-3673.
- Bergan JF, Ro P, Ro D, Knudsen EI (2005) Hunting increases adaptive auditory map plasticity in adult barn owls. *Journal of Neuroscience* 25:9816-9820.

- Bertelson P, De Gelder B (2004) The psychology of multimodal perception. In: Crossmodal space and crossmodal attention (Spence C, Driver J, eds), pp 141-177. New York: Oxford University Press.
- Bertelson P, Radeau M (1981) Cross-modal bias and perceptual fusion with auditory-visual spatial discordance. *Perception & Psychophysics* 29:578-584.
- Brainard MS, Knudsen EI (1998) Sensitive periods for visual calibration of the auditory space map in the barn owl optic tectum. *Journal of Neuroscience* 18:3929-3942.
- Burnett LR, Stein BE, Chaponis D, Wallace MT (2004) Superior colliculus lesions preferentially disrupt multisensory orientation. *Neuroscience* 124:535-547.
- Calvert GA, Spence C, Stein BE (2004) *The handbook of multisensory processes*. Cambridge, Massachusetts: The MIT Press.
- Carpenter RHS, Williams MLL (1995) Neural computation of log likelihood in control of saccadic eye-movements. *Nature* 377:59-62.
- Carr CE, Konishi M (1990) A circuit for detection of interaural time differences in the brain stem of the barn owl. *Journal of Neuroscience* 10:3227-3246.
- Colonus H, Arndt P (2001) A two-stage model for visual-auditory interaction in saccadic latencies. *Perception & Psychophysics* 63:126-147.
- Colonus H, Diederich A (2004) Multisensory interaction in saccadic reaction time: a time-window-of-integration model. *Journal of Cognitive Neuroscience* 16:1000-1009.
- Cornel BD, Munoz DP (1996) The influence of auditory and visual distractors on human orienting gaze shifts. *Journal of Neuroscience* 16:8193-8207.
- Cornel BD, Van Wanrooij M, Munoz DP, Van Opstal AJ (2002) Auditory-visual interactions subserving goal-directed saccades in a complex scene. *Journal of Neurophysiology* 88:438-454.
- Diederich A, Colonius H (2004) Bimodal and trimodal multisensory enhancement: effects of stimulus onset and intensity on reaction time. *Perception & Psychophysics* 66:1388-1404.
- Diederich A, Colonius H, Bockhorst D, Tabeling S (2003) Visual-tactile spatial interaction in saccade generation. *Experimental Brain Research* 148:328-337.

- du Lac S, Knudsen EI (1990) Neural maps of head movement vector and speed in the optic tectum of the barn owl. *Journal of Neurophysiology* 63:131-146.
- Engelken EJ, Stevens KW (1989) Saccadic eye-movements in response to visual, auditory, and bisensory stimuli. *Aviation Space and Environmental Medicine* 60:762-768.
- Ernst MO, Banks MS (2002) Humans integrate visual and haptic information in a statistically optimal fashion. *Nature* 415:429-433.
- Euston DR, Takahashi TT (2002) From spectrum to space: the contribution of level difference cues to spatial receptive fields in the barn owl inferior colliculus. *Journal of Neuroscience* 22:284-293.
- Feldman DE, Knudsen EI (1997) An anatomical basis for visual calibration of the auditory space map in the barn owl's midbrain. *Journal of Neuroscience* 17:6820-6837.
- Frens MA, Van Opstal AJ (1998) Visual-auditory interactions modulate saccade-related activity in monkey superior colliculus. *Brain Research Bulletin* 46:211-224.
- Frens MA, Van Opstal AJ, Van der Willigen RF (1995) Spatial and temporal factors determine auditory-visual interactions in human saccadic eye movements. *Perception & Psychophysics* 57:802-816.
- Glimcher PW (2003) The neurobiology of visual-saccadic decision making. *Annual Review of Neuroscience* 26:133-179.
- Grace MS, Woodward OM, Church DR, Calisch G (2001) Prey targeting by the infrared-imaging snake *Python molurus*: effects of experimental and congenital visual deprivation. *Behavioral Brain Research* 119:23-31.
- Grice GR, Canham L, Boroughs JM (1984) Combination rule for redundant information in reaction-time tasks with divided attention. *Perception & Psychophysics* 35:451-463.
- Gutfreund Y, Zheng W, Knudsen EI (2002) Gated visual input to the central auditory system. *Science* 297:1556-1559.
- Hairston WD, Laurienti PJ, Mishra G, Burdette JH, Wallace MT (2003) Multisensory enhancement of localization under conditions of induced myopia. *Experimental Brain Research* 152:404-408.

