

Temporal weighting of binaural cues for sound localization

Andrew D. Brown

A dissertation
submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

University of Washington
2012

Reading Committee:
George C. Stecker, Chair
Julie A. Bierer
Joseph A. Sisneros
Lynne A. Werner

Authorized to Offer Degree:
Speech and Hearing Sciences

University of Washington

Abstract

Temporal weighting of binaural cues for sound localization

Andrew D. Brown

Chair of the Supervisory Committee:
Associate Professor G. Christopher Stecker

Normal-hearing human listeners localize sound sources accurately and effortlessly across an effectively limitless variety of acoustic environments by responding to, among other cues, small differences in the timing and intensity of sound arriving at the two ears (interaural time differences [ITD] and interaural level differences [ILD]). This dissertation explores the temporal dynamics of normal sensitivity to these cues by assessing normal-hearing listeners' weighting of ITD or ILD cues over the duration of binaurally dynamic stimuli in a variety of psychophysical experiments. Experimental stimuli were rapidly amplitude-modulated signals mimicking those typical of echoic environments (e.g., rooms), including rapid sequences of filtered impulses carrying varied ITD or ILD over their duration, and pairs and trains of pairs of clicks carrying discrete ITD or ILD in first and second click of each pair. Consistent with past studies of the precedence effect in sound localization, onset dominance in lateralization and binaural adaptation, data demonstrated a prime salience of ILD and especially ITD at signal onset. Somewhat more surprisingly, data also indicated greater sensitivity to post-onset ILD than post-onset ITD (i.e., a weaker precedence effect for ILD). These psychophysical data were qualitatively consistent with outputs of a simple model that included peripheral auditory filtering effects, auditory nerve adaptation, and parallel computation of ITD and ILD. These investigations collectively suggest that accurate sound localization in ordinary listening environments depends on acute ITD and ILD sensitivity at signal onset, followed (on scale of a few to tens of milliseconds) by post-onset attenuation of ITD sensitivity and dependence on ILD for detection of changes in the environment (e.g., introduction of a novel source).

Table of Contents

List of Figures	iv
-----------------------	----

Chapter 1: Sound localization in the horizontal plane: An introduction

1.1. Binaural cues for sound localization	2
1.2. ITD carried by amplitude envelopes	4
1.3. Sensitivity to ILD across the spectrum	5
1.4. The importance of signal onsets.....	6
1.5. Clinical considerations	7

Chapter 2: Temporal weighting of interaural time and level differences carried by high-rate trains of filtered impulses

2.1. Previous temporal weighting studies	10
2.2. Methods	12
2.2.1. Subjects	12
2.2.2. Stimuli and procedure	13
2.2.3. Statistical methods.....	14
2.3. Results	16
2.3.1. Temporal weighting functions for ITD and ILD	16
2.3.2. Reduced onset dominance for ILD	20
2.4. Discussion.....	22
2.4.1. Onset dominance for ITD and ILD	22
2.4.2. Cue-averaging versus sensitivity to peak interaural value.....	24
2.4.3. ROC analysis for TWFs	25
2.4.4. Cues in agreement versus cues in conflict	27
2.4.5. Possible effects of the task employed	27
2.4.6. The range of interaural variation	28
2.5. Summary and conclusions	29

Chapter 3: Temporal weighting of interaural time and level differences. II. The effect of binaurally synchronous temporal jitter

3.1. Effects of temporal irregularity on binaural sensitivity	31
3.2. Methods	35
3.2.1. Subjects	35
3.2.2. Stimuli and procedure	35
3.2.3. Statistical methods.....	37
3.3. Results	38
3.3.1. Temporal weighting functions for ITD and ILD	38
3.3.2. Longest versus shortest gaps.....	41
3.4. Discussion.....	42
3.4.1. The effect of jitter on TWFs for ITD and ILD	42
3.4.2. Greater cue-averaging for ILD than ITD	43
3.4.3. “Restarting” versus listening after the longest gaps.....	44
3.5. Summary and conclusions	47

Chapter 4: The precedence effect in sound localization: Fusion and lateralization measures for pairs and trains of lead-lag stimuli lateralized by interaural time and level differences

4.1. Different contributions of ITD and ILD to precedence	49
4.2. Common experimental methods	51
4.2.1. Subjects	51
4.2.2. Procedure	51
4.2.2.1. Training	52
4.2.2.2. ITD-ILD matching task	53
4.2.2.3. Main experimental task	54
4.2.3. Analysis	56
4.3. Experiment I: Dynamic precedence effects for ITD and ILD.....	57
4.3.1. Results	58
4.3.1.1. Echo thresholds	58
4.3.1.2. Lateralization responses	59
4.3.2. Interim discussion	62
4.4. Experiment II: Cross-cue transfer of buildup.....	63
4.4.1. Results and interim discussion	64
4.5. Experiment III: Dynamic precedence effects within a single hemifield.....	66
4.5.1. Stimuli	68
4.5.2. Results and interim discussion	68
4.5.2.1. Echo thresholds	68
4.5.2.2. Lateralization responses	70
4.6. Summary and general discussion	72

Chapter 5: Consequences of peripheral auditory processing on the temporal dynamics of ITD and ILD sensitivity: A dual-display model

5.1. Existing binaural models	76
5.1.1. Peripheral components.....	76
5.1.2. Jeffress-Colburn models.....	78
5.1.3. Models based on binaural “cancellation”	80
5.2. A dual-display binaural model.....	81
5.2.1. Auditory nerve model: Zilany et al., 2009	81
5.2.2. ITD display: Frequency-weighted interaural cross-correlation.....	83
5.2.3. ILD display: Count-comparison of left- and right-ear PSTHs	85
5.3. Model demonstrations	88
5.3.1. Gabor click trains	88
5.3.2. Lead-lag click pairs: Baseline and Buildup conditions	91
5.3.2.1. Baseline ITD stimuli	91
5.3.2.2. Baseline ILD stimuli	92
5.3.2.3. Buildup ITD stimuli	94
5.3.3. Binaural acoustic recordings.....	96

Chapter 6: Concluding remarks

97

Appendix A: Left-lead, right-lag lateralization data	101
Appendix B: Model implementation details	104
Bibliography	105

List of Figures

Chapter 1

1.1. Binaural cues for sound localization	3
1.2. Envelope ITD	4
1.3. ILD discrimination as a function of frequency	5
1.4. Recordings of balloon pops	6

Chapter 2

2.1. Schematic illustration of stimuli	13
2.2. Averaged temporal weighting functions	17
2.3. Individual temporal weighting functions	18
2.4. Onset dominance for ITD and ILD	21
2.5. Weight estimation by ROC analysis	25

Chapter 3

3.1. Schematic illustration of stimuli	36
3.2. Averaged temporal weighting functions	39
3.3. Individual temporal weighting functions	40
3.4. Onset dominance for Jittered and Isochronous stimuli	44

Chapter 4

4.1. Illustration of 120 μ s impulse	52
4.2. Main experimental task	55
4.3. Schematic illustration of Experiment I stimuli	58
4.4. Mean echo thresholds, Experiment I	59
4.5. Individual echo thresholds, Experiment I	59
4.6. Lateralization data, Experiment I	60
4.7. ITD-ILD matching data, Experiment II	63
4.8. Mean echo thresholds, Experiment II	64
4.9. Individual echo thresholds, Experiment II	65
4.10. Lateralization data, Experiment II	66
4.11. ITD-ILD matching data, Experiment III	68
4.12. Mean echo thresholds, Experiment III	69
4.13. Individual echo thresholds, Experiment III	69
4.14. Lateralization data, Experiment III	71

Chapter 5

5.1. PSTHs for an approximate impulse	77
5.2. Zilany model output for click with ITD	82
5.3. Zilany model output for click with ILD	83
5.4. ITD display	84
5.5. ILD display	86
5.6. Summary of complete model	87
5.7. Model outputs for Chapter 2 and 3 stimuli	90

5.8. ITD display for Baseline ITD stimuli	92
5.9. ILD display for Baseline ILD stimuli	93
5.10. Baseline ITD versus Baseline ILD	94
5.11. Baseline ITD versus Buildup ITD	95
5.12. Model output for balloon pop	96

Acknowledgements

The author thanks the members of his doctoral committee and especially his doctoral advisor, Chris Stecker, for professional guidance and wisdom in things academic, members of his lab past and present for their willing assistance and humor, his many mentors, professors, and teachers at various institutions over many years, the funding agencies and scholarship donors that have financially supported his work and education, his friends and family for their support and encouragement, and most especially, his mother and father.

To my parents.

Sound localization in the horizontal plane: An introduction

The ability of organisms to perceive and orient to biologically relevant stimuli is basic to their survival. *Spatial orientation* facilitates predator avoidance, prey capture and, in humans and many other complex organisms, communication (for an excellent review, see Schone, 1986). Spatial orientation requires that information about the location or directionality of external stimuli be extracted and encoded internally, in the central nervous system. Despite the morphological diversity of sensory systems, both across sensory modalities within single species and across species within single sensory modalities, most all sensory systems facilitate spatial orientation (Schone, 1986). Such systems are in operation when a migrating salmon traces a chemical gradient back to its natal stream, when an owl swoops down to capture a rustling field mouse, or when a party guest shouts across a noisy and crowded room to a friend, who turns to face him without hesitation.

This dissertation is concerned with spatial orientation in the auditory domain. While the fundamental mechanisms of audition (i.e., auditory transduction and peripheral neural encoding of acoustic features such as frequency and intensity) are largely preserved across vertebrates, substantial variability exists across species in more central aspects of audition (see Pickles, 2008). Nonetheless, most vertebrates (even water-bound species, including many fish) possess the ability to locate and respond adaptively to a wide variety of biologically relevant sound sources. Among studied species, humans exhibit an extraordinarily high degree of sensitivity to auditory spatial information¹. For sound sources directly or nearly directly ahead, normal-hearing adult listeners are able to discriminate differences in sound source location as small as 1° (Mills, 1958). That this feat may be achieved in the presence of competing sound sources including background noise, signal reflections and diffuse reverberation (Wallach *et al.*, 1949; Cherry, 1953) is remarkable, particularly when one

considers that the complete ensemble of acoustic information is condensed to a one-dimensional deflection of each tympanum prior to auditory transduction.

Assuredly listeners must exploit a variety of acoustic cues for robust localization across the effectively limitless variety of acoustic environments they may encounter, and assuredly a variety of physiologic mechanisms must subserve sensitivity to these cues. Toward an improved understanding of the foregoing issues, this dissertation examines the sensitivity of normal-hearing adult listeners to acoustic cues for sound localization in the horizontal plane (i.e., in the left-right dimension in which most biologically relevant sound sources exist) across a variety of stimulus conditions and listening paradigms: The focus will be on the temporal dynamics of psychophysical sensitivity to *binaural* localization cues. After several psychophysical experiments are described (Chapters 2, 3 and 4), a simple peripheral auditory model is proposed to account for some of the reported psychophysical observations (Chapter 5). Before the details of the studies conducted may be set forth, however, some foundations in binaural hearing must be established. The remainder of this introductory chapter is devoted to this purpose.

1.1 Binaural cues for sound localization

Sound traveling from a source located directly before or behind a listener will reach the two ears at the same instant and at the same intensity. In any other case – that is, if the sound source is displaced to either side of the listener's median-sagittal plane – the sound will arrive at the ear nearer the source (1) shortly before it arrives at the other, introducing an interaural time difference (ITD), and (2) at a relatively greater intensity than at the other, introducing an interaural level difference (ILD) (see Figure 1.1). The maximum ITD for humans, attributable to the average speed of airborne sound (~340 m/s) and the average interaural distance of an adult (22-23 cm), is around 670 μ s, occurring for a sound source directly to the left or right (i.e., at $\pm 90^\circ$ azimuth, orthogonal to the listener's median plane) (Blauert, 1997). The maximum ILD is strongly dependent on the spectral content of the incident sound, attributable mainly to the

frequency-dependent shadowing effects of the head, which reduce the intensity of the sound at the ear further from the source. For sufficiently high-frequency sound sources slightly less than directly to the left or right, the ILD may be 20 dB or larger². ITD and ILD thus provide two physically distinct but complementary cues to sound source location that listeners might exploit for localization.

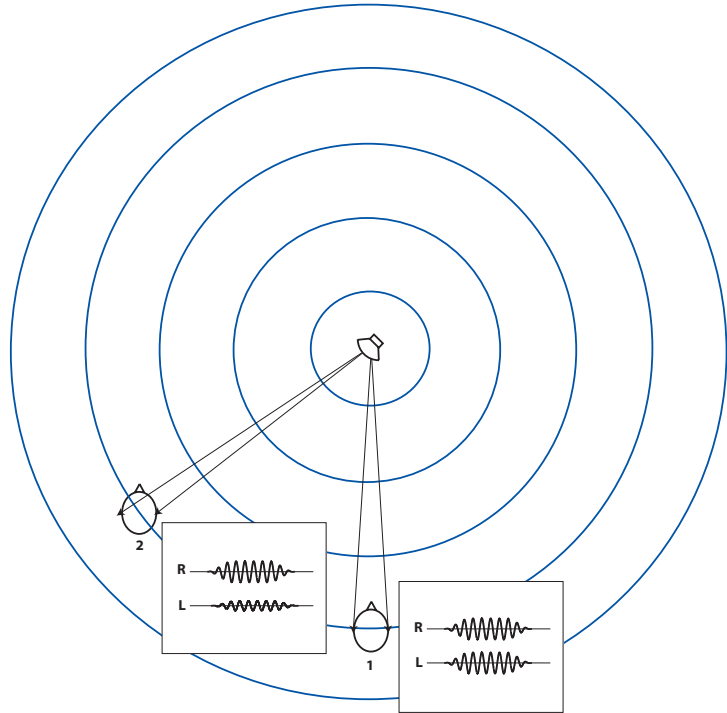


Figure 1.1: Binaural cues for sound localization. Sound traveling from a source directly in front of (or behind) a listener (1) reaches the right and left ears at the same instant with equal intensity (i.e., with 0 ITD and 0 ILD). In any other case (e.g., 2), the signal is delayed and attenuated at the ear further from the source relative to the ear nearer the source, giving non-zero ITD and ILD.

Until the early 20th century, it was believed that humans localized sound solely on the basis of ILD. As Rayleigh (1907) observed, however, low-frequency sounds, the wavelengths of which may be up to several meters, are little impeded by the head; thus, the ILD occurring for low-frequency sound is typically very small, and a seemingly insufficient cue for accurate localization at low frequencies. Listeners Rayleigh tested clearly were sensitive to the location of low-frequency sound sources, however (a 128 Hz. tuning fork, in one example); in the absence of another explanation, Rayleigh (1907) “reluctantly” concluded that listeners must be sensitive to disparities in signal *timing* at the two ears (i.e., ITD), which he proved with a series of investigations using pairs of slightly mistuned tuning forks (which therefore drifted in and out of phase with one another, causing listeners to report “beating” lateralization). Rayleigh (1907) is thus credited with proposing the “duplex theory” of sound localization, which posits that low-frequency sound is localized on the basis of ITD, and high-

frequency sound on the basis of ILD. Psychophysical studies in the century since Rayleigh have mainly supported his duplex theory (e.g., Zwislocki and Feldman, 1956; Mills, 1960; Wightman and Kistler, 1992), but two important revisions are noted in the following sections.

1.2 ITD carried by amplitude envelopes

All real-world sounds are amplitude-modulated (dynamic in intensity over time) on some scale. Different modulation rates produce qualitatively different percepts. At very slow modulation rates (e.g., <1 Hz), listeners perceive gradual changes in the loudness of the sound: Tuning forks like those used by Rayleigh (1907), for example, resonate with the greatest intensity immediately after being struck and decay gradually over the seconds that follow. When such nearly

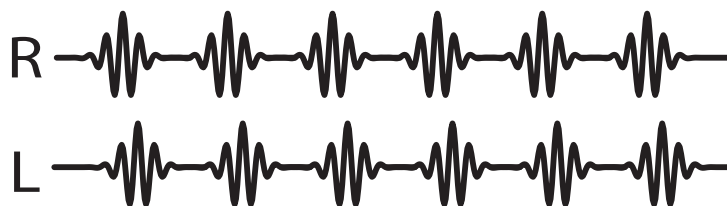


Figure 1.2: Envelope ITD. In addition to onset ITD and cycle-by-cycle ITD carried by the signal fine-structure, ITD is carried by the amplitude envelope. Listeners can detect such “envelope ITD” for carrier frequencies as high as ~ 10 kHz, provided the modulation rate is sufficiently low (Majdak and Laback, 2009; see text).

steady-state sounds reach a listeners’ ears, binaural cues are effectively conveyed only by the signal onset and in the cycle-by-cycle pressure fluctuations of the waveform. Consistent

with Rayleigh’s duplex theory, when the frequency of such signals exceeds approximately 1500 Hz, listeners appear to be completely insensitive to the cycle-by-cycle ITD (Zwislocki and Feldman, 1956; Dunai and Hartmann, 2011). Many real world sounds are modulated at much faster rates, however, such that the amplitude envelope fluctuates on a scale of milliseconds. For such sounds, ITD are effectively conveyed not only by the signal onset and the cycle-by-cycle fluctuations of the waveform, but also by the millisecond fluctuations of the amplitude envelope (see Figure 1.2). Henning (1974) was the first to demonstrate that listeners are indeed sensitive to “envelope ITD,” and it has since been shown that listeners can detect ITD in amplitude-modulated signals with carrier frequencies as high as at least ~ 10 kHz (provided the AM rate is beyond rates which produce the perception of “fluttering” loudness, and below

rates too high for effective transduction of the amplitude envelope, i.e., rates $>500 \text{ Hz}^3$, Majdak and Laback, 2009). Thus, high-frequency sound may be localized on the basis of both ILD and envelope ITD, an important revision to the duplex theory as originally proposed (Rayleigh, 1907).

1.3 Sensitivity to ILD across the spectrum

While sensitivity to ITD is strongly dependent on the spectral content and the modulation profile of the signal, sensitivity to ILD is relatively constant over a wide range of signal

parameters. Most notably, ILD discrimination thresholds measured under headphones are nearly invariant across the spectrum from 250-10000 Hz (Mills, 1960; see Figure 1.3). While the observation of ILD *sensitivity* at low frequencies may seem in conflict with the duplex theory, naturally-occurring ILD for signals $<1000 \text{ Hz}$ are typically negligible⁴

(Macauley *et al.*, 2010), and such signals are normally localized on the basis of the more salient ITD cue (Wightman and Kistler, 1992). An important exception was noted by Rakerd and Hartmann (1985), who demonstrated that listeners localize even low-frequency sound (albeit not very well) on the basis of ILD when ITD cues are “implausible,” in their experiment due to (1) the presence of an intentionally placed reflecting surface which severely distorted cycle-by-cycle ITD and (2) the removal of reliable onset cues by gating of the stimulus. The topic of onset- versus post-onset sensitivity and perceptual weighting of ITD versus ILD,

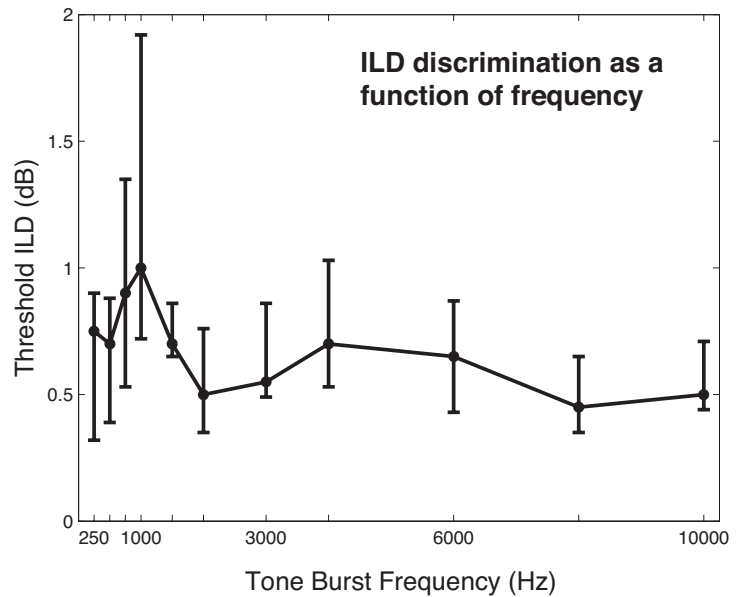


Figure 1.3: Replotted from Mills (1960). Sensitivity to ILD is nearly invariant across the spectrum. Thus, in addition to utilizing ILD for localization of high-frequency sounds, listeners may under some conditions utilize small but nonzero ILD carried by low-frequency sound. Error bars give the entire range of subjects' thresholds in Mills' study.

particularly in the context of acoustically complex environments, is central to the experiments described in the following chapters (see following section).