- Harrington LK, Peck CK (1998) Spatial disparity affects visual-auditory interactions in human sensorimotor processing. *Experimental Brain Research* 122:247-252.
- Hartline PH, Kass L, Loop MS (1978) Merging of modalities in the optic tectum: infrared and visual integration in rattlesnakes. *Science* 199:1225-1229.
- Heil P (2004) First-spike latency of auditory neurons revisited. *Current Opinion in Neurobiology* 14:461-467.
- Heil P, Irvine DRF (1997) First-spike timing of auditory-nerve fibers and comparison with auditory cortex. *Journal of Neurophysiology* 78:2438-2454.
- Heron J, Whitaker D, McGraw F (2004) Sensory uncertainty governs the extent of audio-visual interaction. *Vision Research* 44:2875-2884.
- Honda H (2005) The remote distractor effect of saccade latencies in fixation-offset and overlap conditions. *Vision Research* 45:2773-2779.
- Hughes HC, Nelson MD, Aronchick DM (1998) Spatial characteristics of visual-auditory summation in human saccades. *Vision Research* 38:3955-3963.
- Hughes HC, Reuter-Lorenz PA, Nozawa G, Fendrich R (1994) Visual-auditory interactions in sensorimotor processing: saccades versus manual responses. *Journal of Experimental Psychology Human Perception and Performance* 20:131-153.
- Hyde PS, Knudsen EI (2001) A topographic instructive signal guides the adjustment of the auditory space map in the optic tectum. *Journal of Neuroscience* 21:8586-8593.
- Hyde PS, Knudsen EI (2002) The optic tectum controls visually guided adaptive plasticity in the owl's auditory space map. *Nature* 415:73-76.
- Institute of Laboratory Animal Resources (U.S.), NetLibrary Inc. (1997) Guide for the care and use of laboratory animals, Ed 7. Washington, D.C.: National Academy Press.
- Jay MF, Sparks DL (1987) Sensorimotor integration in the primate superior colliculus. II. Coordinates of auditory signals. *Journal of Neurophysiology* 57:35-55.
- Jiang W, Jiang H, Stein BE (2002) Two corticotectal areas facilitate multisensory orientation behavior. *Journal of Cognitive Neuroscience* 14:1240-1255.

- Keller CH, Hartung K, Takahashi TT (1998) Head-related transfer functions of the barn owl: measurement and neural responses. *Hearing Research* 118:13-34.
- Keller CH, Takahashi TT (1996) Responses to simulated echoes by neurons in the barn owl's auditory space map. *Journal of Comparative Physiology A, Neuroethology, Sensory, Neural, and Behavioral Physiology* 178:499-512.
- Keller CH, Takahashi TT (2005) Localization and identification of concurrent sounds in the owl's auditory space map. *Journal of Neuroscience* 25:10446-10461.
- King AJ, Palmer AR (1985) Integration of visual and auditory information in bimodal neurones in the guinea-pig superior colliculus. *Experimental Brain Research* 60:492-500.
- Kirchner H, Colonius H (2004) Predictiveness of a visual distractor modulates saccadic responses to auditory targets. *Experimental Brain Research* 155:257-260.
- Knudsen EI (1982) Auditory and visual maps of space in the optic tectum of the owl. *Journal of Neuroscience* 2:1177-1194.
- Knudsen EI (1999) Mechanisms of experience-dependent plasticity in the auditory localization pathway of the barn owl. *Journal of Comparative Physiology A, Neuroethology, Sensory, Neural, and Behavioral Physiology* 185:305-321.
- Knudsen EI, Blasdel GG, Konishi M (1979) Sound localization by the barn owl (*Tyto alba*) measured with the search coil technique. *Journal of Comparative Physiology A, Neuroethology, Sensory, Neural, and Behavioral Physiology* 133:1-11.
- Knudsen EI, Brainard MS (1995) Creating a unified representation of visual and auditory space in the brain. *Annual Review of Neuroscience* 18:19-43.
- Knudsen EI, Knudsen PF (1983) Space-mapped auditory projections from the inferior colliculus to the optic tectum in the barn owl (*Tyto alba*). *Journal of Comparative Neurology* 218:187-196.
- Knudsen EI, Knudsen PF (1989) Vision calibrates sound localization in developing barn owls. *Journal of Neuroscience* 9:3306-3313.
- Knudsen EI, Knudsen PF, Esterly SD (1982) Early auditory experience modifies sound localization in barn owls. *Nature* 295:238-240.