1.4 The importance of signal onsets

Acoustic events (perhaps obviously) consist of an “onset” portion and a “post-onset” portion. For sounds of some duration, the post-onset portion of the sound naturally exceeds the duration of the onset portion, but for all sounds (even very brief ones), the effective duration of the signal is in many environments (e.g., rooms) extended by tens or hundreds of milliseconds due to

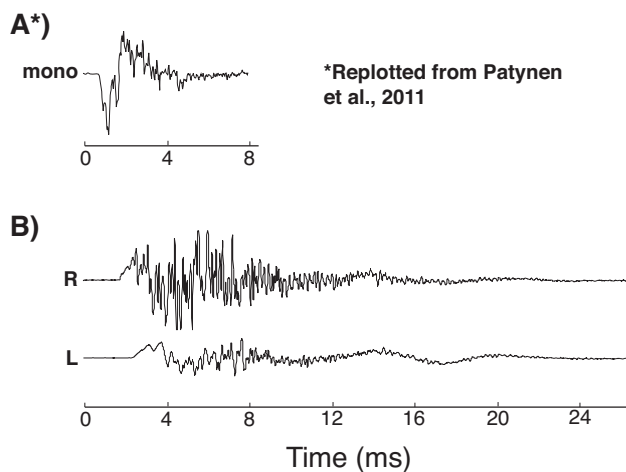


Figure 1.4: Recordings of balloon pops (approximate impulses) recorded in (A) an anechoic chamber with a single microphone and (B) a sound-attenuated booth with binaural manikin. In all but truly anechoic environments, reflected sound extends the signal duration substantially and adds spurious binaural information to post-onset portions of the signal.

reflections and reverberation of the sound from nearby surfaces. Thus, the physical localization cues (the ITD and ILD) reaching the ears are dominated by post-onset portions of the signal (see Figure 1.4). Paradoxically, a host of psychophysical studies have demonstrated that for many types of signals (low-frequency, high-frequency, and broadband, modulated and unmodulated alike), sound localization itself is dominated – and in some cases

determined entirely – by the cues present at the signal onset. Although the adaptive value of this phenomenon is clear – post-onset localization cues are frequently spurious (or at best redundant), arising from a combination of the direct sound and its reflected copies – onset-dominated localization has been given a number of different terminological labels including “the precedence effect” (Wallach *et al.*, 1949), “binaural adaptation” (Hafter and Dye, 1983; Hafter *et al.*, 1983; Hafter *et al.*, 1988), or simply “onset dominance” (Saberri and Perrot, 1995; Saberri, 1996; Freyman *et al.*, 1997), with little consensus on the

physiologic mechanisms responsible. Relevant past studies of these topics, which provided the context for the psychophysical experiments described in the following chapters, will be reviewed in due course. At present, it is interesting to note, in the context of spatial orientation as a general problem for organisms, that the perceptual dominance of localization cues at sound onset has been reported in species as diverse as cats, owls, and crickets (to name a few), and also for human localization in elevation (i.e., localization in the median-sagittal plane, a topic not substantively addressed in this dissertation) (reviewed in Litovsky *et al.*, 1999; see also Tollin and Yin, 2003). Thus, dependence on signal onsets seems to be a basic characteristic of auditory spatial orientation. Perhaps not coincidentally, robust onset responses and less robust post-onset responses are also a basic characteristic of primary auditory neurons (e.g., Kiang *et al.*, 1965; see Chapter 5).

1.5 Clinical considerations

Deaf and hearing-impaired individuals have benefited immensely from developments in assistive listening technologies over the past several decades. Although considerable progress has been made in the engineering of hearing aids and auditory prostheses, most notably cochlear implants – today many cochlear implant users and even some auditory brainstem implant users can hold telephone conversations (Shannon, 2012) – most patients continue to experience substantial difficulty in acoustically complex environments (e.g., social gatherings, Loizou *et al.*, 2009; or classrooms, Neuman *et al.*, 2012). The benefits of binaural hearing (versus monaural hearing) in such environments are substantial. In addition to more effective sound source localization and segregation, binaural unmasking can improve speech reception thresholds in noise by up to 15 dB (e.g., Levitt and Rabiner, 1967). Thus, toward restoration of binaural hearing, many patients now receive two hearing aids or cochlear implants rather than one. While some improvements in performance have been reported (e.g., Loizou *et al.*, 2009), the observed benefits among bilateral cochlear implant users in particular fall well short of those expected on the basis

of normal hearing data. Although device limitations and patient factors (e.g., age of onset and duration of deafness, Litovsky *et al.*, 2010) must contribute to this disparity, a complete understanding of limitations on performance among patient populations will require a deeper understanding of limitations on performance in the normal auditory system. The experiments described in the following chapters are also thus motivated by the author's interest in translational research. Relevant connections between normal-hearing and clinical data will be made where appropriate (see Chapters 3 and 6).

Textual footnotes

¹ Perception can probably be measured with greater acuity in humans than in non-human animals (Schone, 1986). Nonetheless, it has been demonstrated that barn owls, which depend on sound localization to eat (Knudsen and Konishi, 1978), and some species of frog, which depend on sound localization to find mates (Shen *et al.*, 2008), are at least as good at sound localization as humans. One would suspect that other as-yet unstudied species are also as good or better.

² The magnitude of ILD for a given source location increases nonmonotonically with increasing frequency. A degree of idiosyncrasy (across listeners) in ILD across the spectrum, attributable to the particularities of individual head shapes and sizes, has also been reported (e.g., Macauley *et al.*, 2010). Nonetheless, higher-frequency sound from sources slightly less than perpendicular to the interaural axis (e.g., $\pm 55^\circ$ azimuth) generally produce the largest ILD.

³ Listeners are neither able to reliably detect changes in amplitude modulation rate (a monaural task) nor detect ITD carried by amplitude envelopes for modulation rates much beyond ~ 500 Hz (e.g., Formby, 1988; Bernstein and Trahiotis, 2002), assumedly because the "ringing" cochlea cannot follow such rapid changes in amplitude with any fidelity. Limitations on temporal resolution in the auditory system comprise an expansive area of psychophysical and physiological auditory research, the majority of which is beyond the scope of this dissertation.

⁴ For sources nearer than ~ 1 m, ILD arises as a result of both shadowing by the head (the usual cause) and attenuation of sound with distance from the source (an effect of travel). Thus, ILD can be substantial for proximate low-frequency sources (e.g., Shinn-Cunningham *et al.*, 2000). For more distant sources, effects of travel are negligible and thus large ILD are relegated to higher frequencies.

2

Temporal weighting of interaural time and level differences carried by high-rate trains of filtered impulses*

An extensive body of psychophysical evidence indicates that listeners are able to accurately localize sound sources in echoic environments by responding to the cues carried by early-arriving direct sound rather than the spurious cues carried by later-arriving reflected sound. The various phenomena associated with this observation are known collectively as the “precedence effect” (Wallach *et al.*, 1949; for reviews, see Blauert, 1997; Litovsky *et al.*, 1999), a term coined by Wallach *et al.* (1949) in their seminal paper on the topic. In the paradigm used by Wallach *et al.* (1949) and dozens of authors since, two clicks simulating a direct source and single echo, hereafter referred to as the *lead* click and *lag* click, are presented from loudspeakers positioned to each side of a subject in the free field (Wallach *et al.*, 1949). The delay between the lead and lag is reduced until the subject perceives a single auditory event. This delay, typically 5-10 ms for impulsive stimuli such as clicks and up to 50 ms for running speech (see Litovsky *et al.*, 1999), is known as the echo threshold, the temporal limit of lead-lag *fusion*. At delays shorter than echo threshold, subjects localize the fused image in the direction of the lead (a phenomenon termed localization dominance), and are impaired in their ability to discriminate changes in the location of the lag (discrimination suppression). The terms “precedence effect” and “echo suppression” are often used synonymously. As Hartung and Trahiotis (2001) noted, however, “suppression” (in the sense of active suppression or inhibition of the lag) may not be necessary to produce the precedence effect physiologically (Chapter 5). Precedence may thus be understood more generally as a temporally dynamic “weighting” of localization cues, characterized by high weighting of cues carried by the lead and relatively lower weighting of cues carried by the lag.

The present experiment was designed to assess such “temporal weighting” for stimuli somewhat more complex than Wallach’s lead-lag pairs:

* Published under same title in [Brown, A.D. and Stecker, G.C. (2010). *J. Acoust. Soc. Am.* **128**: 332-341.]

trains of clicks carrying ITD or ILD, after studies which have historically been conducted under the terminological labels of “binaural adaptation” (e.g., Hafter *et al.*, 1983) or “onset dominance” (Freyman *et al.*, 1997). In the author’s view, the phenomena of precedence, binaural adaptation, and onset dominance are essentially identical. The question of interest was whether temporal weighting appeared to be identical for ITD and ILD, since the cues may be affected differently and may be differently utile over the stimulus duration in echoic environments (e.g., Rakerd and Hartmann, 1985; see Chapters 1 and 5).

2.1 Previous temporal weighting studies

A number of investigators over the past two decades have manipulated brief stimuli to assess the temporal dynamics of sensitivity to binaural cues over durations of tens to hundreds of milliseconds (e.g., Hafter and Dye, 1983; Freyman *et al.*, 1997). Generally, such studies have demonstrated that, just as “lead-lag” click pairs are spatially fused (i.e., appear to originate from a single location) at lead-lag delays shorter than around 5-10 ms (Wallach *et al.*, 1949), trains (rapid sequences) of clicks or other impulsive stimuli such as noise bursts are spatially fused for inter-click intervals (ICI) in the same temporal range. Fusion has been shown to persist for up to several seconds at short ICI (e.g., 2 ms), giving rise to a single auditory event at the location cued by the first click of a train, despite tens or hundreds of “lag” clicks carrying conflicting interaural information (Saberri and Perrot, 1995; Freyman *et al.*, 1997). This exaggerated salience of the onset click in relation to the ongoing clicks (i.e., onset dominance) has been quantified in several studies where the sensitivity of listeners’ judgments to variation in spatial cues has been measured independently for each temporal portion of a stimulus (e.g., each click in a train), and plotted as a function of time to form a *temporal weighting function* (TWF). To date, TWFs have been measured for ITD discrimination (under headphones) using narrowband click trains (Saberri, 1996) and broadband noise bursts (Dizon *et al.*, 1998), for free field localization using narrowband click trains (Stecker and Hafter, 2002; 2009), for median-sagittal plane localization using directionally

filtered noise bursts (Macpherson and Wagner, 2008), and for lateralization of electrical impulse trains in bilateral cochlear implant (CI) users (van Hoesel, 2008). Consistent with studies designed to measure the precedence effect explicitly, a typical finding in TWF studies has been that, for high-rate stimuli (e.g., ICI <5 ms), weights are high at onset (i.e., for the “lead”) and relatively low for the remainder of the stimulus. For lower-rate stimuli (e.g., ICI >5 ms), weights are roughly equivalent for each portion of the stimulus, resulting in a “flat” TWF indicative of uniform binaural sensitivity over the stimulus duration. That is, onset dominance has been evidenced at high rates, but not at lower rates, in agreement with maintained post-onset binaural sensitivity (reduced precedence) at lower rates (e.g., Hafter and Dye, 1983; Bernstein and Trahiotis, 2002).

A second feature observed in TWFs for sound localization in the free field (Stecker and Hafter 2002, 2009) and in virtual acoustic space (Macpherson and Wagner 2008) that is somewhat more difficult to relate to other studies of precedence, is an *increase* in the weighting of information near the end of the stimulus. This TWF feature, termed “upweighting” by Stecker and Hafter (2002, 2009), was observed over a wider range of ICI (3-8 ms) than was onset dominance (3 ms ICI only) in their study of free field localization (Stecker and Hafter 2002). Although upweighting appears consistent with the effects of temporal integration in ITD detection in noise (Zurek, 1980; Akeroyd and Bernstein, 2001; see discussion by Stecker and Hafter, 2009), upweighting has not been observed in headphone ITD studies explicitly measuring TWFs for ITD (Sabeti, 1996; Dizon *et al.*, 1998; van Hoesel, 2008), a discrepancy which led Stecker and Hafter (2002, 2009) to speculate that upweighting might be observed only in the presence of non-ITD cues (e.g., ILD and/or spectral cues), or alternatively, only in explicit localization (and not discrimination) tasks.

Recent investigations have attempted to characterize the time course of listeners’ sensitivity to ILD (e.g., van Hoesel, 2008; Stecker and Brown, 2010). Stecker and Brown (2010) demonstrated a profound impairment in high-rate ITD discrimination when click trains carried a diotic onset and dichotic offset (i.e., no ITD until after onset), but found no such impairment in ILD discrimination under

corresponding conditions, suggesting equal sensitivity to ILD at onset and offset. Comparable ILD sensitivity at onset and offset suggests either uniform sensitivity to ILD over the stimulus duration, or, alternatively, a combination of onset dominance and upweighting (as seen in free field TWFs, Stecker and Hafter, 2002, 2009). In a study explicitly measuring TWFs for ITD and ILD in bilateral CI users, van Hoesel (2008) found that ILD after onset contributed more to listeners' lateralization judgments than did ITD after onset (i.e., onset dominance was reduced for ILD), while neither cue appeared to produce upweighting. Bilateral CI users, however, typically demonstrate much better overall sensitivity to ILD than ITD (e.g., van Hoesel, 2004; Grantham *et al.*, 2008). Moreover, the effects of cochlear filtering, which heavily contribute to amplitude envelope processing and thus envelope ITD sensitivity in normal-hearing individuals (e.g., Bernstein and Trahiotis, 2002), are removed in electric hearing. Therefore, the time course of *normal-hearing* sensitivity to ILD (versus the more thoroughly characterized time course of sensitivity to ITD) remains to be established. Toward elucidating temporal weighting of ITD and ILD in the context of prior studies of the precedence effect and related phenomena, the present study directly measured TWFs for ITD and ILD discrimination in the same subjects.

2.2 Methods

All procedures, including recruitment, consenting, and testing of human subjects followed the guidelines of the University of Washington Human Subjects Division and were reviewed and approved by the cognizant Institutional Review Board.

2.2.1 Subjects

Six normal-hearing subjects aged 21-36 participated in this experiment. Subjects 0601, 0804, 0805 and 0815 were naive to the purposes of the experiment and were compensated for their participation. Subject 0502 was the author; 0501 was his advisor. All subjects demonstrated pure-tone detection thresholds < 20 dB HL at octave frequencies over the range 250-8000 Hz.

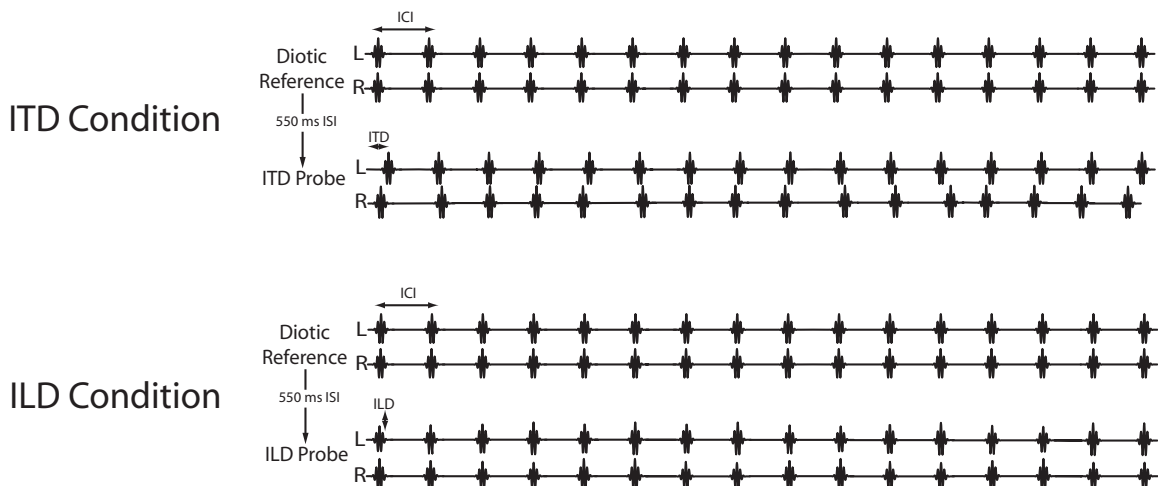


Figure 2.1: Schematic illustration of stimuli (not to scale). Each trial consisted of a diotic reference stimulus followed by a probe stimulus. The reference stimulus was comprised of 16 equal-amplitude Gabor clicks presented synchronously to the left and right earphones. Following a 550 ms silent interval (ISI), an ITD or ILD probe stimulus was presented, comprised of 16 Gabor click pairs with random ITD or ILD imposed on each. In the ITD probe, each click pair carried an ITD drawn from a uniform distribution of $\pm 100 \mu\text{s}$ centered on a base ITD of -100, 0, or 100 μs (0 in the above illustration); clicks were presented at equal amplitude to the two earphones in this condition. In the ILD probe, each click pair carried an ILD drawn from a uniform distribution of $\pm 2 \text{ dB}$ centered on a base ILD of -1, 0, or 1 dB (0 in the above illustration); clicks were presented synchronously to the two earphones in this condition. The inter-click interval (ICI), which corresponded to the rate of click presentation, was held constant within and between trials of a single run.

2.2.2 Stimuli and procedure

Stimuli were trains of 16 Gabor clicks (Gaussian-windowed tone bursts). Each click consisted of a 4 kHz cosine multiplied by a Gaussian temporal envelope with $\sigma=221 \mu\text{s}$. The resulting spectral bandwidth was also Gaussian, with $\sigma=750 \text{ Hz}$ (half-maximal bandwidth $\sim 1.8 \text{ kHz}$). Clicks were synthesized at 48.848 kHz (Tucker-Davis Technologies RP2.1, Alachua FL) and presented using STAX model 4070 closed-back electrostatic headphones (STAX LTD., Saitama, Japan). Click trains were presented with one of four peak-to-peak inter-click intervals (ICI) (1.25, 2.5, 5, or 10 ms ICI) at 65-74 dB SPL (65 dB SPL at 10 ms ICI; 74 dB SPL at 1.25 ms ICI). A 4-button response box (TDT RBOX) recorded subject responses. Testing was completed in a double-walled sound booth (IAC, Bronx, NY). Each trial consisted of two equal-duration click trains separated by a 550-ms silent interval. The click train duration was dependent on the ICI. The first train, presented as a reference, was always diotic; the probe train followed, presented with one of three different “base” interaural values with

additional interaural variation imposed on each click (see Figure 2.1). Two separate conditions employed ITD and ILD cues.

The base interaural value was varied pseudo-randomly from trial to trial. In the ITD conditions, the base values were -100, 0, or 100 μs ; in the ILD conditions, the base values were -1, 0, or 1 dB^1 . By convention, negative ITD and ILD values led to or favored the left ear; positive ITD and ILD values led to or favored the right. Additional random ITD or ILD variation, drawn from a uniform distribution of $\pm 100 \mu\text{s}$ or $\pm 2 \text{ dB}$, was added independently to each click in a train. These ranges were selected to correspond roughly to the $\pm 11^\circ$ azimuthal variation employed by Stecker and Hafter (2002, 2009), based on the physical correspondence between azimuth, ITD, and ILD described by Gulick *et al.* (1989), and an ITD/ILD trading ratio of approximately 50 $\mu\text{s}/\text{dB}$ (Hafter and Jeffress, 1968). Due to the per-click variation in ITD or ILD, trials with a base value of 0 carried nonzero interaural differences on each click, and thus a nonzero average ITD or ILD per trial.

The subject's task was to indicate via button press whether the probe stimulus was presented to the left or right of the reference stimulus. All subjects received training with immediate visual feedback until performance stabilized. Feedback was then turned off and testing began. One hundred trials comprised a single run; in each run, 60 trials had 0 base interaural difference, 20 were left-leading and 20 were right-leading. Base values were presented in random order over the course of each run; rate and cue type were randomized and counterbalanced between runs. Each subject completed four runs at each of the four ICIs in ITD and ILD conditions, giving 3200 trials in total (400 trials x 8 conditions). Only 0 base trials were included in the TWF analysis (1920 trials in total; 240 x 8 conditions; see Footnote 1).

2.2.3 Statistical methods

TWFs were computed following the logic of previous observer-weighting approaches (e.g., Berg, 1989; Saberi, 1996; Stecker and Hafter, 2002). In the present investigation, weights were computed using a receiver operating

characteristic analysis (ROC) analysis to quantify the accuracy with which listener responses could be predicted on the basis of interaural differences applied on each trial. For example, the “weight” of click #1 in a particular ITD condition for a particular subject was determined by classifying the subject’s responses (one “left” or “right” response per trial) on the basis of the values of ITD applied to click #1 (the decision variable); an ROC curve was generated by plotting hits (probability of classification “left” given the subject responded “left”) versus false alarms (probability of classification “left” given response “right”) for all presented criterion values of click #1 ITD (i.e., for ITD from $-100 \mu\text{s}$ to $+100 \mu\text{s}$). The area under the ROC (termed “ AUC_1 ” for click #1) can thus be construed as the weight assigned to click #1, since its value corresponds directly to the accuracy with which click #1 ITD predicts the subject’s response, expected to range between chance ($\text{AUC}=0.5$), indicating zero weight, and perfect ($\text{AUC}=1$). To generate the TWF, this procedure was repeated separately for each click in the train (giving AUC_2 , AUC_3 ... AUC_{16}). Note that this application of ROC analysis reverses the roles of stimulus and response variables relative to their usual roles in ROC analysis. Here, the analysis quantifies the accuracy with which stimulus values predict subject responses rather than the accuracy with which subjects discriminate stimulus values (see section 2.4.3 and Figure 2.5 for further discussion of the use of ROC analysis in the present study). An additional analysis classified subject responses using the mean ITD or ILD across clicks in a train as the decision variable. The resulting AUC (termed “ AUC_{mean} ”) describes the performance of a classifier that explicitly averages the ITD or ILD of all clicks in the train, after multiple-looks theory (cf. Viemeister and Wakefield, 1991) (note that AUC_{mean} is *not* the average of $\text{AUC}_1 \dots \text{AUC}_{16}$). Comparing AUC_1 against AUC_{mean} (i.e., $\text{AUC}_1 - \text{AUC}_{\text{mean}}$) thus provides a measure of onset dominance, since onset dominance entails salience of onset cues over the long-term interaural value (cf. Saberi and Perrot, 1995), where optimal performance in this task is facilitated by a cue-averaging strategy that minimizes decision variance (Saberi, 1996). A final analysis classified responses using only the maximum leftward or rightward interaural difference of each trial as the decision variable.

The resulting AUC (hereafter AUC_{max}) describes the performance of a classifier dominated by the most lateral click in a train, irrespective of that click's temporal position (see section 2.4.2). $AUC_1 \dots AUC_{16}$, AUC_{mean} and AUC_{max} were computed separately for each combination of subject, ICI and cue type (ITD or ILD) and subsequently compared in several null-hypothesis tests described in sections 2.3 and 2.4.

Finally, confidence intervals on weight estimates were generated using a permutation test (e.g., Moore and McCabe, 2005). For each combination of subject and condition, the subject's responses were randomly permuted across trials, and a TWF was generated following the AUC procedure described above. This procedure was repeated 1000 times; the standard deviation of estimated weights across repetitions was then used to estimate the standard error and to generate 95% statistical confidence intervals for each weight estimate. For the purpose of illustration, and because standard errors did not vary systematically with click position (ranging less than 0.002 across the train), summary 95% confidence intervals were computed from the mean standard error across clicks for each combination of subject and condition, indicated by gray shading around the chance performance level ($AUC=0.5$) in Figure 2.3. Similar confidence intervals are shown for the group-mean TWFs in Figure 2.2.