- Knudsen EI, Knudsen PF, Masino T (1993) Parallel pathways mediating both sound localization and gaze control in the forebrain and midbrain of the barn owl. *Journal of Neuroscience* 13:2837-2852.
- Knudsen EI, Konishi M (1978) A neural map of auditory space in the owl. *Science* 200:795-797.
- Knudsen EI, Konishi M (1979) Mechanisms of sound localization in the barn owl (*Tyto alba*). *Journal of Comparative Physiology A, Neuroethology, Sensory, Neural, and Behavioral Physiology* 133:13-21.
- Konishi M (1973) How the owl tracks its prey. *American Scientist* 61:414-424.
- Koppl C, Yates G (1999) Coding of sound pressure level in the barn owl's auditory nerve. *Journal of Neuroscience* 19:9674-9686.
- Litovsky RY, Colburn HS, Yost WA, Guzman SJ (1999) The precedence effect. *Journal of the Acoustical Society of America* 106:1633-1654.
- Luksch H (2003) Cytoarchitecture of the avian optic tectum: neuronal substrate for cellular computation. *Reviews in the Neurosciences* 14:85-106.
- Ma W-J, Beck J, Pe L, Pouget A (in press) Bayesian inference with probabilistic population codes. *Nature Neuroscience*.
- Masino T, Knudsen EI (1992) Anatomical pathways from the optic tectum to the spinal cord subserving orienting movements in the barn owl. *Experimental Brain Research* 92:194-208.
- Masino T, Knudsen EI (1993) Orienting head movements resulting from electrical microstimulation of the brainstem tegmentum in the barn owl. *Journal of Neuroscience* 13:351-370.
- McSorley E, Haggard P, Walker R (2005) Spatial and temporal aspects of oculomotor inhibition as revealed by saccade trajectories. *Vision Research* 45:2492-2499.
- Meredith MA, Nemitz JW, Stein BE (1987) Determinants of multisensory integration in superior colliculus neurons. I. Temporal factors. *Journal of Neuroscience* 7:3215-3229.

- Meredith MA, Stein BE (1986a) Visual, auditory, and somatosensory convergence on cells in superior colliculus results in multisensory integration. *Journal of Neurophysiology* 56:640-662.
- Meredith MA, Stein BE (1986b) Spatial factors determine the activity of multisensory neurons in cat superior colliculus. *Brain Research* 365:350-354.
- Meredith MA, Stein BE (1996) Spatial determinants of multisensory integration in cat superior colliculus neurons. *Journal of Neurophysiology* 75:1843-1857.
- Miller JO (1982) Divided attention: evidence for coactivation with redundant signals. *Cognitive Psychology* 14:247-279.
- Moiseff A, Konishi M (1981) Neuronal and behavioral sensitivity to binaural time differences in the owl. *Journal of Neuroscience* 1:40-48.
- Moiseff A, Konishi M (1983) Binaural characteristics of units in the owls brain-stem auditory pathway - Precursors of restricted spatial receptive-fields. *Journal of Neuroscience* 3:2553-2562.
- Moller P (2002) Multimodal sensory integration in weakly electric fish: a behavioral account. *Journal of Physiology-Paris* 96:547-556.
- Mordkoff JT, Yantis S (1991) An interactive race model of divided attention. *Journal of Experimental Psychology Human Perception and Performance* 17:520-538.
- Munoz DP, Dorris MC, Pare M, Everling S (2000) On your mark, get set: brainstem circuitry underlying saccadic initiation. *Canadian Journal of Physiology and Pharmacology* 78:934-944.
- Munoz DP, Istvan PJ (1998) Lateral inhibitory interactions in the intermediate layers of the monkey superior colliculus. *Journal of Neurophysiology* 79:1193-1209.
- Nemec P, Altmann J, Marhold S, Burda H, Oelschläger HH (2001) Neuroanatomy of magnetoreception: the superior colliculus involved in magnetic orientation in a mammal. *Science* 294:366-368.
- Nozawa G, Reuter-Lorenz PA, Hughes HC (1994) Parallel and serial processes in the human oculomotor system: bimodal integration and express saccades. *Biological Cybernetics* 72:19-34.

- Olivier E, Dorris MC, Munoz DP (1999) Lateral interactions in the superior colliculus, not an extended fixation zone, can account for the remote distracter effect. *Behavioral and Brain Sciences* 22:694-695.
- Papoulis A (1991) *Probability, random variables, and stochastic process*. New York: McGraw Hill.
- Payne RS (1971) Acoustic location of prey by barn owls (*Tyto alba*). *Journal of Experimental Biology* 54:535-573.
- Payne RS, Drury W (1958) Marksman of the darkness. *Natural History* 67:316-323.
- Peck CK, Baro JA, Warder SM (1995) Effects of eye position on saccadic eye movements and on the neuronal responses to auditory and visual stimuli in cat superior colliculus. *Experimental Brain Research* 103:227-242.
- Poganiatz I, Nelken I, Wagner H (2001) Sound-localization experiments with barn owls in virtual space: influence of interaural time difference on head-turning behavior. *Journal of the Association for Research in Otolaryngology* 2:1-21.
- Populin LC (2005) Anesthetics change the excitation/inhibition balance that governs sensory processing in the cat superior colliculus. *Journal of Neuroscience* 25:5903-5914.
- Populin LC, Tollin DJ, Yin TCT (2004) Effect of eye position on saccades and neuronal responses to acoustic stimuli in the superior colliculus of the behaving cat. *Journal of Neurophysiology* 92:2151-2167.
- Populin LC, Yin TC (2002) Bimodal interactions in the superior colliculus of the behaving cat. *Journal of Neuroscience* 22:2826-2834.
- Raab DH (1962) Statistical facilitation of simple reaction times. *Transactions of the New York Academy of Sciences* 24:574-590.
- Radeau M, Bertelson P (1987) Auditory-visual interaction and the timing of inputs - Thomas (1941) revisited. *Psychological Research-Psychologische Forschung* 49:17-22.
- Ramon y Cajal S (1995) The optic lobe of lower vertebrates. In: *Histology of the nervous system*, pp 161-173. Oxford: Oxford University Press.