2.3 Results

2.3.1 Temporal weighting functions for ITD and ILD

Figure 2.2 displays TWFs averaged across subjects for each of the eight conditions tested. Filled circles plot the AUC obtained using each click ITD or ILD ($AUC_1 \dots AUC_{16}$), dashed lines plot the AUC obtained using average click train ITD or ILD (AUC_{mean}), and crosses plot the AUC obtained using maximum ITD or ILD within each click train (AUC_{max}). The shaded region plots the range of chance classification (95% CI of $AUC=0.5$), as described above. Overall, AUC_1 decreased with increasing ICI in both the ITD and ILD conditions. A monotonic increase in AUC_{mean} (dashed lines) with increasing ICI was also observed in both the ITD and ILD conditions. Onset dominance, indicated by $AUC_1 > AUC_{mean}$,

was apparent for ITD at ICI values of 1.25 and 2.5 ms, but only at 1.25 ms ICI for ILD. Although the overall shape of the TWF for ILD at 2.5 ms was similar to that for ITD, the mean ILD appeared to be the best overall predictor of subjects' responses in that condition; that is, although AUC_1

exceeded the values of $AUC_2 \dots AUC_{16}$, it did not exceed the value of AUC_{mean} . Related to this point, the mean value of

AUC_{mean} for ILD was greater than the value of AUC_{mean} for ITD at all ICI (AUC_{mean} values at 1.25, 2.5, 5 and 10 ms ICI for ILD=0.63, 0.67, 0.68, 0.69; for ITD=0.58, 0.61, 0.65, 0.67; also see Figure 2.4). Consistent with previous studies (Saberi, 1996; Stecker and Hafter, 2002), onset dominance was minimal at 5 ms and 10 ms ICI for both cues ($AUC_1 < AUC_{\text{mean}}$), and weights were nearly equal across clicks at 10 ms ICI. Finally, the value of AUC_{max} was less than the value of both AUC_1 and AUC_{mean} across conditions (see section 2.4.2). Taken together, the averaged TWFs suggest onset dominance for both ITD and ILD at short ICI, with a reduction in onset dominance for ILD relative to ITD, attributable specifically to greater cue-averaging for ILD (see section 2.3.2). The data do not suggest “upweighting” for ILD, as proposed by Stecker and Hafter (2002, 2009).

Averaged Temporal Weighting Functions

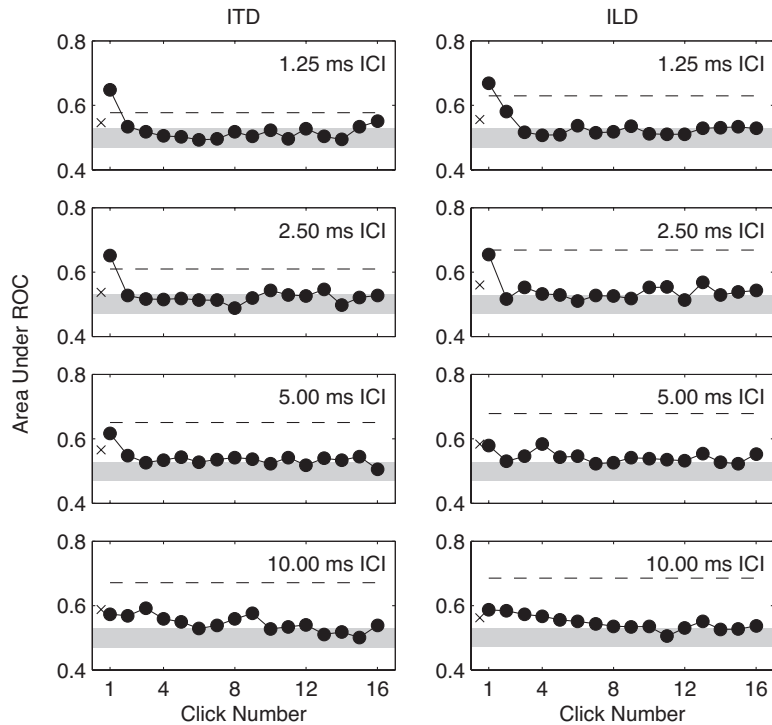


Figure 2.2: Temporal weighting functions (TWFs) averaged across subjects for ITD (left column) and ILD (right column) conditions for each inter-click interval (ICI; rows). Within each panel, per-click ITD or ILD weights ($AUC_1 \dots AUC_{16}$, filled circles) are plotted against click number to form the TWF; the weight of the mean ITD or ILD (AUC_{mean} , dashed line) and the maximum ITD or ILD (AUC_{max} , cross) are plotted for comparison (see text). The shaded region plots the range of chance classification (see text).

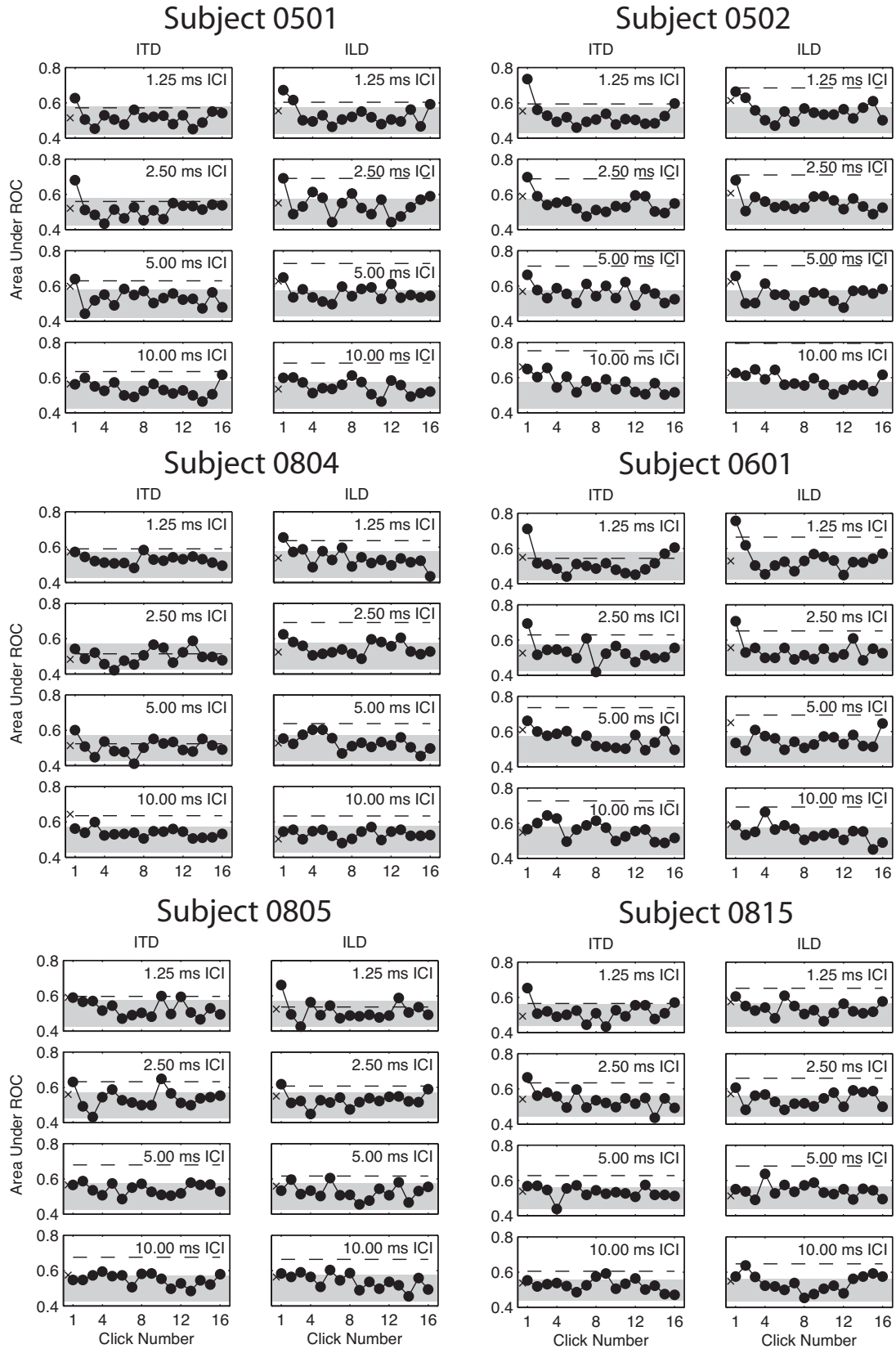


Figure 2.3: TWFs for individual subjects. Legend as in Fig. 2.2.

Figure 2.3 displays TWFs for individual subjects across each of the eight conditions tested (legend as in Fig. 2.2). Substantial intersubject variability is evident, with the data of subjects 0501, 0502, 0601 and 0815 generally in agreement regarding the effects of ICI and cue type on onset (AUC_1) and mean (AUC_{mean}) weights, but the data of subjects 0804 and 0805 diverging from this pattern. Most strikingly, neither 0804 nor 0805 exhibited onset dominance in the ITD condition. This result is inconsistent with past studies which have measured TWFs for ITD discrimination (Sabeti, 1996; Dizon *et al.*, 1998) and with the TWFs of other subjects in the present investigation, although it is consistent with the individual variability in ITD lag discrimination thresholds measured by Sabeti and Antonio (2003) in a population study of the precedence effect, which suggested that onset dominance in ITD discrimination occurs for most but not all listeners. Both subjects 0804 and 0805 reported that it was often difficult to lateralize the ITD stimulus. Our initial assumption was that these subjects simply had higher overall ITD discrimination thresholds and thus lacked sensitivity to $\pm 100 \mu\text{s}$ ITD variation about the midline. Thus, we attributed their unique TWFs to lower overall ITD sensitivity rather than to different time courses of sensitivity, and, accordingly, attempted testing both subjects with a greater number of suprathreshold ITD clicks per trial by increasing levels of ITD variation to $\pm 200 \mu\text{s}$ and subsequently $\pm 150 \mu\text{s}$. For both manipulations, the subjects reported that the probe sounded “broken up,” but no easier to lateralize. The perception of “broken” spatial fusion in this context is most interesting, as it might be construed to imply that these subjects actually were sensitive to the presented ITD but, in the absence of onset dominance, perceived a diffuse image centered at the midline, reflecting the expanded range of left- and right-leading ITD. Nonetheless, since the wider range of ITD variation resulted in a percept that was qualitatively different from the spatially compact images perceived by other subjects with less ITD variation, we elected to proceed with $\pm 100 \mu\text{s}$ variation for all subjects, similarly maintaining ± 2 dB variation for all subjects in the ILD conditions. Subjects 0804 and 0805 did not report difficulty lateralizing the ILD probe, and their data agreed more closely with the averaged TWFs in that

condition. Whether subjects 0804 and 0805 simply had higher ITD than ILD thresholds (McFadden *et al.*, 1973) or were rather more similar to listeners lacking onset dominance in ITD discrimination (Saberri and Antonio, 2003) is a matter that cannot be definitively resolved in the present investigation since absolute sensitivity was not measured, although additional intersubject variability in ILD TWFs (see section 2.3.2) may lend support to the latter idea.

Importantly, upweighting was not observed in any of the conditions tested. The suggestion of Stecker and Hafter (2002, 2009) that upweighting might be attributable to ILD sensitivity near stimulus offset is thus not supported by the present data. The prediction of upweighting in this experiment would be for AUC values near the end of the train to exceed other post-onset AUC values at intermediate ICI (e.g., 5 ms ICI). In all of the conditions tested, for both ILD and ITD, AUC_1 invariably exceeded AUC_{16} . Thus, contrary to the suggestion of Stecker and Hafter (2002, 2009), the time course of ILD sensitivity does not appear to feature upweighting. Rather, TWFs measured by the present method for both ITD and ILD are consistent in terms of individual click weights with those described by Saberri (1996) for ITD lateralization and by van Hoesel (2008) for ITD and ILD lateralization by bilateral CI users, with high onset weights followed by relatively low and uniform post-onset weights at short ICI. Possible differences in the TWFs for ITD versus ILD (van Hoesel, 2008) – specifically, the apparently higher salience of post-onset ILD than ITD suggested by higher weighting of mean ILD than mean ITD – are considered in the following sections.

2.3.2 Reduced onset dominance for ILD

Although averaged TWFs (Fig. 2.2) for ITD and ILD appeared similar in terms of *individual* click weights, mean ILD appeared to be a more salient cue than mean ITD at all ICI, and a more salient cue than onset ILD at 2.5 ms ICI, suggesting that the salience of onset ILD may be tempered by the salience of the mean ILD even at short ICI, while onset ITD alone dominates perception at short ICI. Clear intersubject variability was present, however, in individual TWFs (Fig. 2.3) for both ITD and ILD conditions: of the four subjects demonstrating clear

onset dominance for ITD at 1.25 and 2.5 ms ICI, two (0502 and 0815) exhibited no onset dominance for ILD regardless of ICI, while one (0501) experienced onset dominance for ILD at 1.25 ms only. One subject (0601) demonstrated onset dominance for ILD at both rates, but also relatively greater weighting of the mean ILD at those rates than in the corresponding ITD conditions. As noted above, the data of subjects 0804 and 0805 differed from the mean data. Subject 0805 was particularly

divergent from the others, exhibiting the greatest degree of onset dominance among any subject at 1.25 ms ICI in the ILD condition, while exhibiting no onset dominance in ITD conditions and greater AUC_{mean} values for ITD than ILD at all rates.

Bearing in mind the

possible importance of individual differences discussed heretofore, the overall observation that subjects exhibited onset dominance at high rates, but to a reduced degree for ILD compared to ITD (particularly at 2.5 ms ICI), is broadly consistent with the studies of van Hoesel (2008) and Stecker and Brown (2010), as well as with the interpretation of results presented by Saberi *et al.*, (2004), which suggested onset dominance is stronger in ITD than ILD discrimination.

Figure 2.4 plots a derived measure of onset dominance, $AUC_1 - AUC_{\text{mean}}$, across ICI for both ITD and ILD (averaged across subjects). $AUC_1 - AUC_{\text{mean}}$ for ITD exceeded that for ILD at all ICI, anticipated from the observation that AUC_{mean} was greater for ILD than ITD at all ICI (see Fig. 2.2), although the difference was minimal at 1.25 and 10 ms ICI. With the caveat of individual

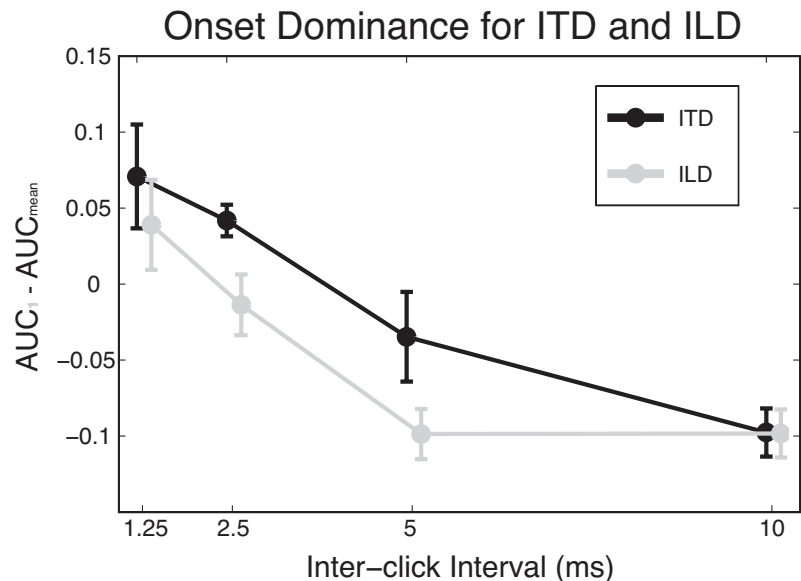


Figure 2.4: Cross-subject averages of $AUC_1 - AUC_{\text{mean}}$ against inter-click interval (ICI) for ITD (black) and ILD (gray) conditions. Error bars give \pm one standard error of the mean. Data suggested a shift from onset dominance at high rates to cue-averaging at low rates, with greater cue-averaging for ILD than ITD.

differences, the cross-subject effects of ICI and cue type on AUC_1-AUC_{mean} were assessed with a 4x2 repeated measures ANOVA. The main effect of ICI was significant ($F(3,15)=17.53, p<0.05$), while neither the main effect of cue type ($F(1,5)=3.64, p=0.12$) nor the ICI by cue type interaction ($F(3,15)=1.20, p=0.34$) were significant, doubtless the result of a large subject-by-cue type error term arising from the disparity of subjects 0804 and 0805 and the others. Although performing the analysis without these “divergent” subjects did yield a significant main effect of cue ($F(1,3)=26.89, p<0.05$) – a result consistent with reduced onset dominance in ILD relative to ITD (van Hoesel, 2008, Stecker and Brown, 2010) - the work of Saberi and colleagues (2003, 2004) suggests that the large subject-by-cue type “error” term may actually represent significant natural variation in the time course of ITD versus ILD sensitivity among the general population. These findings are discussed further in the context of the precedence effect (and binaural adaptation and upweighting) in the following sections.

2.4 Discussion

2.4.1 Onset dominance for ITD and ILD

The present results suggest that normal hearing listeners are more sensitive to the ITD or ILD carried by the onset of a rapidly fluctuating signal envelope than to the ITD or ILD carried by any discrete post-onset portion of the signal, consistent with previous studies of the precedence effect and related phenomena (e.g., Wallach *et al.*, 1949; Zurek, 1980; Hafter *et al.*, 1990; Saberi, 1996); the ITD or ILD presented at the onset of a brief 16-click train was found to be more predictive of subjects’ discrimination judgments than ITD or ILD presented in any of the 15 subsequent clicks at $ICI \leq 5$ ms. The present results additionally suggest, however, that listeners’ lateralization of rapidly fluctuating signal envelopes carrying ILD is influenced by the mean signal ILD; the influence of mean ILD exceeded that of onset ILD at $ICI > 1.25$ ms on average. This was not the case for ITD, and mean ILD was a better predictor of subjects’ responses than mean ITD across the four tested ICI for all but one subject. This finding compares favorably to TWFs measured for ITD and ILD lateralization by bilateral

CI listeners (van Hoesel, 2008), although larger differences between ITD and ILD were reported in that study, and substantial intersubject variability was noted in the present study, with one subject in particular (0805) displaying a pattern exactly the opposite of other subjects – an observation itself consistent with previous reports of variable onset dominance in both ITD and ILD “precedence” among the general population (Sabeti and Antonio, 2003; Sabeti *et al.*, 2004).

The present report is the only report, to our knowledge, of TWFs for ILD in normal-hearing listeners. Previous TWF investigations in normal-hearing listeners have indicated near-complete onset dominance for high-rate stimuli carrying ITD (Sabeti, 1996; Dizon *et al.*, 1998), and strong onset dominance for high-rate free field stimuli carrying ITD, ILD and spectral cues in combination (Stecker and Hafter, 2002). Pairing our results with existing findings suggests that listeners are acutely sensitive to both ITD and ILD in the early-arriving sound (i.e., cues of the “lead”), while reduced onset dominance for ILD relative to ITD in the majority of our subjects (specifically, higher weighting of mean ILD than mean ITD), suggests that sensitivity to ILD must somehow persist over the duration of the stimulus. This finding is inconsistent with the suggestion of equal “binaural adaptation” for ITD and ILD arising from a single mechanism (Hafter *et al.*, 1990), but *is* consistent with both the “level-meter” hypothesis of Hartmann and Constan (2002), which posits that ILD sensitivity entails a running integration of ILD over the stimulus duration, and with evidence for higher “lead-lag” echo thresholds for ITD than ILD (Krumbholz and Nobbe, 2002; see Chapter 4).

In spite of better sensitivity to ILD than ITD over the duration of the stimulus, upweighting for ILD, as proposed by Stecker and Hafter (2002, 2009), was clearly not present in the measured TWFs; an interesting interpretation of the results of Stecker and Brown (2010) may be given in this context: Although equivalent thresholds for ILD detection at stimulus onset and offset reported in that study could arise from an elevation of sensitivity at onset and again at offset (cf. Stecker and Hafter, 2002, 2009), the same results can be explained by sensitivity to the *mean* ILD, since the mean ILD information available was identical in the onset and offset ILD conditions of that study.

A final and important consideration is that the several investigations discussed here (including the present), which have sampled small numbers of subjects, may appear to offer partially conflicting results because, in fact, the time course of sensitivity to binaural cues varies significantly among normal-hearing individuals (Saber and Antonio, 2003; Saber *et al.*, 2004). The importance of careful attention to individual differences in studies of this type rather than exclusive emphasis of group statistics might thus be underscored; the question of why the seemingly low-level phenomena under study (cf. Hafter *et al.*, 1990; Chapter 5) would vary among normal hearing listeners is a fascinating one which certainly cannot be addressed by discussing what “average” subjects do.

2.4.2 Cue-averaging versus sensitivity to peak interaural value

In several conditions, elevated values of AUC_{mean} were observed despite low values for individual click weights. We have interpreted this result as evidence for cue-averaging in those conditions. Importantly, the relatively greater AUC_{mean} values for ILD versus ITD suggest a greater cue-averaging of ILD than ITD in the majority of our subjects. An alternative interpretation, however, might be that judgments reflected a minority of clicks which carried large interaural differences but which varied in temporal position from trial to trial (e.g., the *peak* interaural value on each trial). In that case, low individual-click weights would reflect the low probability (1/16) that the peak value occurred at a given click number, and the high value of AUC_{mean} would reflect the strong influence of these outlying values on the mean ITD or ILD across clicks, rather than cue-averaging by the subjects (note that such an account would still imply sensitivity to post-onset clicks, despite the lack of explicit cue-averaging). We examined this possibility by computing AUC_{max} , classifying subject responses on the basis of the maximum ITD or ILD occurring on each trial. AUC_{max} values are plotted as crosses in Figs. 2.2 and 2.3, where they appear consistently lower than both AUC_{mean} and AUC_1 . A 2x4x2 repeated measures ANOVA with factors of measure (AUC_{mean} , AUC_{max}), ICI, and cue type confirmed that AUC_{mean} was a significantly better predictor of subject response than AUC_{max} ($F(1,5)=184.81$, $p<0.05$),

indicating that occasionally large interaural values did not account for the salience of the mean. Indeed, the value of AUC_{max} across conditions and rates was similar to the AUC values of individual post-onset clicks ($AUC_2 \dots AUC_{16}$), as would be expected under the null hypothesis since the long-term AUC_{max} samples, on average, each click position equally. In contrast, AUC_{mean} increased monotonically with increasing ICI, giving rise to a significant measure by ICI interaction ($(F(1,5)=9.93, p<0.05)$), consistent with better cue-averaging at slower rates, where subjects were sensitive to interaural information in each click.

2.4.3 ROC analysis for TWFs

The application of ROC analysis in the present investigation differs from its usual application and from previous approaches to observer weighting

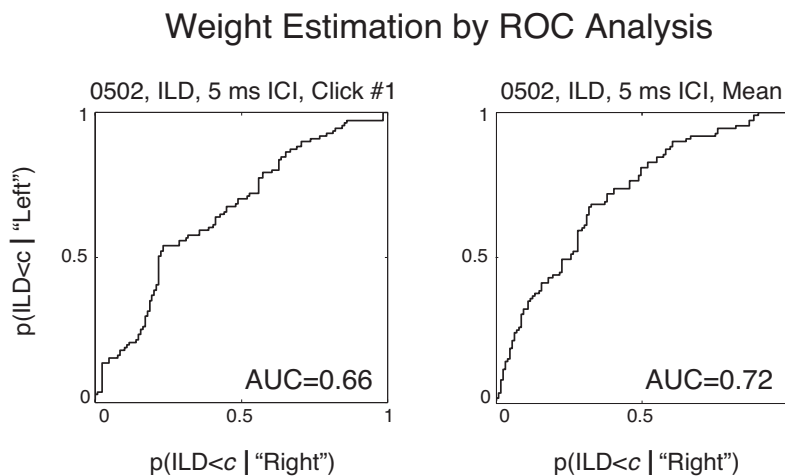


Figure 2.5: Two ROC plots generated by the present method (subject 0502, ILD condition, 5 ms ICI, for click #1 ILD [left panel] and mean ILD [right panel]). In this example, the ROC was given by the accumulation of hits (ILD < criterion c for response “Left”) and false alarms (ILD < criterion c for response “Right”) as the criterion c was adjusted across all presented values of ILD, plotted as a probabilities ranging from 0 to 1. The AUC, computed by integrating the ROC with respect to false-alarm probability, gives the percent correct classification of responses by the ILD of click #1 (left) and the mean ILD of the click train (right). The same procedure was applied for each combination of subject, condition, and classifier to generate the AUC weights given throughout the paper.

analysis such as COSS analysis (Berg, 1989; Saberi, 1996), linear regression (Stecker and Hafter 2002; Macpherson and Wagner, 2008; van Hoesel, 2008), and point-biserial correlation (Richards and Zhu, 1994, Lutfi, 1995). In such studies, computed weights have conventionally been normalized to some standard

referent (e.g., the weight of the onset, van Hoesel, 2008) and thenceforth presented as “relative weights.” Normalized weights can then be evaluated in terms of relative salience (i.e., the perceptual weight of a stimulus component

relative to other components) and, normalized for each subject, are amenable to intersubject comparisons, while raw computed weights are more cumbersome in their interpretation (e.g., in COSS analysis, Saberi, 1996) and may differ in magnitude between subjects (e.g., Stecker and Hafter, 2002). Thus, an immediate advantage of ROC analysis is that it produces raw weights which can be interpreted directly and compared across conditions and across subjects without normalization. Specifically, computed AUC values in the present investigation correspond to the degree to which particular stimulus features predicted the listener's response (i.e., proportion correct classification of responses). This use of ROC analysis resembles previous applications of ROC analysis to neural data, where it has been used to assess the degree to which a neural response measure (e.g., spike count) correctly predicts the stimulus eliciting that response (e.g., Skottun *et al.*, 2002). Importantly, the interpretation of AUC weights does not change with the statistics of the feature used for classification. Here, we used three types of features: (1) ITD or ILD imposed on single clicks in the train ($AUC_1 \dots AUC_{16}$), (2) the mean ITD or ILD across clicks in a train (AUC_{mean}), and (3) the largest click ITD or ILD imposed in each train (AUC_{max}). Regarding (2) specifically, ROC analysis produces, by the same procedure applied in (1) and (3), a "cue-averaging" weight which can be directly compared to individual click weights – a parsimonious means of evaluating the salience of mean ITD or ILD.

Figure 2.5 displays two example ROC plots generated by the present method (subject 0502, ILD condition, 5 ms ICI, ROC for click #1 [left panel] and mean ILD [right panel]). Note that ROC curves were not interpolated; rather, in this example, the ROC was given by the accumulation of "hits" ($ILD < \text{criterion } c$ for response "Left") and "false alarms" ($ILD < \text{criterion } c$ for response "Right") for the 240 trials with 5 ms ICI and 0 dB base ILD as the criterion c was adjusted across all presented values of ILD. The AUC was computed directly by numerical integration, corresponding in this case to the percent correct classification of subject 0502's responses by click #1 ILD (left panel) and mean ILD (right panel). To ascertain that the ROC method of weight estimation was comparable to a

more conventional method, we also computed TWFs using logistic regression. The resulting logistic TWFs did not differ qualitatively from those computed using ROC analysis, but lacked both the simplicity of interpretation and parsimony provided by the ROC method. We thus suggest that the ROC method may be a useful approach in future studies of observer weighting employing discrimination paradigms.

2.4.4 Cues in agreement versus cues in conflict

In considering the differences between TWFs measured in the free field by Stecker and Hafter (2002, 2009) and those measured in this and other studies where ITD or ILD were manipulated independently, (e.g., Saberi, 1996; van Hoesel, 2008), it should be kept in mind that free field stimulation presents ITD and ILD in agreement, whereas the current and related headphone studies present a situation in which the two cues conflict (i.e., a dichotic ITD or ILD combined with a diotic ILD or ITD). Although a number of studies have investigated binaural sensitivity to various combinations of ITD and ILD in conflict and agreement (e.g., Hafter and Jeffress, 1968, Hafter *et al.*, 1990), little is known about the possible effects of cue conflict on the temporal integration of individual binaural cues. While it seems unlikely that the effect of cue agreement versus conflict could give rise to drastically different TWFs (i.e., the presence or absence of upweighting), the notion of combination-dependent changes in sensitivity to one cue over the other is generally consistent with the “plausibility hypothesis” of Rakerd and Hartmann (1985).

2.4.5 Possible effects of the task employed

One explanation for upweighting in TWFs measured by Stecker and Hafter (2002, 2009) concerns the nature of the task employed in the estimation of TWFs. Stecker and Hafter (2002, 2009) measured TWFs for free field localization by asking subjects to point to the location of click trains comprised of clicks presented randomly from speakers in an array about the subject; TWFs were then computed by multiple linear regression (per-click “slope” plotted against click number). Free field localization is different than intracranial lateralization

(e.g., Plenge, 1974), and a number of studies have suggested additional important differences between explicit localization (or lateralization) versus discrimination of binaural stimuli (Bernstein and Trahiotis, 1994; Buell *et al.*, 1994, Tollin and Henning, 1998). Stecker and Hafter (2002, 2009) thus suggested that the difference between orienting to a sound source and simple discrimination might underlie differences in TWFs for localization (Stecker and Hafter, 2002, 2009; Macpherson and Wagner, 2008) versus discrimination tasks (e.g., Saberi, 1996) – specifically, the finding of clear upweighting in the former case versus no upweighting in the latter case. That is, they speculated that upweighting might reflect the operation of memory mechanisms (e.g., temporal integration) required for programming motor responses in an explicitly spatial task, but not for other types of tasks. The current results are consistent with that view in that discrimination of either cue presented over headphones failed to produce upweighting. Future work explicitly comparing TWFs across multiple tasks could provide further insight on such discrepancies in the literature.

2.4.6 The range of interaural variation

A final but critical consideration is whether the nature of interaural variation across clicks exerted any influence on the measured TWFs. For example, a restricted range of variation might be expected to result in less sensitivity to TWF features (such as onset dominance or upweighting) than would a larger range of variation. Too large a range of variation, however, could cause the auditory image to become diffuse, as indeed occurred when we increased the range of ITD variation for subjects 0804 and 0805 (see section 2.3.1). Interaural variation in the current study utilized a uniform distribution spanning $\pm 100 \mu\text{s}$ of ITD or $\pm 2 \text{ dB}$ of ILD. Those ranges were selected to correspond roughly to the $\pm 11^\circ$ azimuthal variation employed by Stecker and Hafter (2002, 2009), based on acoustic measurements of ITD and ILD (Gulick *et al.*, 1989). The ITD range was also roughly similar to that employed by Saberi (1996), who employed Gaussian distributions of ITD with $\sigma = 100 \mu\text{s}$. The similarity of TWFs obtained in that study and the current report suggest that the range of ITD variation was sufficient for

the stated purpose. Likewise, the onset dominance observed at 1.25 ms ICI in the ILD condition suggests that the ILD range was similarly sufficient for detection of non-uniform TWFs. Thus, reduced onset dominance at longer ICI and the failure to observe upweighting does not likely reflect insensitivity of the method due to insufficient cue variation.

2.5 Summary and conclusions

The current study measured temporal weighting functions for ITD and ILD sensitivity in the same subjects in order to directly compare the temporal dynamics of sensitivity to the two cues for stimuli consisting of a “lead” and multiple “lag” clicks. The findings of the current study suggest the following conclusions:

- 1) Consistent with prior studies, discrimination of ITD in high-rate (short ICI) click trains was dominated by the ITD of the onset click. For ICI ≥ 5 ms, ITD discrimination appeared to reflect the average ITD across clicks.
- 2) Discrimination of ILD followed a pattern qualitatively similar to discrimination of ITD, but strong onset dominance was restricted to the shortest ICI tested (1.25 ms), and cue-averaging appeared more robust for ILD than for ITD at all ICI.
- 3) Substantial intersubject variability in (1) and (2) was evident, suggesting onset dominance for ITD and ILD occurs to varying degrees across listeners, consistent with the reports of Saberi and Antonio (2003) and Saberi *et al.* (2004).
- 4) The similar pattern in individual click weights for ITD and ILD is consistent with studies reporting similar time-courses of sensitivity to the two cues (Zurek, 1980; Hafter and Dye, 1983; Hafter *et al.*, 1983), but the apparently greater influence of average (including post-onset) ILD suggests that the “precedence effect” may be weaker for ILD than ITD (see Chapters 4 and 5).
- 5) Neither ITD nor ILD TWFs showed “upweighting” of late-arriving sound (Stecker and Hafter, 2002, 2009), contradicting the hypothesis of Stecker

and Hafter (2002, 2009) that upweighting observed in free-field localization might reflect a specific contribution of late-arriving ILD.

- 6) Thus, the discrepancy between free field and headphone TWFs is likely not explained by the simple difference in binaural cues available. Rather, the discrepancy might relate to differences between conflicting and congruent cues, spectral cues, differences in the task employed, or differences in the externalized vs. intracranial perception of the sounds. Future work should address these possibilities.

Textual footnotes

¹ AUC values (and thus TWFs) were computed and compared only on trials where the base ITD or ILD was 0. Trials with nonzero base ITD and ILD were included in the experiment to mediate the perceived task difficulty and to provide a basis for feedback during training. Although per-click variation in ITD or ILD was likely effective in biasing the *degree* of lateralization perceived, our discrimination task was inherently insensitive to gradations of lateralization: nonzero base trials, being in general strongly lateralized according to the sidedness of the base, were instead useful in sustaining vigilance between 0 base trials (on which biasing across the midline could be measured), and in ascertaining that subjects could perform the task at the outset of the experiment. Mean values for nonzero base conditions (-1,1 dB ILD, -100,100 μ s ITD) were selected to produce approximately equivalent lateralization in both conditions.

3

Temporal weighting of interaural time and level differences. II. The effect of binaurally synchronous temporal jitter*

As evidenced in Chapter 2, listeners' sensitivity to binaural cues carried by amplitude-modulated (AM) signals – particularly their sensitivity to envelope ITD – is limited by modulation rate: At modulation rates near 100 Hz (or 10 ms ICI in the case of click trains), normal-hearing listeners are comparably sensitive to ITD carried by high-frequency AM signals and ITD carried by low-frequency pure tones, while for modulation rates beyond 200-300 Hz ITD (ICI <5 ms), sensitivity declines precipitously (Henning, 1974; Hafter and Dye, 1983; Buell *et al.*, 2008; Majdak and Laback, 2009). Hafter and colleagues (and others already mentioned) have attributed this rate-dependence to a loss of sensitivity to post-onset binaural information, a phenomenon they termed “binaural adaptation” (Hafter *et al.*, 1988; see previous chapters). Although, as earlier suggested, binaural adaptation, the precedence effect, and onset dominance seem to be essentially similar phenomena, indeed dependent on a rapid decay of binaural sensitivity after signal onset, the paradigm employed by Hafter and colleagues differs somewhat from those typically employed in studies explicitly focused on the precedence effect or onset dominance. In their paradigm, the binaural cues in the “leading” and “lagging” portions of the stimulus are identical: reduced sensitivity to post-onset binaural information (versus onset or “lead” information) is evidenced only by listeners' very limited improvement in binaural cue discrimination with increasing stimulus duration (i.e., a failure to integrate post-onset binaural information for improved sensitivity) at high AM rates (short ICI).

3.1 Effects of temporal irregularity on binaural sensitivity

One unexpected feature of binaural adaptation described by Hafter and Buell (1990) and later Freyman *et al.* (1997) is a so-called “binaural restarting” phenomenon, wherein a temporally irregular modulation in a high-rate AM signal (e.g., a lengthening or shortening of the interval between successive click pairs)

* Published under same title in [Brown, A.D. and Stecker, G.C. (2011). *J. Acoust. Soc. Am.* **129**: 293-300.]

improves ITD discrimination performance. More recent studies have shown that ITD sensitivity at high rates may be similarly enhanced by randomly perturbing the timing of electrical stimulation (Laback and Majdak, 2008) or the temporal envelopes of acoustic signals (Goupell *et al.*, 2009), leading the authors of those studies to suggest that the benefit of binaurally synchronous “temporal jitter” might be attributable to a restarting of the “adapted” binaural system (after Hafter and Buell, 1990). An alternative explanation for the results of Laback and Madjak (2008), first offered by van Hoesel (2008a, 2008b), is that jitter improves ITD sensitivity simply by slowing the pulse rate over discrete portions of the stimulus and reducing the inter-pulse ambiguity of post-onset ITD (cf. Freyman *et al.*, 1997). Goupell *et al.* (2009) evaluated this possibility, but found that normal-hearing listeners were better able to discriminate ITD carried by jittered 1200 pulse-per-second (pps) stimuli than isochronous 600 pps stimuli, even though the 1200 pps stimuli contained inter-pulse intervals (IPI) at most equal in duration to the IPI of the isochronous 600 pps stimuli.

The effects of temporal irregularity on the temporal dynamics of binaural sensitivity are of interest for two quite different reasons. Firstly, normal-hearing data on temporal weighting of binaural cues are frequently extrapolated (particularly in paper discussions and conference presentations) to “real-world” listening situations: Studies of the precedence effect using single lead-lag pairs (after Wallach *et al.* 1949), for example, purport to offer insight on how reflected sound might be treated by the auditory system in real-world environments. Real-world environments, however, contain multiple reflecting surfaces and thus introduce multiple, perhaps dozens of discrete early reflections (moreover, many real-world sounds are ongoing or repeated, an issue returned to in Chapter 4). Although studies of “onset dominance” (including the one comprising the previous chapter) have often employed trains of clicks, which may be construed as a lead click and a *series* of lag clicks, these stimuli have almost always been (with the exception of Goupell *et al.*, 2009) temporally isochronous. Real-world environments *do not* produce temporally regular reflections. Thus, understanding the ecological utility of the precedence effect will require, among a host of other

considerations (e.g., considering the consequences of frequency distortion and attenuation of reflected sound), elucidating the effects of temporal irregularity on binaural and especially onset versus post-onset binaural sensitivity.

The second reason to study the effect of temporal irregularity on binaural sensitivity is clinical in nature. Bilateral CI users' sensitivity to ITD is notoriously poor. Even when the timing of the electrical pulse trains delivered by the two implants is precisely controlled by a research processor (Laback and Majdak, 2008), bilateral CI users exhibit marked deficits in binaural task performance. Laback and Majdak (2008) found, however, that imposing binaurally synchronous temporal jitter during bilaterally synchronized CI stimulation significantly improved ITD sensitivity in all of the CI patients they tested. Although the possible application of these findings for the clinical improvement of ITD sensitivity in bilateral CI users is intriguing, a number of questions concerning the effect of jitter remain. The effect of jitter on listeners' sensitivity to ILD, for example, has not been measured, despite the fact that sound localization by bilateral CI users (using their clinical processors) currently depends almost entirely on sensitivity to ILD (see van Hoesel and Tyler, 2003; Grantham *et al.*, 2008, Litovsky *et al.*, 2010). Examining the effect of jitter on ILD sensitivity is thus essential to further consideration of its clinical implementation; unintended disruption of ILD sensitivity by jitter could at least negate any gains in ITD sensitivity.

In the previous investigation (Chapter 2), normal-hearing listeners' sensitivity to ITD and ILD was characterized by TWFs for discrimination of ITD or ILD carried by brief *isochronous* click trains (Brown and Stecker, 2010). The basic premise of binaural adaptation – that binaural sensitivity is dominated by onset ITD and ILD for rapidly amplitude-modulated stimuli, while post-onset information contributes relatively less to subjects' perception – was borne out in the TWFs measured for normal hearing listeners at the highest click rate tested (800 Hz) for both ITD and ILD. The data also suggested reduced onset dominance for ILD relative to ITD – a finding in conflict with the notion of *equal* binaural adaptation for ITD and ILD (Haftner *et al.*, 1990), but consistent with more recent studies of onset dominance and the precedence effect (e.g., Saberi and

Antonio, 2003; Saberi *et al.*, 2004; Stecker and Brown, 2010). Nonetheless, since improved sensitivity to post-onset portions of the stimulus (i.e., reduced onset dominance) is the main consequence of binaural restarting, the restarting hypothesis (after Hafter and Buell, 1990) adopted by Laback and Majdak (2008) and Goupell *et al.* (2009) to explain their finding of improved ITD sensitivity for rapid AM stimuli clearly predicts that TWFs measured for brief *anisochronous* or “jittered” click trains should feature more comparable weights across the duration of the stimulus than isochronous click trains. That is, to the degree that onset weights are elevated relative to post-onset weights given isochronous stimulation, jitter-induced restarting should “flatten” TWFs for both ITD and ILD. Baseline differences in temporal weighting of the cues (i.e., less onset dominance/precedence/binaural adaptation for ILD) could temper the magnitude of difference between isochronous and jittered conditions, but more uniform temporal weighting should be expected for both ITD and ILD stimuli given jitter.

In the present study we thus evaluated the effect of binaurally synchronous temporal jitter on the time-course of normal-hearing listeners’ sensitivity to ITD and ILD by measuring TWFs for “jittered” high-rate click trains (see section 3.2.2). TWFs measured for jittered click trains were compared to TWFs for isochronous click trains from the previous experiment (Chapter 2). In an additional analysis, we evaluated the weights of ITD and ILD presented after the longest and shortest gaps in the jittered click trains. van Hoesel’s (2008a, 2008b) interpretation of the results of Laback and Majdak (2008) – that bilateral CI users’ improvement in ITD discrimination was attributable to occasional “slowing” of the effective pulse rate rather than binaural restarting (Hafter and Buell, 1990) – suggests that the effect of jitter might be directly observed as increased weighting of ITD presented following the longest ICI in the jittered stimuli. Alternatively, the hypothesis of restarting of binaural adaptation adopted by Laback and Majdak (2008) and Goupell *et al.* (2009) suggests, by comparison to the data of Hafter and colleagues (Hafter and Buell, 1990; Hafter *et al.*, 1990), that both the longest ICI (“gaps”) and shortest ICI (“squeezes”) should produce increased weights, with a similar effect for both ITD and ILD.

3.2 Experimental methods

The procedural methods employed in the current study were identical, except for differences in the stimuli presented, to those described in section 2.2. All procedures, including recruitment, consenting, and testing of human subjects followed the guidelines of the University of Washington Human Subjects Division and were reviewed and approved by the cognizant Institutional Review Board.

3.2.1 Subjects

Six normal-hearing subjects (0501, 0502, 0601, 0804, 0805, 0815) aged 21-36 participated in this experiment. The same six subjects participated in the previous experiment, from which the “Isochronous” data in the present report are taken. Subjects 0601, 0804, 0805 and 0815 were naive to the purposes of the experiment and were compensated for their participation. Subject 0501 was the author; 0502 was his advisor. All subjects reported normal hearing and demonstrated pure-tone detection thresholds < 20 dB HL over the range 250-8000 Hz.