- Robinson DA (1972) Eye movements evoked by collicular stimulation in the alert monkey. *Vision Research* 12:1795-1808.
- Schwarz W (1996) Further tests of the interactive race model of divided attention: the effects of negative bias and varying stimulus - onset asynchronies. *Psychological Research-Psychologische Forschung* 58:233-245.
- Sparks DL (2002) The brainstem control of saccadic eye movements. *Nature Reviews Neuroscience* 3:952-964.
- Spence C, Driver J (2004) *Crossmodal space and crossmodal attention*. New York: Oxford University Press.
- Spezio ML, Takahashi TT (2003) Frequency-specific interaural level difference tuning predicts spatial response patterns of space-specific neurons in the barn owl inferior colliculus. *Journal of Neuroscience* 23:4677-4688.
- Spitzer MW, Bala AD, Takahashi TT (2003) Auditory spatial discrimination by barn owls in simulated echoic conditions. *Journal of the Acoustical Society of America* 113:1631-1645.
- Spitzer MW, Takahashi TT (2006) Sound localization by barn owls in a simulated echoic environment. *Journal of Neurophysiology* 95:3571-3584.
- Stanford TR, Quessy S, Stein BE (2005) Evaluating the operations underlying multisensory integration in the cat superior colliculus. *Journal of Neuroscience* 25:6499-6508.
- Stein BE, Gaither NS (1981) Sensory representation in reptilian optic tectum: some comparisons with mammals. *Journal of Comparative Neurology* 202:69-87.
- Stein BE, Huneycutt WS, Meredith MA (1988) Neurons and behavior: the same rules of multisensory integration apply. *Brain Research* 448:355-358.
- Stein BE, Meredith MA, eds (1993) *The merging of the senses*. Cambridge, MA: MIT Press.
- Stein BE, Meredith MA, Huneycutt WS, McDade L (1989) Behavioral indices of multisensory integration: orientation to visual cues is affected by auditory stimuli. *Journal of Cognitive Neuroscience* 1:12-24.
- Takahashi TT (1989) Construction of an auditory space map. *Annals of the New York Academy of Sciences* 563:101-113.

- Takahashi TT, Bala ADS, Spitzer MW, Euston DR, Spezio ML, Keller CH (2003) The synthesis and use of the owl's auditory space map. *Biological Cybernetics* 89:378-387.
- Todd JW (1912) Reaction to multiple stimuli. In: *Archives of psychology*, No.25 (Woodworth RS, ed), pp 1-65. New York: The Science Press.
- Townsend JT, Wenger MJ (2004) A theory of interactive parallel processing: new capacity measures and predictions for a response time inequality series. *Psychological Review* 111:1003-1035.
- Valentine DE, Sinha SR, Moss CF (2002) Orienting responses and vocalizations produced by microstimulation in the superior colliculus of the echolocating bat, *Eptesicus fuscus*. *Journal of Comparative Physiology A, Neuroethology, Sensory, Neural, and Behavioral Physiology* 188:89-108.
- Van Opstal AJ, Munoz DP (2004) Auditory-visual interactions subserving primate gaze orienting. In: *The handbook of multisensory processes* (Calvert GA, Spence C, Stein BE, eds), pp 373-393. Cambridge, MA: The MIT Press.
- Vanegas H, ed (1984) *Comparative neurology of the optic tectum*. New York: Plenum Press.
- Vliegen J, Van Grootel TJ, Van Opstal AJ (2004) Dynamic sound localization during rapid eye-head gaze shifts. *Journal of Neuroscience* 24:9291-9302.
- Wagner H (1993) Sound-localization deficits induced by lesions in the barn owls auditory space map. *Journal of Neuroscience* 13:371-386.
- Wallace MT, Meredith MA, Stein BE (1998) Multisensory integration in the superior colliculus of the alert cat. *Journal of Neurophysiology* 80:1006-1010.
- Wallace MT, Wilkinson LK, Stein BE (1996) Representation and integration of multiple sensory inputs in primate superior colliculus. *Journal of Neurophysiology* 76:1246-1266.
- Walls GL (1972) Eagle owl eye. In: *The vertebrate eye and its adaptive radiation*, p 785. New York: Hafner Publishing Company.
- Welch R, Warren D (1986) Intersensory interactions. In: *Handbook of perception and human performance* (Boff KR, Kaufman L, Thomas JP, eds), pp 21-36. New York: Wiley.

- Whitchurch EA, Takahashi TT (2006) Combined auditory and visual stimuli facilitate head saccades in the barn owl (*Tyto alba*). *Journal of Neurophysiology* 96:730-745.
- Wilkinson LK, Meredith MA, Stein BE (1996) The role of anterior ectosylvian cortex in cross-modality orientation and approach behavior. *Experimental Brain Research* 112:1-10.
- Zheng W, Knudsen EI (2001) Gabaergic inhibition antagonizes adaptive adjustment of the owl's auditory space map during the initial phase of plasticity. *Journal of Neuroscience* 21:4356-4365.