3.2.2 Stimuli and procedure

Stimuli were trains of 16 Gabor clicks (Gaussian-windowed tone bursts). Each click consisted of a 4 kHz cosine multiplied by a Gaussian temporal envelope with $\sigma=221 \mu\text{s}$ ($367 \mu\text{s}$ duration at 3 dB below peak). The resulting spectral bandwidth was also Gaussian, with $\sigma=750 \text{ Hz}$ (-3 dB bandwidth = 1250 Hz). Clicks were synthesized at 48.848 kHz (Tucker-Davis Technologies RP2.1, Alachua FL) and presented using STAX model 4070 closed-back electrostatic headphones (STAX LTD., Saitama, Japan). Click trains were presented at approximately 70 dB SPL with mean peak-to-peak inter-click intervals (ICI) of 1.25 or 2.5 ms. The timing of individual clicks was determined using the temporal jitter procedure described by Goupell *et al.* (2009). For each trial, the ICI between successive clicks was drawn randomly from a uniform distribution centered at the nominal ICI (1.25 or 2.5 ms) with a width equal to $2k \times$ nominal ICI. The parameter k thus defined the degree of temporal jitter, with $k=0$ corresponding to temporal isochrony (no jitter) and $k=1$ to maximal jitter (individual ICI ranging

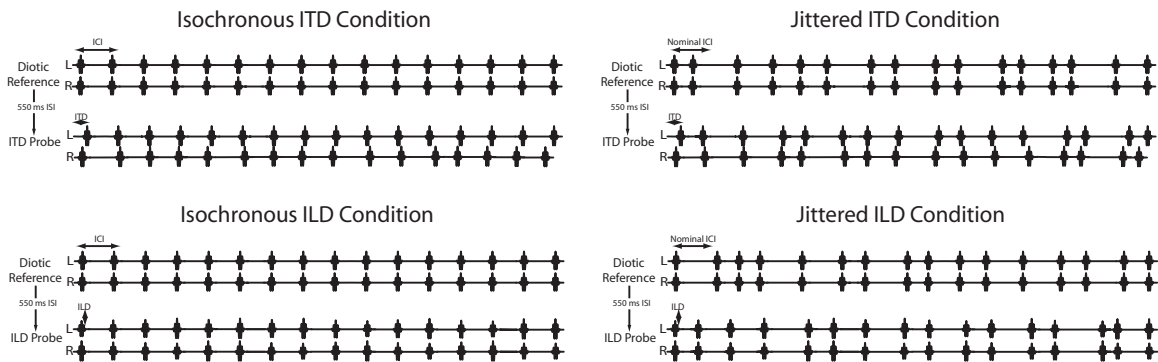


Figure 3.1: Schematic illustration of stimuli (not to scale). Each trial consisted of a diotic reference stimulus followed by a probe stimulus. In Isochronous (left) conditions (tested in a previous investigation), the reference stimulus was comprised of 16 equal-amplitude Gabor clicks presented synchronously to the left and right earphones. Following a 550 ms silent interval (ISI), and ITD or ILD probe stimulus was presented, comprised of 16 Gabor click pairs with random ITD or ILD imposed on each. In the ITD probe, each click pair carried an ITD drawn from a uniform distribution of $\pm 100 \mu\text{s}$ about a base ITD of -100, 0, or +100 μs (0 in the above illustration); clicks were presented at equal amplitude to the two earphones in this condition. In the ILD probe, each click pair carried an ILD drawn from a uniform distribution of $\pm 2 \text{ dB}$ about a base ILD of -1, 0, or +1 dB (0 in the above illustration); clicks were presented synchronously to the two earphones in this condition. The inter-click interval (ICI), which corresponded to the rate of click presentation, was held constant within and between trials of a single run. The Jittered conditions (right) were identical to the Isochronous conditions, with the exception that the ICI in both the reference and probe stimuli were varied randomly about the nominal ICI according to the parameter k (see text).

from 0 to $2 \times \text{ICI}$). In the current experiment, “Jittered” stimuli were generated with $k=0.9$, while “Isochronous” stimuli in the previous experiment were generated with $k=0$.

A button box (TDT RBOX) recorded subject responses. Testing was completed in a double-walled sound booth (IAC, Bronx, NY). Each trial consisted of two equal-duration click trains separated by a 550 ms silent interval. Click train duration was dependent on the ICI. The first train, presented as a reference, was always diotic; the probe train followed, presented with one of three different “base” interaural values. Additional *ITD* or *ILD* variation was imposed on each click to facilitate the computation of TWFs (see section 2.2.3). Note that such variation was independent of the binaurally synchronous temporal jitter imposed on the *ICI*. Separate conditions employed ITD and ILD cues (see Figure 3.1).

The base interaural value was varied pseudo-randomly from trial to trial. In the ITD conditions, the base values were -100, 0, or 100 μs ; in the ILD conditions, the base values were -1, 0, or 1 dB. By convention, negative ITD and ILD values led to or favored the left ear; positive ITD and ILD values led to or

avored the right. Additional random ITD or ILD variation, drawn from a uniform distribution of $\pm 100 \mu\text{s}$ or $\pm 2 \text{ dB}$, was added independently to each click in a train. These ranges were selected to correspond roughly to the $\pm 11^\circ$ azimuthal variation employed by Stecker and Hafter (2002), based on the physical correspondence between azimuth, ITD, and ILD described by Gulick *et al.* (1989), and an ITD/ILD trading ratio of approximately $50 \mu\text{s}/\text{dB}$ (Hafter and Jeffress, 1968; also see Chapter 2). Due to the per-click variation in ITD or ILD, trials with a base value of 0 carried nonzero interaural differences on each click, and thus a nonzero average ITD or ILD per trial.

Subjects were instructed to indicate via button press whether the probe stimulus was presented to the left or right of the reference stimulus. All subjects received training with visual feedback until performance stabilized. Feedback was then turned off and testing began. One hundred trials comprised a single run; in each run, 60 trials had 0 base interaural difference, 20 were left-leading and 20 were right-leading. Base values were presented in random order over the course of each run; rate and cue type were randomized and counterbalanced between runs. For the Jittered condition ($k=0.9$), each subject completed four runs at each combination of ICI (1.25 or 2.5 ms) and cue type (ITD or ILD), giving 1600 trials in total (400 trials \times 2 ICI \times 2 cues). Only 0 base trials were included in the TWF analysis (960 trials in total; 240 trials \times 2 ICI \times 2 cues). An additional and equivalent set of runs was completed without temporal jitter ($k=0$). Those data, reported in Chapter 2, comprise the “Isochronous” data included in the present report.

3.2.3 Statistical methods

TWFs were computed following the logic of previous observer-weighting approaches (e.g., Berg, 1989; Saberi, 1996; Stecker and Hafter, 2002). In the present investigation, weights were computed using a receiver operating characteristic (ROC) analysis to quantify the accuracy with which listener responses could be predicted on the basis of interaural differences applied on each trial. (For a detailed discussion of the application of ROC analysis for the

computation of TWFs, see section 2.4.3). Classification performance was quantified by the area under the ROC curve (AUC) obtained for classification of subjects' responses by several independently assessed classification variables, including the ITD or ILD applied to individual clicks (termed "AUC₁" for click #1, "AUC₂" for click #2, etc., plotted against click number to form the TWF itself), the mean ITD or ILD across clicks in a train ("AUC_{mean}," assessing the degree to which listener responses followed the mean ITD or ILD across clicks in the trains), and the ITD or ILD carried by the click pairs following the longest and shortest ICI of each trial ("AUC_{maxICI}" and "AUC_{minICI}," respectively). Since AUC values corresponded directly to the proportion of the subject's responses correctly classified by each classification variable, ranging from chance (AUC=0.5) (zero weight) to perfect classification (AUC=1), obtained AUC values can be construed as perceptual "weights," quantifying the influence of interaural differences carried by each classification variable on listeners' performance. Comparing AUC₁ against AUC_{mean} (i.e., AUC₁ – AUC_{mean}) thus provides a measure of onset dominance, since onset dominance entails weighting of onset cues over long-term interaural information (cf. Saberi and Perrott, 1995), where optimal performance in this task is obtained by a cue-averaging strategy that minimizes decision variance (Saberi, 1996). Values of AUC₁ – AUC_{mean} and other AUC measures were compared in several null hypothesis tests described in sections 3.3 and 3.4. Statistical confidence intervals on weight estimates were generated using 1000-repeat permutation tests to define the range of chance classification about AUC=0.5 (Figures 3.2 and 3.3; see Chapter 2).

3.3 Results

3.3.1 Temporal weighting functions for ITD and ILD

Figure 3.2 displays mean TWFs for the four Jittered conditions tested in the present investigation and four corresponding Isochronous conditions tested in Chapter 2 (see also Brown and Stecker, 2010). Filled circles in each panel plot the AUC obtained using each click ITD or ILD (AUC₁...AUC₁₆), dashed lines plot the AUC obtained using average ITD or ILD (AUC_{mean}), and triangles plot the

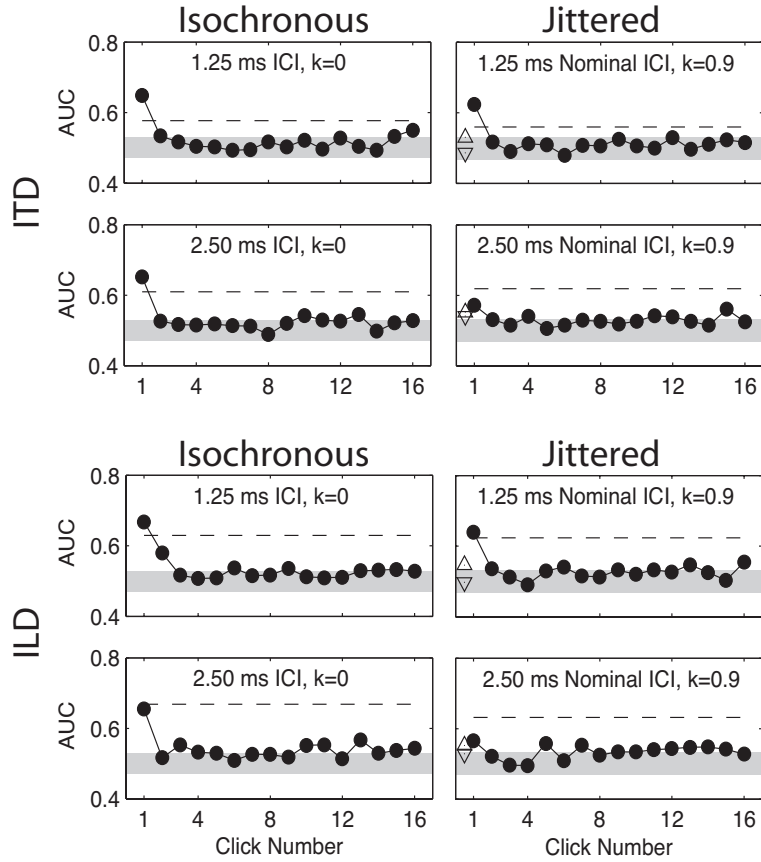


Figure 3.2: Temporal weighting functions (TWFs) averaged across subjects for Isochronous (left column) and Jittered (right column) conditions across inter-click interval (ICI) (rows) for both ITD (upper panels) and ILD (lower panels) stimuli. Within each panel, per-click ITD or ILD weights ($AUC_1 \dots AUC_{16}$, filled circles) are plotted against click number to form the TWF; the weight of mean ITD or ILD (AUC_{mean} , dashed line) and the weight of ITD or ILD following the longest (AUC_{maxICI} , upward triangle) and shortest (AUC_{minICI} , downward triangle) ICI per click train are plotted for comparison (see text). The shaded region plots the range of chance classification (see text).

AUC obtained using the ITD or ILD of clicks following the maximum and minimum ICI within each click train (AUC_{maxICI} , upward triangles; AUC_{minICI} , downward triangles; see sections 3.3.2 and 3.3.3). The shaded region plots the range of chance classification (95% CI of $AUC=0.5$), according to the distribution of AUC values obtained by classification of randomized responses (1000-repeat permutation test). Visually evident across both Isochronous (left column) and Jittered (right column) conditions is the decrease in AUC_1

relative to $AUC_2 \dots AUC_{16}$ and increase in AUC_{mean} (dashed line) at 2.5 ms ICI versus 1.25 ms ICI. Correspondingly, for both ITD (upper panels) and ILD (lower panels), the value of $AUC_1 - AUC_{\text{mean}}$ (i.e., the degree of onset dominance) is uniformly lower at 2.5 ms ICI than at 1.25 ms ICI. Of greater interest, the trend is more pronounced in the Jittered conditions; the value of AUC_1 falls below AUC_{mean} for both ITD and ILD at 2.5 ms ICI. Individual TWFs (Figure 3.3) support these observations, although individual variability is clear (for an extended discussion of individual differences in TWFs, see section 2.4 of Chapter 2).

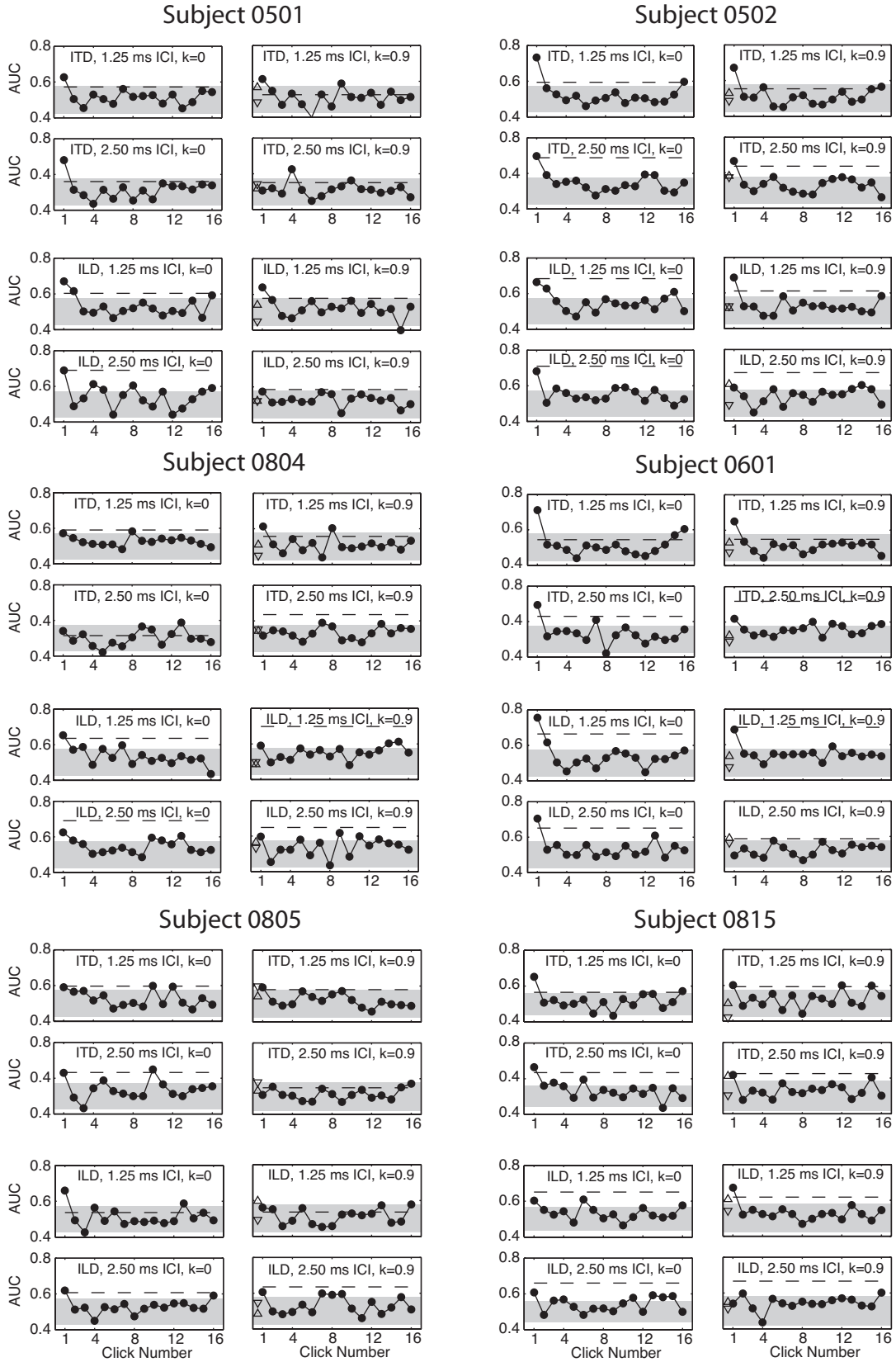


Figure 3.3: Individual subject TWFs for Jittered and Isochronous conditions. Legend as in Fig. 3.2.

Individual subject values of $AUC_1 - AUC_{\text{mean}}$ were submitted to a 2x2x2 repeated-measures ANOVA with factors of rate (1.25, 2.5 ms ICI), jitter (Isochronous, Jittered) and cue (ITD, ILD). The main effect of rate was significant ($F(1,5)=87.87, p<0.05$), as was the rate x jitter interaction ($F(1,5)=7.04, p<0.05$), while neither the main effect of jitter ($F(1,5)=5.56, p=0.065$) nor the main effect of cue ($F(1,5)=6.43, p=0.052$) were significant. Taken together, these results suggest a general reduction of onset dominance with increasing ICI, augmented by the effect of jitter at 2.5 ms ICI in particular. Follow-up pairwise *t*-tests verified that $AUC_1 - AUC_{\text{mean}}$ was significantly reduced by jitter for both ITD ($t(5)=3.21, p<0.025$) and ILD ($t(5)=5.29, p<0.025$) at 2.5 ms ICI, but for neither cue at 1.25 ms ICI (ITD: $t(5)=0.20, p=0.85$; ILD: $t(5)=0.46, p=0.66$). These results are summarized in Figure 3.4.

3.3.2 Longest versus shortest gaps

Under a strict interpretation of the “restarting” hypothesis adopted by Laback and Majdak (2008) and Goupell *et al.* (2009), both lengthened ICI (“gaps”) and shortened ICI (“squeezes”) should contribute comparably to ITD discrimination, with the result expected to extend to both ITD and ILD discrimination on the basis of the work of Hafter and colleagues (cf. Hafter and Buell, 1990; Hafter *et al.*, 1990). Thus, to test the effectiveness of ITD and ILD carried by clicks following gaps and squeezes, we computed AUC weights for ITD and ILD presented after the longest (AUC_{maxICI}) and shortest (AUC_{minICI}) ICI in each train (Jittered conditions only). The obtained values of AUC_{maxICI} and AUC_{minICI} are plotted in the right panels of Figs. 3.2 and 3.3 as upward and downward triangles, respectively. Both AUC_{maxICI} and AUC_{minICI} appear generally comparable to other individual post-onset weights (i.e., $AUC_2 \dots AUC_{16}$), while AUC_{maxICI} appears marginally higher than AUC_{minICI} . To test this difference statistically, the values of AUC_{maxICI} and AUC_{minICI} were compared in a 2x2x2 repeated-measures ANOVA with factors of measure ($AUC_{\text{maxICI}}, AUC_{\text{minICI}}$), rate (1.25, 2.5 ms ICI) and cue type (ITD, ILD). The analysis yielded a significant

main effect of measure ($F(1,5)=7.29$, $p<0.05$), but no significant main effect of cue ($F(1,5)=0.11$, $p=0.75$) nor measure by cue interaction ($F(1,5)=0.34$, $p=0.58$). The main effect of rate was measured at $F(1,5)=5.47$, $p=0.066$. Taken together, these results indicate that *both* ITD and ILD following the longest ICI exerted more influence on subjects' perception than ITD and ILD following the shortest ICI, although both AUC_{minICI} and AUC_{maxICI} were comparable to other post-onset weights and less than AUC_1 and AUC_{mean} across conditions. Further, the average value of both AUC_{maxICI} and AUC_{minICI} fell above the range of chance classification in the 2.5 ms ICI ITD condition – the condition of the present investigation most similar to “restarting” conditions measured by Hafter and Buell (1990) and Freyman *et al.* (1997) (see Fig. 3.2).

3.4 Discussion

3.4.1 The effect of jitter on TWFs for ITD and ILD

The present results indicate that imposing binaurally synchronous temporal jitter on high-rate (2.5 ms ICI) click trains reduces listeners' dependence on onset ITD and ILD cues for lateral discrimination. The effect was not apparent, however, at the highest rate tested (1.25 ms ICI), for which onset dominance was only marginally lower in Jittered than in Isochronous conditions (see Fig. 3.4). Reduced onset dominance for jittered stimuli, attributable to improved sensitivity to post-onset binaural information (i.e., reduced binaural adaptation), is consistent with the overall improvements in ITD discrimination with jittered stimuli described by Laback and Majdak (2008) and Goupell *et al.* (2009). The finding of greater onset dominance at 1.25 ms ICI than 2.5 ms ICI with or without jitter is further consistent with results of previous TWF investigations using filtered clicks (e.g., Saberi, 1996; Stecker and Hafter, 2002), noise bursts (Dizon *et al.*, 1998) and electrical impulses (van Hoesel, 2008a). It is interesting that the present investigation did not reveal an effect of jitter at 1.25 ms ICI (800 Hz), as Laback and Majdak (2008) and Goupell *et al.* (2009) demonstrated improved ITD discrimination with jittered stimuli at still higher rates. This discrepancy may be attributable to differences in the stimuli or performance measures employed

across studies. For example, the two previous studies employed trains of microsecond biphasic or monophasic pulses carrying fixed interaural cues, while the present investigation employed trains of millisecond Gabor clicks carrying variable interaural cues. Further, while the investigations of Laback and Majdak (2008) and Goupell *et al.* (2009) employed stimuli with gradual onsets so that subjects were forced to rely exclusively on post-onset information for discrimination judgments, the present study employed stimuli with strong onset cues. Finally, the present investigation did not explicitly measure listeners' ITD or ILD discrimination performance but rather estimated relative sensitivity to ITD or ILD over discrete portions of the stimuli.

3.4.2 Greater cue-averaging for ILD than ITD

Although binaural adaptation has been shown to occur for both ITD and ILD (Haftner and Dye, 1983; Haftner *et al.*, 1983) and has been attributed to common pre-binaural processing (Haftner *et al.*, 1990; cf. Goupell *et al.*, 2009), there are strong indications in the literature that the time course of processing of ITD and ILD diverges at one or more levels of the auditory system. For example, discrimination of ILD presented at signal offset appears to be relatively better than discrimination of ITD at signal offset (Stecker and Brown, 2010; see also Saberi *et al.*, 2004). Additionally, TWFs for lateralization suggest that onset dominance is reduced for ILD relative to ITD in both bilateral CI users (van Hoesel, 2008) and normal-hearing listeners (Chapter 2). Suggestive of further differences in the time courses of ITD and ILD processing, Krumbholz and Nobbe (2002) demonstrated a striking disparity between ITD- and ILD-based “echo thresholds” in “buildup” and “breakdown” aspects of the precedence effect, with a stronger buildup effect for click pairs carrying ITD than ILD, and a much stronger breakdown effect for click pairs carrying ILD than ITD (see Chapter 4).

In the present investigation (and the previous) we have operationally defined onset dominance as salience of onset binaural information over long-term average binaural information ($AUC_1 - AUC_{\text{mean}}$) after Saberi and Perrott (1995). As described above, jitter systematically lowered the value of AUC_1 for

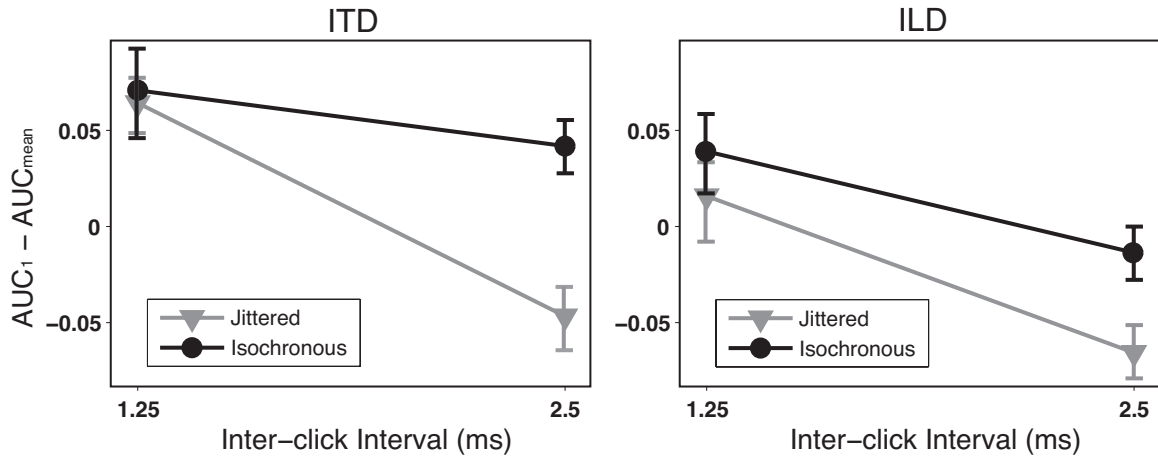


Figure 3.4: Cross-subject averages of $AUC_1 - AUC_{mean}$ against inter-click interval (ICI) for Isochronous (black) and Jittered (gray) conditions for both ITD (left panel) and ILD (right panel). Error bars indicate one standard error of the mean.

both ITD and ILD, with no significant difference in the effect between ITD and ILD conditions. However, cue-averaging alone, as measured by the value of AUC_{mean} – an explicit measure of the perceptual salience of mean ITD or ILD of all clicks in each train without respect to the weighting of individual clicks– appeared to be greater for ILD than ITD in all conditions (dashed lines, Fig. 3.2). To test this difference statistically, individual subject values of AUC_{mean} were submitted to a $2 \times 2 \times 2$ repeated-measures ANOVA with factors of jitter (Isochronous, Jittered), rate (1.25, 2.5 ms ICI), and cue type (ITD, ILD). The main effects of both rate ($F(1,5)=25.76$, $p<0.05$) and cue ($F(1,5)=33.38$, $p<0.05$) were significant, suggesting that listeners not only relied on cue-averaging to a greater degree at 2.5 ms than at 1.25 ms ICI (cf. Hafter and Dye, 1983; Hafter *et al.*, 1983), but also to a greater degree in ILD than ITD conditions (cf. Brown and Stecker, 2010; Stecker and Brown, 2010). Importantly, jitter neither affected the value of AUC_{mean} (main effect of jitter, $F(1,5)=0.58$, $p=0.48$), nor differentially affected AUC_{mean} for ITD versus ILD (jitter by cue interaction, $F(1,5)=0.36$, $p=0.58$).

3.4.3 “Restarting” versus listening after the longest gaps

Hafter and Buell’s (1990) original account of binaural restarting indicated that the effect could be induced by either lengthening or shortening the ICI (i.e., by insertion of both gaps and squeezes). This result was tentatively verified by Freyman *et al.* (1997), who demonstrated using two subjects that either

lengthening or shortening the inter-pulse interval (IPI) in an ongoing pulse train induced a shift in the perceived laterality of the stimulus to the side cued by the ongoing ITD, while perception was dominated by the opposing ITD cued at stimulus onset without such manipulations. Under the restarting hypothesis adopted by Laback and Majdak (2008) and Goupell *et al.* (2009), cues following both gaps and squeezes introduced by jitter should thus contribute to subjects' discrimination judgments, although this prediction is not perfectly clear for the highest rates tested in this and the previous investigations (Laback and Majdak, 2008; Goupell *et al.*, 2009): Hafter and Buell (1990) originally reported the restarting effect for 5 ms gaps inserted in 2.5 ms ICI click trains and for 2.5 ms squeezes inserted in 5 ms ICI Hz click trains. For the jittered 800 Hz (1.25 ms ICI) stimuli tested in the present investigation, the minimum possible ICI was $125 \mu\text{s}$ ($k=0.9$), while in the 1200 pps condition of Goupell *et al.* (2009), the minimum possible ICI was $0 \mu\text{s}$ ($k=1$). Thus, at very high rates, jittered stimuli are likely to include clicks which overlap partially or completely, such that, for acoustic stimulation, "squeezes" *must* be less effective or completely ineffective compared to "gaps." This consideration does not necessarily apply to electrical stimulation (Laback and Majdak, 2008), however, and benefits of jitter for acoustic stimulation have also been demonstrated at relatively low values of k (e.g., $k=1/3$, Goupell *et al.*, 2009).

Our test of AUC_{maxICI} versus AUC_{minICI} demonstrated that ITD and ILD following "gaps" were in fact more salient than ITD and ILD following "squeezes." This result is consistent with TWFs measured for free field sound localization by Stecker (2000; also see Stecker and Hafter 2002), who reported restarting (as evidenced by increased weight following altered ICI) following 5 ms gaps, but not 2 ms squeezes in 3 ms ICI click trains. Although the finding is also roughly consistent with the suggestion of van Hoesel (2008a, 2008b) that jitter enhances sensitivity to ITD following long intervals specifically, we note that the difference between AUC_{maxICI} and AUC_{minICI} was not statistically different for stimuli carrying ITD and stimuli carrying ILD across all conditions. The result is therefore difficult to understand in terms of reduced temporal ambiguity of ITD (cf. Freyman *et al.*,

1997 and van Hoesel, 2008a, 2008b), because temporal ambiguity *per se* (i.e., multiple ITD cued by “slipped cycles” in a periodic click train with short ICI) should not influence sensitivity to ILD (cf. Hartmann and Constan, 2002). Thus, it may be more appropriate to consider the effect of temporal jitter on peripheral filtering (Stecker and Brown, 2009) or neural refractoriness (see Goupell *et al.*, 2009; van Hoesel, 2008b). Although the physiological bases of the “restarting” effect remain uncertain, Goupell *et al.* (2009) modeled the effects of temporal jitter on the firing patterns of auditory nerve fibers, and showed that cross-correlation of responses to “jittered” stimuli led to better representation of stimulus ITD than did cross-correlation of responses to isochronous stimuli. While this account does not explicitly address the complementary question of ILD coding and is otherwise rather tentative, the fact that it does not require an active restarting mechanism suggests that at least some aspects of the effect of “temporal jitter” (and binaural adaptation/onset dominance/precedence itself) may reflect fairly low-level mechanisms such as peripheral filtering and adaptation in primary auditory neurons (see also Hartung and Trahiotis, 2001; Bernstein and Trahiotis, 2002). The influence of peripheral auditory processing (including neural adaptation) on the temporal dynamics of binaural sensitivity for both ITD and ILD is treated in detail in Chapter 5.

Concerning the theoretical motivation for the present investigation, our data suggest, in the context of the precedence effect, that temporal irregularity of “lagging” information may reduce the perceptual dominance of “leading” information (see section 3.1). These data may thus be taken to suggest that temporally isochronous stimuli overestimate the typical magnitude of the precedence effect. Experiments using room simulations, which yield more realistic, temporally irregular reflections, have become more common over the past decade (e.g., Djelani and Blauert, 2001; Brandewie and Zahorik, 2010); it will be instructive to compare data from such experiments to data from psychophysical studies using more traditional stimuli. Finally, toward the translational motivation for the present experiment, our data suggested no *drastic* effect of binaurally synchronous temporal jitter on listeners’ ILD sensitivity.

Listeners appeared to retain sensitivity to average stimulus ILD (AUC_{mean}). Obviously, future work aiming to assess the clinical feasibility of using temporal jitter to improve binaural outcomes in bilateral CI users should further examine the effects of temporal jitter on both ITD and ILD sensitivity in CI users directly.

3.5 Summary and conclusions

- 1) TWFs for discrimination of ITD and ILD carried by isochronous high-rate click trains demonstrated greater weighting of cues applied to onset clicks than to individual post-onset clicks, consistent with previous studies demonstrating dominance of onset cues over ongoing cues in the localization of sounds with regular, rapid modulation.
- 2) The introduction of binaurally synchronous random temporal jitter in the envelopes of high-rate modulated stimuli reduced the dependence of listeners' discrimination judgments on onset ITD and ILD. This finding is consistent with the view that improved ITD discrimination performance evidenced in previous investigations using jittered stimuli might be attributed to improved sensitivity to post-onset information (reduced binaural adaptation). Importantly, these data suggest a weaker precedence effect given temporally irregular "lagging" information.
- 3) Cue-averaging (quantified by weight applied to the mean ITD or ILD of the complete click train) was found to be greater for ILD than for ITD stimuli, consistent with a greater role for temporal integration in ILD than ITD processing. This difference was not affected by the introduction of temporal jitter.
- 4) Thus, improved discrimination with jitter is not likely attributable to better cue-averaging over the duration of the stimulus or to reduced stimulus "ambiguity"; rather, the combinatorial effect of multiple long and short ICI may enhance the neural representation of the stimulus envelope (e.g., by increasing the effective modulation depth) to facilitate more robust encoding of interaural differences by central mechanisms (Goupell *et al.* 2009; Stecker and Brown, 2009; see Chapter 5).

4

The precedence effect in sound localization: Fusion and lateralization measures for pairs and trains of lead-lag stimuli lateralized by interaural time and level differences

In 1987, Clifton offered a most surprising addition to the precedence effect literature by demonstrating that the range of lead-lag fusion (i.e., the echo threshold) was dynamic and dependent on prior stimulation. Using click stimuli in a free field lead-lag paradigm (after Wallach *et al.*, 1949), Clifton (1987) showed that listeners' echo thresholds could be elevated by repeated presentation of a particular lead-lag click pair ("buildup"), similar to a result initially described in brief by Thurlow and Parks (1961), but returned to baseline by presentation of a binaurally reversed lead-lag pair ("breakdown," achieved by switching the lead and lag speakers to create an "echo-source" stimulus). Multiple studies of this phenomenon over the past two decades (e.g., Clifton and Freyman, 1989; Clifton *et al.*, 1994; Grantham, 1996; McCall *et al.*, 1998; Djelani and Blauert, 2001) have established that a listener's echo threshold for lead-lag pairs of clicks or other impulsive stimuli, generally on the order of 5-10 ms at baseline, may be elevated to 15-25 ms following buildup, and reduced back to 5-10 ms following breakdown. These observations have been taken to suggest a role for cognition in sound localization, centered around the notion that acquired "knowledge" of the natural order of sources and echoes in a particular listening environment can modulate sensitivity to reflected sound (see Blauert, 1997). A dynamic precedence effect should enable listeners to localize adaptively in a continuously changing listening environment. Recent investigations using virtual auditory space and free field stimuli have pursued this notion further, demonstrating that multiple lead-lag stimuli (carrying different sets of binaural cues) can be "built-up" concurrently (Djelani and Blauert, 2001; Freyman and Keen, 2006; Keen and Freyman, 2009). That is, normal hearing listeners may establish built-up echo thresholds for multiple sources at once – doubtless an invaluable faculty in complex sound fields comprised of multiple sources and their echoes.

Interestingly, the finding that built-up echo thresholds are apparently maintained for a lead-lag stimulus of a particular binaural configuration after presentation of an intervening stimulus of a different binaural configuration (Djelani and Blauert, 2001) has been taken to suggest that previous measurements of echo thresholds in breakdown paradigms simply reflected baseline echo thresholds for novel stimuli (rather than “reset” thresholds for the original stimuli). Findings concerning the contributions of individual spatial cues (ITD and ILD) to precedence phenomena, however, complicate this interpretation.

4.1 Different contributions of ITD and ILD to precedence

The breakdown of the precedence effect is rather a non-intuitive result considering the proposed function of the precedence effect (i.e., to facilitate the localization of direct sound by suppressing the localization of echoes), because it suggests that listeners must be capable of detecting a change in an echo that is not consciously perceived – perhaps the reason the effect went undiscovered for nearly 40 years following the seminal paper of Wallach *et al.* (1949). The simplifying interpretation that breakdown is merely the absence of buildup for a novel stimulus (Djelani and Blauert, 2001), although parsimonious, neglects the fact that breakdown occurs not only for a switching of the lead and lag sources, but also for a simple adjustment of the lead-lag delay (Clifton *et al.*, 1994) or a shift in the spectrum of the echo (McCall *et al.*, 1998). That is, breakdown is not readily explainable as the introduction of a new “not-yet-built-up” lead source, because it also occurs when the “source” is fixed and a change is made only to the cues carried by the “echo.” A study by Krumbholz and Nobbe (2002) may offer valuable insight on the subject: Krumbholz and Nobbe (2002) presented listeners with lead-lag pairs of clicks over headphones to measure the buildup and breakdown of precedence for stimuli lateralized by either ITD or ILD. This experiment was a departure from most other studies of precedence, which have almost exclusively employed headphone ITD or free field (or virtual auditory space) stimuli, where the contributions of ITD and ILD to listeners’ judgments were not independently evaluated (e.g., Wallach *et al.*, 1949; Thurlow and Parks,

1961; Clifton 1987; Shinn-Cunningham *et al.*, 1993; Yang and Grantham, 1997; Djelani and Blauert, 2001; Keen and Freyman, 2009; see Litovsky *et al.*, 1999). Krumbholz and Nobbe (2002) measured significant buildup of precedence for ITD stimuli, with echo thresholds increasing from a mean of 7 ms at baseline to 16 ms after stimulus repetition, and lesser but significant buildup for ILD stimuli, with echo thresholds increasing from a mean of 5 ms at baseline to 11 ms after stimulus repetition. The *breakdown* of precedence, in contrast, was evidenced *only* for ILD stimuli. Echo thresholds for an ILD lead-lag test stimulus carrying lead and lag cues opposite those in 12 preceding pairs were comparable to baseline ILD thresholds (mean of ~4 ms), while echo thresholds in a corresponding ITD condition were comparable to those in the buildup ITD condition (mean of ~14 ms). That is, presentation of a lead ITD stimulus from a new location produced no breakdown of echo suppression. The difference between breakdown ITD and ILD conditions was statistically significant, and robust across listeners. The findings of Krumbholz and Nobbe (2002) are thus striking in the context of previous studies of buildup and breakdown (e.g., Clifton, 1987; Djelani and Blauert, 2001; Keen and Freyman, 2009): based on their data, breakdown in the free field (or even baseline sensitivity to a not-yet-built-up signal), would seem necessarily mediated by sensitivity to changes in the ILD, while the concomitant “switching” of the ITD may be inconsequential.

A number of questions about the precedence effect might be asked on the basis of this result. For example, Krumbholz and Nobbe (2002) did not include a “re-buildup” condition (after Djelani and Blauert, 2001) in their experiment to assess whether echo threshold had actually “broken down” for ILD, or whether thresholds reflected, as now widely assumed, baseline thresholds for novel stimuli. Of greater interest, while it is frequently assumed (e.g., Getzmann, 2004; Phillips and Hall, 2005; Maier *et al.*, 2010) that ITD and ILD are combined at some level of the brain into a common “spatial” code, the data of Krumbholz and Nobbe (2002; see their Discussion in particular) suggest that precedence effects for ITD and ILD are to some degree independent. While it is difficult to rigorously test for the existence or nonexistence of an internal cross-cue representation of

“auditory space” within a purely psychophysical study of the precedence effect, the view that dynamic precedence reflects listeners’ immediate experience with the spatial configuration of sources and echoes in a given listening environment holds that the effects of stimulus repetition should depend only on the *spatial* perception they induce, and not on the specific acoustic cues they carry, i.e., that the precedence effect should be “space-specific,” *not* “cue-specific.” Assessment of the foregoing ideas requires measurement of both echo thresholds and the spatial perception induced by stimuli carrying ITD or ILD presented in isolation or following appropriately designed conditioner stimuli. The present chapter describes several experiments designed for this purpose. Data are presented and briefly discussed for each experiment individually, followed by general discussion in the context of the precedence literature at large.

4.2 Common experimental methods

All procedures, including recruitment, consenting, and testing of subjects followed the guidelines of the University of Washington Human Subjects Division and were reviewed and approved by the cognizant Institutional Review Board.

4.2.1 Subjects

Ten subjects (0601, 0915, 1002, 1004, 1007, 1012, 1013, 1014, 1104, 1106) aged 20-58 (4 female) completed participation in this experiment. Partial data (not reported here) were collected for an additional 4 subjects. All subjects were naïve to the purpose of the experiment and were compensated for their participation. Subjects reported normal hearing and demonstrated thresholds <20 dB HL bilaterally with <10 dB asymmetry at octave frequencies 250-8000 Hz.

4.2.2 Procedure

All testing was completed in a sound-attenuated booth (IAC, Bronx, NY). Subjects were seated in a swivel chair facing a large (80-cm diagonal) touch-sensitive display (elo Touchsystems 3200L, Tyco Electronics, Bermuda). Stimuli across all phases of all experiments were comprised of monophasic pulses

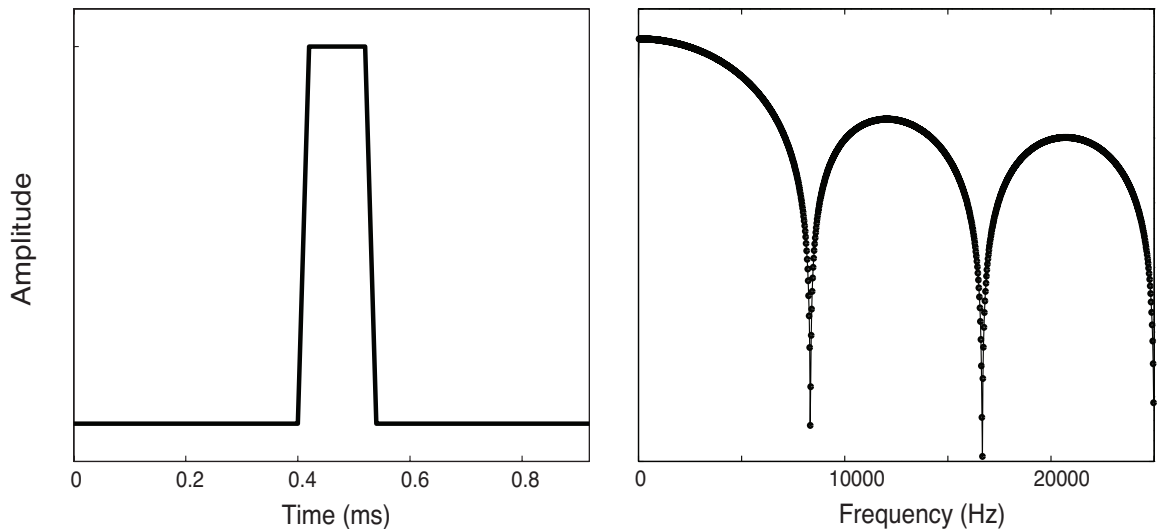


Figure 4.1: Time (left) and frequency domain (right) representations of the 120 μ s impulse used throughout experiments described in the present chapter. Impulse amplitude was set at \sim 60 dB SPL.

(“clicks”) 120 μ s in duration with an average binaural level of \sim 60 dB SPL (Figure 4.1 illustrates a single such click in time and frequency domains). Pairs and trains of clicks (to be detailed subsequently for each experiment) were programmed in MATLAB (MathWorks, Natick, NJ), synthesized at 48.848 kHz (Tucker-Davis Technologies RP2.1, Alachua FL) and presented under closed-back electrostatic headphones (STAX Model 4070, Saitama, Japan). Non-zero ITD were imposed by delaying the signal to the left channel for right-favoring ITD or delaying the signal to the right channel for left-favoring ITD. Non-zero ILD were imposed by amplifying the signal to the favored earphone by half the total ILD and attenuating the signal to the opposite earphone by an equal amount. Subjects completed each experiment by interacting with the touch-sensitive display as described in the following sections.

4.2.2.1 Training

Prior to participation in Experiment I (section 4.3), subjects were trained in a simple two-alternative-forced-choice task. The purpose of this training was (1) to familiarize subjects with the “lead-lag” experimental stimuli, and (2) to provide a standard for judgments in the inherently subjective experimental task. Specifically, stimuli in the training task were pairs of lead-lag dichotic clicks (separated by 1-50 ms) in which the lag carried ITD or ILD either identical to or

opposite the ITD or ILD carried by the lead. On each trial, the subjects' task was to indicate whether the stimuli had consisted of signals from "one location" (true if lead and lag cues were identical, indicated by touching the panel at the top of display) or "two locations" (true if lead and lag cues favored opposite ears, indicated by touching the panel at the bottom of the display). Immediate correct/incorrect feedback was displayed on the monitor after each trial. In addition to verifying that subjects were sensitive to the nonzero ITD or ILD carried by the stimuli, this task served to train subjects to touch the "two locations" panel only when two discrete *locations* were perceived, rather than simply when two sounds were perceived. This method emphasizes a more stringent definition of echo threshold (summarized by Blauert, 1997), which requires that the lead and the lag be perceived at discrete locations (though not necessarily their veridical locations). All subjects completed at least two hours of training runs, evenly divided between ITD stimuli and ILD stimuli. Several subjects (including the four who did not complete testing) required additional training to reach criterion (90 percent correct or better at lead-lag delays >25 ms) for either ITD or ILD conditions (or both), suggestive of lesser sensitivity to the cues in those subjects.¹

4.2.2.2 ITD-ILD matching task

The questions addressed by Experiments II and III (see sections 4.4 and 4.5) required that the ITD and ILD employed in test stimuli produced *equivalent* lateralization for each subject (approximate equivalence was assumed for Experiment I on the basis of Krumbholz and Nobbe, 2002). Thus, prior to testing in Experiments II and III, an explicit ITD-ILD matching task was completed by each subject. On each trial of the matching task, subjects were presented with a "standard" click carrying +/-300 μ s ITD (Experiment II matching task) or +/-600 μ s ITD (Experiment III matching task), followed 1 s later by a second "pointer" click carrying a random ILD from the range -15 to +15 dB ILD. By convention, positive cue values favor the right ear. On each trial, subjects were instructed to adjust the perceived location of the ILD pointer to match the perceived location the ITD

standard by using buttons displayed on the touch screen monitor. Three right arrow buttons and 3 left arrow buttons provided for “coarse” ± 3 dB), “medium” (± 1 dB), and “fine” (± 0.3 dB) adjustment. After each adjustment, the standard-pointer click pair was automatically replayed. Subjects were free to make as many adjustments as necessary to match the location of the pointer to the location of the standard. Once subjects were satisfied with the pointer-standard match, they pressed a “match” button at the bottom of the display, and the next trial began automatically. Each run consisted of 20 trials; subjects completed 4 runs for left- and right- leading ITD standards (8 total) for each matching task. Matching data for Experiments II and III are presented in sections 4.4 and 4.5.

4.2.2.3 Main experimental task

Following completion of training (Experiment I) or ITD-ILD matching tasks (Experiments II and III), subjects began testing. The subject’s task on each trial was the same across all conditions and all experiments. Following presentation of either a brief silent period or a conditioner stimulus (see sections 4.3-4.5), a lead-lag test stimulus was presented. The subject was instructed to indicate for this stimulus (i.e., ignoring the conditioner stimulus, if present) (1) whether one location (upper panel on the touch screen) or two locations (lower panel on the touch screen) had been perceived and (2) to indicate the apparent lateral location of the stimulus within the selected panel (a perceptual scaling task). If two locations were perceived, subjects were further instructed to indicate the *left-most* location perceived (see Figure 4.2). Each response thus carried two independent components of data. Specifically, *which* panel the subject touched indicated whether the lead and lag clicks had appeared “fused,” and *where* the subject touched within the selected panel indicated the extent of localization dominance (i.e., the extent to which the reported percept agreed with the cues carried by the lead versus cues carried by the lag). Both aspects (fusion and localization dominance²) were expected to change as a function of lead-lag delay, the parameter varied from trial-to-trial (see below). Trials within a given run were of a single stimulus condition (e.g., Baseline ITD; see sections 4.3-4.5);

stimulus conditions were presented in random order across subjects. Subjects completed at least one practice run in each condition before testing commenced.

Both adaptive methods (e.g. Krumbholz and Nobbe, 2002) and methods of constant stimuli (e.g., Clifton and Freyman, 1989) have been used for echo threshold estimation in studies of the precedence effect. Each has its advantages (e.g., the efficiency of an adaptive staircase versus the completeness of an empirically measured psychometric function). In efforts to ensure consistency of

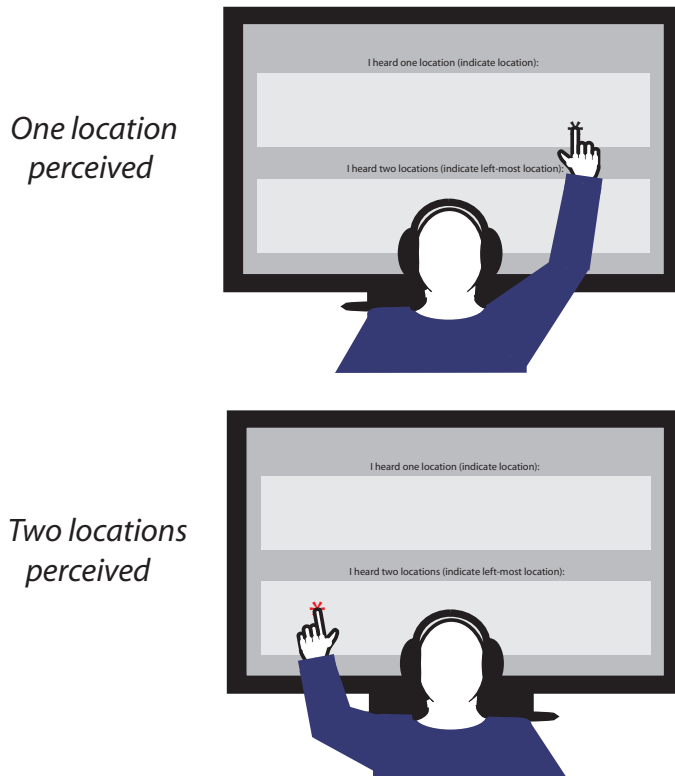


Figure 4.2: Main experimental task. On each trial, the subject was presented with a lead-lag test stimulus and instructed to simultaneously indicate (1) how many sound locations were perceived, done by selecting the upper (“I heard one location”) or lower (“I heard two locations”) panel, and (2) where the sound was perceived, done by indicating within the selected panel the apparent location of the test stimulus. In the event that two locations were perceived, the subject was instructed to always indicate the leftmost location perceived.

measurement, echo thresholds for each run in the present investigation were estimated using three independent procedures (programmed in MATLAB). On each trial within a given run, the lead-lag delay was drawn randomly from one of two adaptive tracks (one ascending, starting value 1 ms; one descending, starting value 50 ms) or from a constant set of delays ([0.42³ 3 6 9 15 25 35 50 ms], 5 trials per stimulus, each equally probable on a given trial). Each adaptive tracker followed a 1-up, 1-down rule to estimate the 50% echo threshold; logarithmic step sizes of 0.2 ($\text{delay}_{\text{new}} = \text{delay}_{\text{old}} \times 10^{\pm 0.2}$) were employed up to the fourth reversal in each track, then decreased to 0.05 ($\text{delay}_{\text{new}} = \text{delay}_{\text{old}} \times 10^{\pm 0.05}$) for the duration of the run. Each track terminated after 8 reversals. The threshold for each was taken as the geometric mean ITD of the final 4 reversals. A third

of measurement, echo thresholds for each run in the present investigation were estimated using three independent procedures (programmed in MATLAB). On each trial within a given run, the lead-lag delay was drawn randomly from one of two adaptive tracks (one ascending, starting value 1 ms; one descending, starting value 50 ms) or from a constant set of delays ([0.42³ 3 6 9 15 25 35 50 ms], 5 trials per stimulus, each equally probable on a given trial). Each adaptive tracker followed a 1-up, 1-down rule to estimate the 50% echo threshold; logarithmic step

threshold was taken as the lead-lag delay at the interpolated 50% point on a psychometric function fit to responses for the constant set of values once the run had finished (using the custom MATLAB function `psignfit`, Wichmann and Hill, 2001). Subjects completed 4 runs in each condition in Experiments I and II, giving 12 total threshold estimates per condition per subject; in Experiment III, subjects completed 2 runs in each condition, giving 6 threshold estimates per condition. Since 2 out of 3 threshold trackers were adaptive, the lead-lag delay presented on a given trial was occasionally related to the lead-lag delay presented on the previous trial; nonetheless, tracker randomization and the presence of a constant stimulus tracker made it nearly impossible to anticipate the trial-to-trial delay. Finally, lateralization responses were independent of echo threshold measurements and did not affect the progression of experimental runs; lateralization data were analyzed offline after completion of the experiment.

4.2.3 Analysis

Echo thresholds were compared across conditions and across subjects by repeated-measures ANOVA and paired or one-sample *t*-tests. Significant differences are given by $p < 0.05$, with corrections for multiple comparisons applied as appropriate. Substantial individual variability in echo thresholds was anticipated on the basis of past studies of the precedence effect and binaural sensitivity in general (e.g., Wallach *et al.*, 1949; McFadden *et al.*, 1973; Freyman *et al.*, 1991; Yang and Grantham, 1997; Chapters 2 and 3); such variability was observed in some conditions of the present investigation. Individual subject data are thus given for each experiment along with mean data. Lateralization data consisted of horizontal position values giving the location within either panel (“one location” or “two locations”) that the subject touched on each trial, ranging from -1 (extreme left) to +1 (extreme right). For all experiments, data were grouped into responses for “one location” trials (i.e., trials on which the subject touched the upper panel), plotted in black, and “two location” trials (i.e., trials on which the subject touched the lower panel), plotted in red. Responses were plotted as a function of lead-lag delay, and weighted lines of best fit were

generated to summarize trends in lateralization across delay (dashed black and red lines in Figures 2.6, 2.10, and 2.14). In some cases, one-sample *t*-tests were conducted on the slopes of these lines (described separately for each experiment).

4.3 Experiment I: Dynamic precedence effects for ITD and ILD

The goal of this experiment was to replicate and extend the study of Krumbholz and Nobbe (2002), in which echo thresholds were measured for pairs of lead-lag clicks lateralized by ITD or ILD across three different stimulus conditions: Baseline, Buildup, and Breakdown. Stimuli in the present experiment were of four different types: (1) *Baseline* stimuli consisted of a single lead-lag click pair, (2) *Buildup* stimuli consisted of 12 “conditioner” lead-lag click pairs and a final test pair identical to the conditioner pairs, (3) *Breakdown* stimuli consisted of 12 conditioner pairs and a “switched” test pair in which the interaural cues were swapped between the lead and lag clicks, and 4) *Retest* stimuli consisted of 11 conditioner pairs, an intervening switched pair, and a final test pair identical to the 11 conditioner pairs (after the condition used by Djelani and Blauert, 2001 to demonstrate maintenance of buildup following “breakdown”). In ITD conditions, lead clicks always carried $\pm 300 \mu\text{s}$ ITD (i.e., $300 \mu\text{s}$ right-favoring or $300 \mu\text{s}$ left-favoring ITD) and 0 dB ILD, and lag clicks always carried $\mp 300 \mu\text{s}$ ITD and 0 dB ILD (i.e., an opposing ITD cue). Correspondingly, in ILD conditions, lead clicks always carried ± 10 dB ILD and 0 μs ITD, and lag clicks always carried ∓ 10 dB ILD and 0 μs ITD. These and other key stimulus parameters are illustrated in Figure 4.3. Stimuli were designed to match those employed by Krumbholz and Nobbe (2002) in their ITD and ILD conditions; cue values were expected to produce approximately equivalent lateralization for ITD and ILD stimuli (see also Figure 4.7). The major novelty of the present experiment was its simultaneous assessment of both fusion and localization dominance, the latter having never been measured for “buildup” stimuli (cf. Clifton, 1987; Freyman *et al.*, 1991; Clifton *et al.*, 1994; Djelani and Blauert, 2001; Freyman and Keen, 2006; Keen and Freyman, 2009).

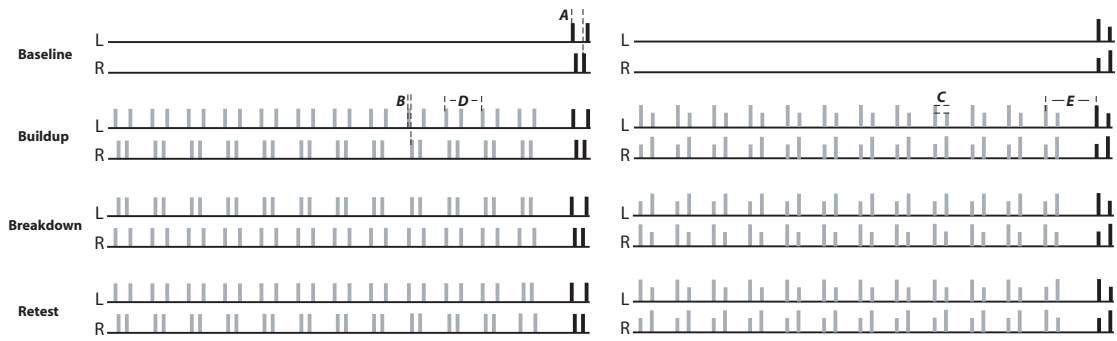


Figure 4.3: Schematic illustration of stimuli for Experiment I. Test lead-lag click pairs (bold ticks) were preceded by silence (Baseline), 12 identical lead-lag click pairs (Buildup), 12 binaurally opposite lead-lag click pairs (Breakdown), or 11 identical lead-lag click pairs and one intervening opposite pair (Retest). Stimuli in the left column were for ITD conditions; stimuli in the right column were for ILD conditions. A, lead-lag delay (varied trial-to-trial, see text); B, 300 μ s ITD; C, 10 dB ILD; D, 250 ms inter-stimulus interval between conditioner lead-lag pairs; E, 500 ms pause between conditioner and test.

4.3.1 Results

4.3.1.1 Echo thresholds

Figure 4.4 gives mean echo thresholds for ITD (dark grey) and ILD (light grey) conditions averaged across subjects (12 threshold estimates per subject). Mean echo thresholds were greater for ITD than ILD in every condition, with the disparity particularly evident in the Breakdown condition. Of interest in the context of the studies of Djelani and Blauert (2001) and Krumbholz and Nobbe (2002), the echo threshold in the ILD Retest condition appeared to be comparable to the echo threshold in the ILD Buildup condition. Thus, lower Breakdown ILD thresholds notwithstanding, buildup was apparently maintained for the original lead-lag ILD stimulus. Individual data, plotted in Figure 4.5, support the mean trends with some individual differences evident (e.g., subjects 0601 and 1014 failing to show any breakdown effect in either ITD or ILD conditions). Mean data were submitted to a 4x2 (condition x cue) repeated-measures ANOVA. The main effects of cue ($F(1,9)=27.51$, $p<0.05$) and condition ($F(3,27)=16.46$, $p<0.05$), and the cue x condition interaction ($F(3,27)=4.31$, $p<0.05$), were all significant. Follow-up paired t -tests with set-wise correction for multiple comparisons demonstrated that ITD-based echo thresholds were significantly higher than ILD-based echo thresholds in Baseline ($t(9)=5.31$, $p<0.0125$), Buildup ($t(9)=3.16$, $p<0.0125$) and Breakdown ($t(9)=4.09$, $p<0.0125$) conditions, but not in the Retest condition ($t(9)=2.57$, $p=0.030$). An additional set

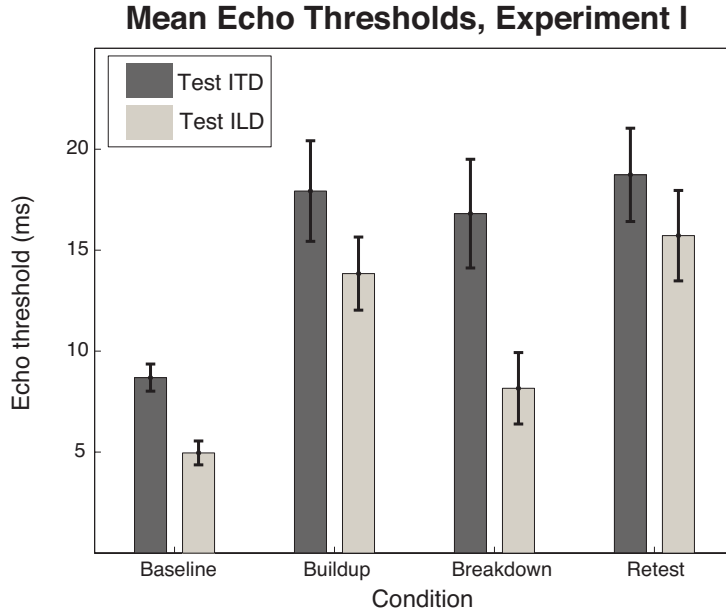


Figure 4.4: Mean echo thresholds for ITD and ILD conditions of Experiment I. Error bars give the standard error of the mean across 10 subjects.

of tests demonstrated that ILD-based echo thresholds were significantly higher in Buildup than Baseline ($t(9)=4.78, p<0.025$) and Breakdown ($t(9)=2.76, p<0.025$) conditions. A final set of tests demonstrated that echo thresholds within ITD conditions were significantly higher than Baseline in Buildup ($t(9)=4.15, p<0.0125$), Breakdown ($t(9)=3.39, p<0.0125$), and Retest ($t(9)=4.52, p<0.0125$) conditions, while Breakdown and Buildup thresholds were not significantly different ($t(9)=1.86, p=0.100$).

4.3.1.2 Lateralization responses

Figure 4.6 displays lateralization responses for all conditions of the present experiment (ITD, upper panels; ILD, lower panels). Responses for trials

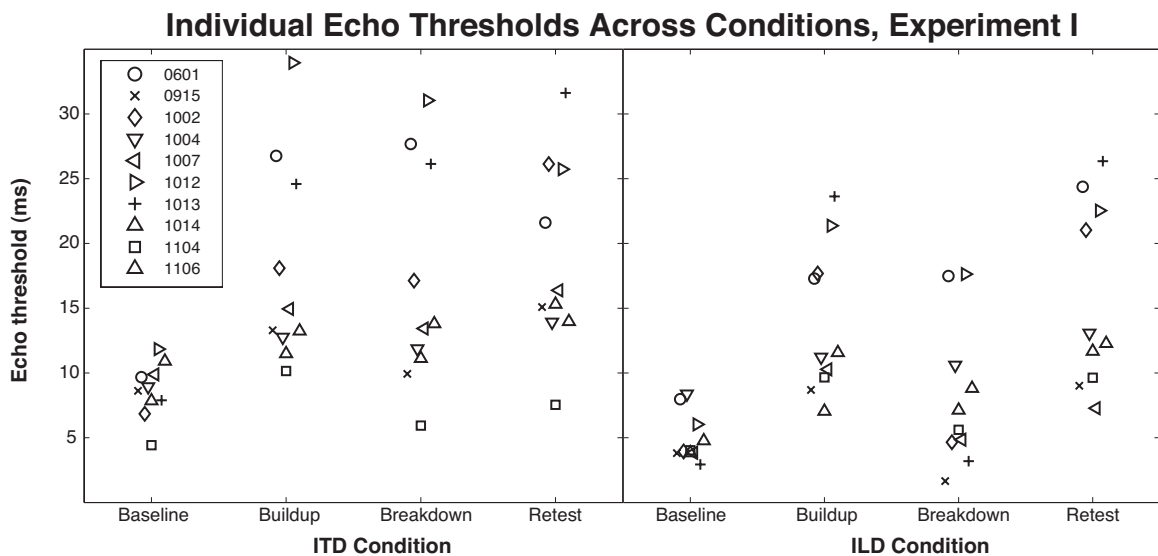


Figure 4.5: Individual subject echo thresholds for ITD (left) and ILD (right) conditions of Experiment I.

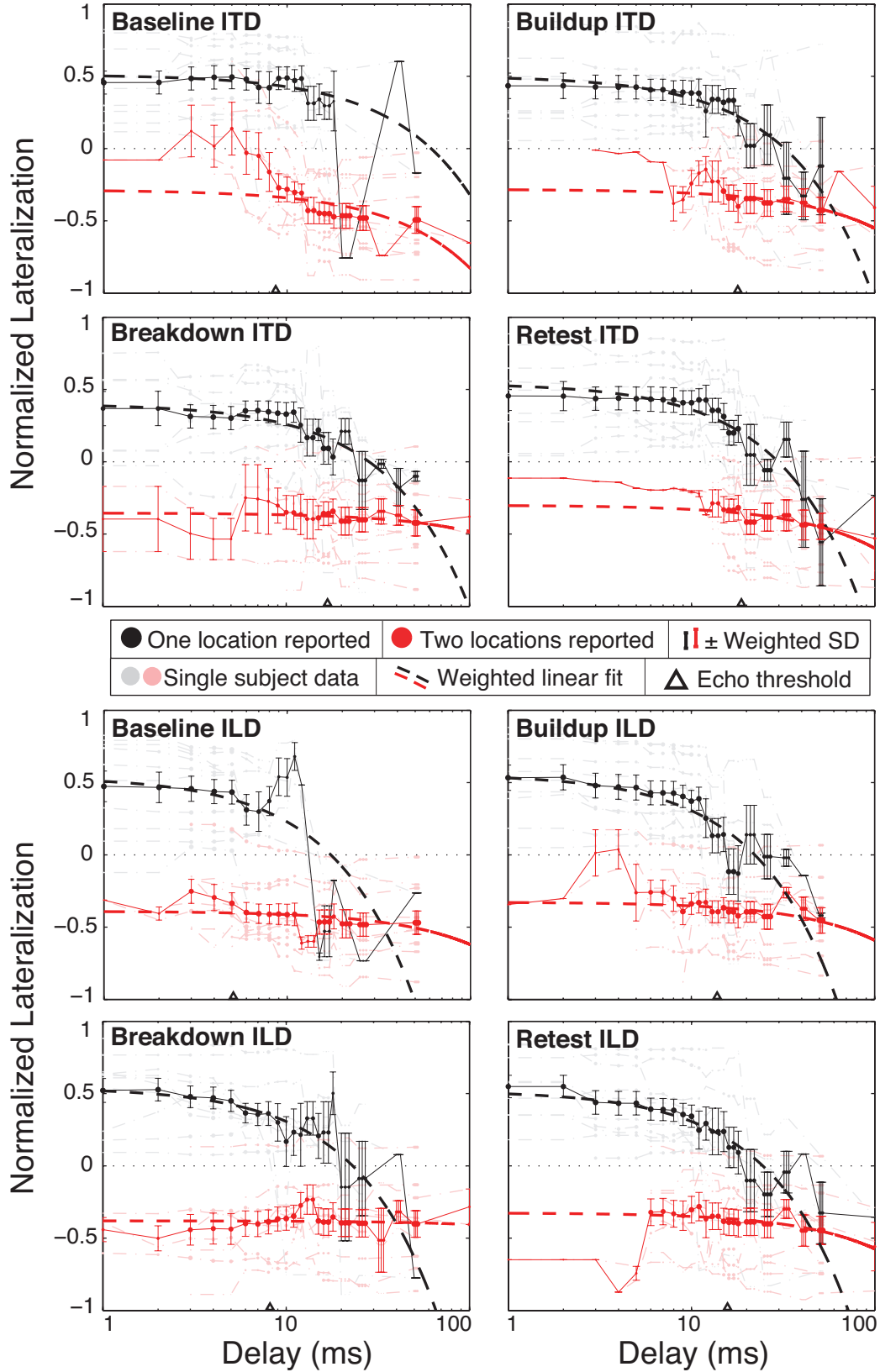


Figure 4.6: Lateralization data for Experiment I. For detailed explanation of legend, see section 4.2.3. Echo thresholds were measured independently and are indicated for visual comparison only.

on which subjects reported “one location” are plotted in black; responses for “two locations” trials are plotted in red. Since the three different threshold estimation procedures introduced dozens of unique lead-lag delays across trials, responses were plotted as the mean for a sliding 3-ms window of lead-lag delay from 1 to 100 ms (interval 0.1 ms). This procedure was repeated for each subject, with individual subject data appearing in lateralization plots as light grey (“one location”) and light red (“two location”) points. Because different subjects were presented with different sets of lead-lag delays (resultant of the unique trajectory of adaptive trackers on each run), cross-subject means per 3-ms window were weighted by the number of subjects (out of 10) for which data existed. These weights are given by the size of bold points (black and red) across lead-lag delays, while error bars give the weighted standard deviation at each delay (standard deviations are not given for windows in which fewer than two subjects responded). Dashed black and red lines give weighted linear fits to the mean lateralization responses. Finally, data are given *only* for right-lead, left-lag trials, which enable clearer assessment of trends in lateralization across lead-lag delay than left-lead, right-lag trials (for further discussion of this point and for left-lead, right-lag lateralization data for all experiments of the present chapter, see Appendix A). Broadly consistent with expectations, lateralization responses on trials for which subjects reported perceiving “one location” (black) generally fell to the right of midline (i.e., to the right of the horizontal midpoint on the selected response button), in agreement with the ITD or ILD carried by the lead, while lateralization responses on trials for which the subject reported perceiving “two locations” (red) generally fell to the left of midline, in agreement with the ITD or ILD carried by the lag. Wholly *inconsistent* with expectations was the finding that the magnitude of lateralization for “one location” responses was systematically reduced as a function of lead-lag delay. This pattern was particularly evident in conditions featuring elevated echo thresholds (section 4.3.1.1), where mean lateralization responses for “one location” trials at lead-lag delays beyond ~20 ms (i.e., those trials making the greatest contribution to the elevation of echo thresholds) fell close to or at the midline. This held true for both ITD and ILD

conditions: weighted lines of best fit (fit to the mean data) had significantly negative slopes across both ITD ($t(3)=-16.19$, $p<0.05$) and ILD conditions ($t(3)=-24.55$, $p<0.05$).

4.3.2 Interim discussion

The results of Experiment I suggest that the precedence effect is more robust for ITD than ILD, consistent with the data of Krumbholz and Nobbe (2002) and several previous studies (e.g., Saberi *et al.*, 2004; Stecker and Brown, 2010; Chapters 2 and 3). Across all tested conditions, echo thresholds for stimuli lateralized by ITD exceeded those for stimuli lateralized by ILD. The difference was most notable in the Breakdown condition, where a sudden perturbation in the ITD of the test stimulus relative to that of the conditioner stimulus (specifically, a switching of the lead and lag ITD) failed to produce a change in echo threshold (i.e., failed to disrupt fusion). These observations support the notion that different precedence effects exist for ITD and ILD, and, by extension, that free field measurements of precedence must reflect a combination of contributions from ITD and ILD: most critically, the data suggest that the breakdown effect demonstrated in the free field depends on the sudden change in lead-lag ILD occurring with a switching of the lead and lag, while the concomitant sudden change in ITD is evidently inconsequential.

The present data additionally suggest that when a fused image is perceived at relatively long lead-lag delays, localization dominance is weak. For many such trials, lateralization responses fell near the midline, suggesting that the lead and lag cues contributed almost equally to the response. As these trials make the greatest contribution to the elevation of measured echo thresholds, it follows that the “built-up” precedence effect may feature fusion without localization dominance, a finding that, if replicated, could significantly shift notions about the ecological utility of the precedence effect. This possibility is explored further in Experiments II and III, while an alternative explanation for near-midline lateralization (concerning the presence of diotic ILD in ITD stimuli and diotic ITD in ILD stimuli) is considered specifically in Experiment III.

4.4 Experiment II: Cross-cue transfer of buildup

Several investigators have suggested that the dynamic precedence effect depends on rapid acquisition of “knowledge” about the *spatial* arrangement of sound sources and their reflections within a given listening environment (Clifton *et al.*, 1994; Grantham, 1996; Blauert, 1997; Litovsky *et al.*, 1999; Freyman and Keen, 2006; Keen and Freyman, 2009). In similar terms, the effect has been construed as a high-level (e.g., cortical) phenomenon dependent on “construction

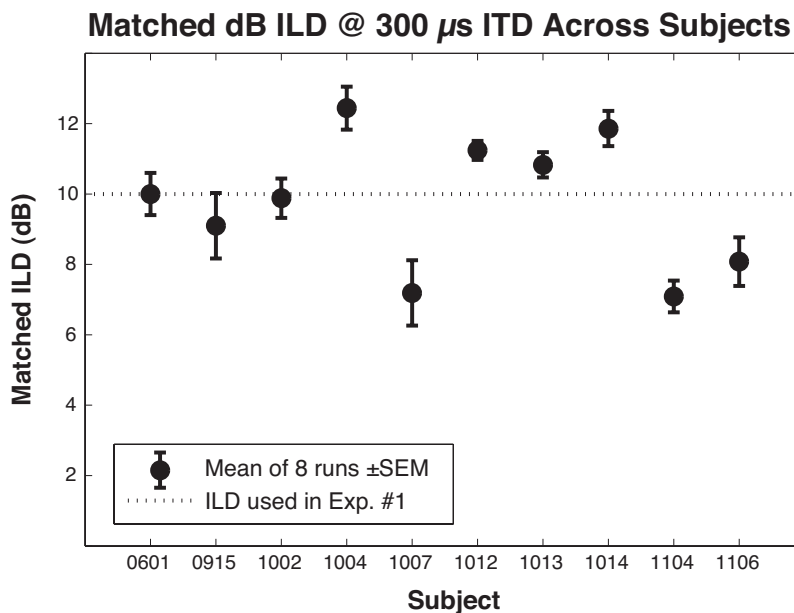


Figure 4.7: Data from Experiment II ITD-ILD matching task. Points give the mean dB ILD matched to a 300 μ s ITD standard by each subject (\pm SEM across 8 runs). For comparison, the dashed line plots the 10 dB ILD used for all subjects in Experiment I.

of an internal model of auditory space”

(Freyman and Keen, 2006, Sanders *et al.*, 2011). As suggested in the introduction to this chapter, while it is difficult to rigorously test for the existence or nonexistence of an internal representation of auditory space within a purely psychophysical study

of the precedence effect, the view that dynamic precedence reflects listeners’ immediate experience with the spatial configuration of sources and echoes in a given listening environment holds that the buildup effect should depend only on the *spatial* percept induced by a repeated stimulus, and not on the specific acoustic cues it carries.

The present experiment was thus designed to assess the “spatial-specificity” versus “cue-specificity” of the dynamic precedence effect in two novel buildup conditions where conditioner and test stimuli were lateralized by *different* cues of *equal* subjective magnitude. Prior to testing in the main experimental task, subjects completed an ITD-ILD matching task to obtain values of ILD

(carried by single clicks) that matched the subjective lateralization of a $\pm 300 \mu\text{s}$ ITD standard (see section 4.2.2.2). These individually determined values of ILD (see Figure 4.7) were used for all ILD conditioner and ILD test stimuli in Experiment II. Stimulus conditions in Experiment II consisted of Baseline ITD and ILD conditions identical to those of Experiment I (with the exception that the individually determined values of ILD in Fig. 4.7 were used in place of the ± 10 dB used in Experiment I), and two novel “buildup” conditions, (1) Buildup ILD, Test ITD and (2) Buildup ITD, Test ILD conditions, in which the conditioner and test were lateralized by ILD and ITD and ITD and ILD, respectively.

4.4.1. Results and interim discussion

Figure 4.8 gives cross-subject mean fusion echo thresholds for conditions of Experiment II. Mean Baseline ITD and ILD thresholds were nearly identical to those measured in Experiment I (thresholds for Baseline ITD Ex. I = 8.69 ms, Ex.

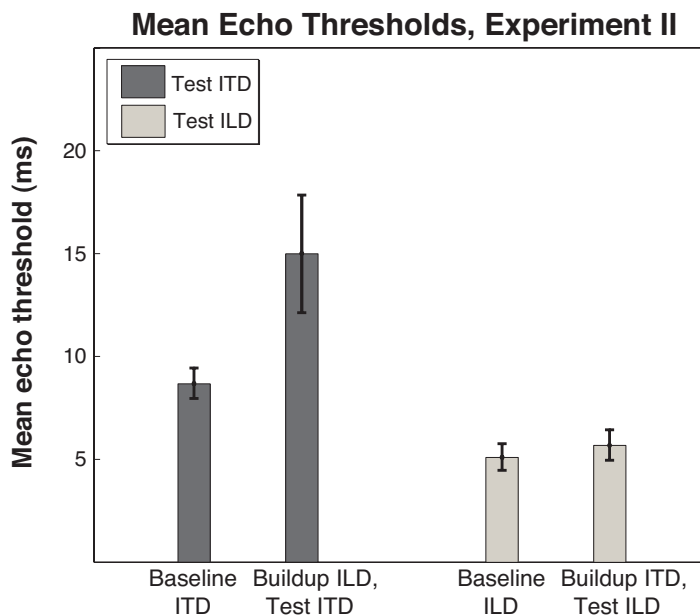


Figure 4.8: Mean echo thresholds for conditions of Experiment II. Legend as in Fig. 4.4.

II = 8.67; Baseline ILD Ex. I = 4.96 ms, Ex. 2 = 5.09 ms).

Toward the primary question addressed by Experiment II, the mean Buildup ILD, Test ITD threshold appeared to be moderately elevated relative to the mean Baseline ITD threshold, while the Buildup ITD, Test ILD threshold appeared to be equal to the threshold in the Baseline ILD condition. Individual subject

data (Figure 4.9) reveal that the mean Buildup ILD, Test ITD threshold was skewed by two subjects, most especially subject 1012, whose threshold in that condition was 37.5 ms. Although this point might be treated as an outlier, subject 1012’s data were not unusual in other conditions, and very high buildup

thresholds are occasionally measured in subjective fusion tasks (e.g., Yang and Grantham, 1997). Paired *t*-tests (corrected for two comparisons) indicated that the Buildup ILD, Test ITD threshold was marginally significantly higher than the Baseline ITD threshold ($t(9)=2.65$, $p=0.027$), while Buildup ITD, Test ILD and Baseline ILD thresholds were

not significantly different ($t(9)=1.16$, $p=0.274$). In comparison to “within-cue” buildup thresholds from Experiment I, the mean Buildup ILD, Test ITD threshold of the present experiment was 14.99 ms (12.54 ms with removal of subject 1012) versus 17.93 ms for Buildup ITD (Experiment I), while the Buildup ITD, Test ILD threshold of the present

experiment was 5.68 ms versus 13.84 ms for Buildup ILD (Experiment I). The data thus suggest that maximal buildup requires static ITD and ILD. Although limited cross-cue transfer of buildup was measured with an ILD conditioner and ITD test (particularly for two subjects), none was measured in the opposite case.

Lateralization data for Experiment II are displayed in Figure 4.10. Trends in lateralization as a function of lead-lag delay closely followed those observed in Experiment I. Most notably, when one location (black) was reported at lead-lag delays beyond ~20 ms, subjects tended to respond near the midline, suggesting weak localization dominance. Consequently, as in Experiment I, the slope of lines fit to “one location” responses taken across conditions was significantly negative ($t(3)=-3.23$, $p<0.05$). Additionally, a subtler trend was observed in responses for “two locations” responses (red): At lead-lag delays near the echo threshold, “two locations” responses tended to be shifted leftward (i.e., toward the lead). This trend, leading to slightly negative slopes for lines fit to “two

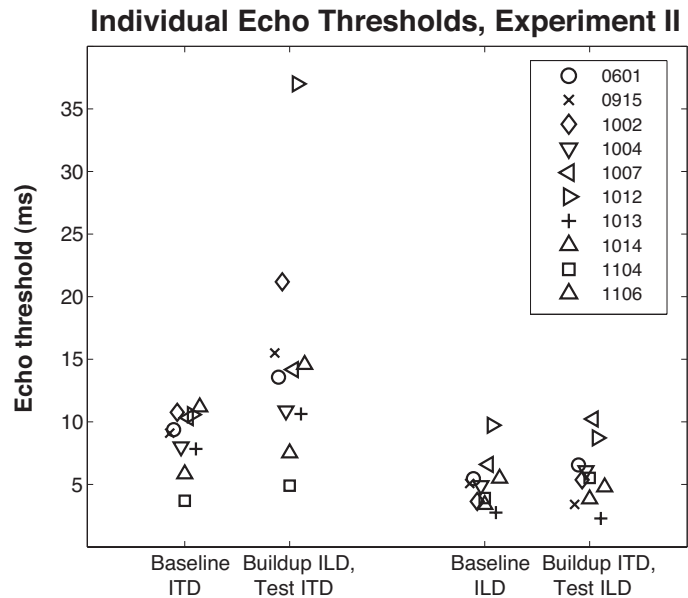


Figure 4.9: Individual echo thresholds for conditions of Experiment II. Legend as in Fig. 4.5.

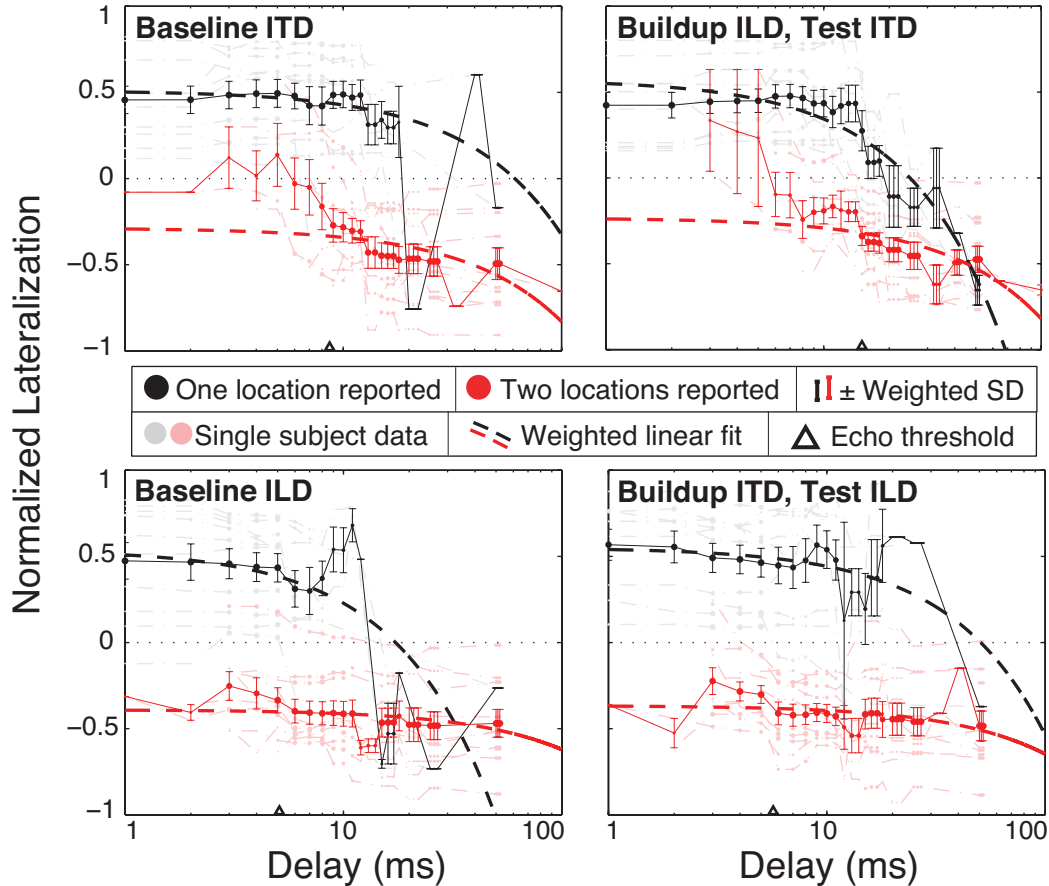


Figure 4.10: Lateralization data for Experiment II. Legend as in Fig. 4.6.

locations” data, was also present in the data of Experiment I, but was particularly evident in the Baseline ITD and Buildup ILD, Test ITD conditions of the present experiment. This observation and the observation of weak lateralization dominance for built-up lead-lag stimuli, as well as differences between ITD and ILD precedence effects discussed hereto, are explored further in Experiment III.

4.5 Experiment III: Dynamic precedence effects within a single hemifield

The majority of precedence effect studies have used “lead” and “lag” stimuli that were symmetrically opposed across the interaural midline (e.g., Wallach *et al.*, 1949; Thurlow and Parks, 1961; Freyman *et al.*, 1991, Krumbholz and Nobbe, 2002; Experiments I and II, and dozens of others). Use of such stimuli offers certain advantages – perhaps most obviously, the ease of comparison of new data and decades of published data from similar paradigms,

but also more substantive methodological advantages such as avoidance of differences in sensitivity to perturbations in the lead versus lag resulting simply from baseline differences in spatial sensitivity across azimuth. Nonetheless, exclusive use of binaurally symmetric stimuli also presents certain disadvantages: Of greatest concern, information is only obtained about one type of synthetic listening condition: While any paradigm involving a direct signal and single equal-intensity reflection is clearly removed from natural listening, where direct sound is likely to be accompanied by at least tens of reflections (see Chapters 2, 3 and 5), extrapolating psychophysical performance in an experiment employing lead-lag stimuli of a single spatial configuration to “listening in rooms” seems rather incomplete.

In the present investigation, another difficulty related to the use of binaurally symmetric stimuli may be identified: We (and many investigators previous) have adopted the terms “ITD stimuli” and “ILD stimuli” to describe stimulus conditions in which the ITD or ILD was manipulated while the ILD or ITD was held constant (usually at 0 dB or 0 μ s). For any binaural stimulus, however, both cues are always present. Thus, the ITD test stimuli of Experiment I can also be described as “+/-300 μ s ITD, 0 dB ILD lead-lag click pairs,” and the ILD stimuli as “+/-10 dB ILD, 0 μ s ITD lead-lag click pairs.” This is a critical consideration in light of the lateralization data obtained in Experiments I and II. When “one location” was reported at lead-lag delays beyond ~ 20 μ s, localization dominance appeared to be very weak: Rather than responding on the side consistent with cues carried by the lead (as at brief delays), subjects responded near the midline. We interpreted these responses to indicate a substantial contribution to lateralization by both lead and lag cues, i.e., “averaging” of the lead and lag. An alternative explanation could be that, given a diffuse image comprised of disparate ITD or ILD lead and lag cues, subjects were compelled to respond to the concomitant and highly stable unmanipulated ILD or ITD cue. To address this concern, Experiment III employed lead and lag stimuli confined to a single “spatial hemifield,” such that the average of the manipulated cue across the lead and lag was nonzero.

4.5.1 Stimuli

Lead clicks in ITD conditions of Experiment III carried $\pm 600 \mu\text{s}$ ITD, while lag clicks carried $0 \mu\text{s}$ ITD. It was assumed on the basis of Experiment II ITD-ILD matching data (Fig. 4.7) that the value of ILD producing lateralization comparable to $\pm 600 \mu\text{s}$ ITD was likely to vary among subjects; subjects thus completed an additional ITD-ILD matching task prior to beginning Experiment III. The task was

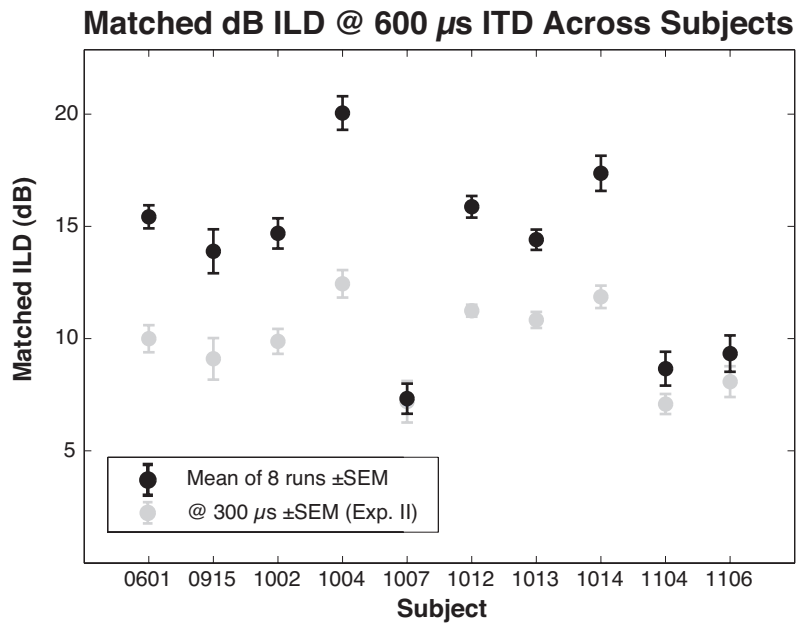


Figure 4.11: Data from Experiment III ITD-ILD matching task. Bold Points give the mean dB ILD matched to a 600 μs ITD standard for each subject (\pm SEM across 8 runs). For comparison, light grey points give matches for a 300 μs standard (Experiment II matching task).

identical to that completed prior to Experiment II (Fig. 4.7), with the exception that the ITD standard to be matched by the ILD pointer carried $\pm 600 \mu\text{s}$ ITD. Results from the matching task are plotted in Figure 4.11.

These individually determined values of

ILD were used in all ILD conditions of

Experiment III: the lead

always carried \pm the matched ILD (mean 13.7 dB), while the lag carried 0 dB ILD. Thus, the “average” of lead and lag ITD in ITD conditions was $\pm 300 \mu\text{s}$, and the average of lead and lag ILD in ILD conditions was half of the matched ILD, while the unmanipulated cue remained at $0 \mu\text{s}$ or dB for both lead and lag clicks.

4.5.2 Results and interim discussion

4.5.2.1 Echo thresholds

Figure 4.12 gives mean echo thresholds for ITD and ILD conditions of Experiment III. Compared to Experiment I, echo thresholds were more similar between ITD and ILD conditions – an unanticipated finding. Nonetheless, a

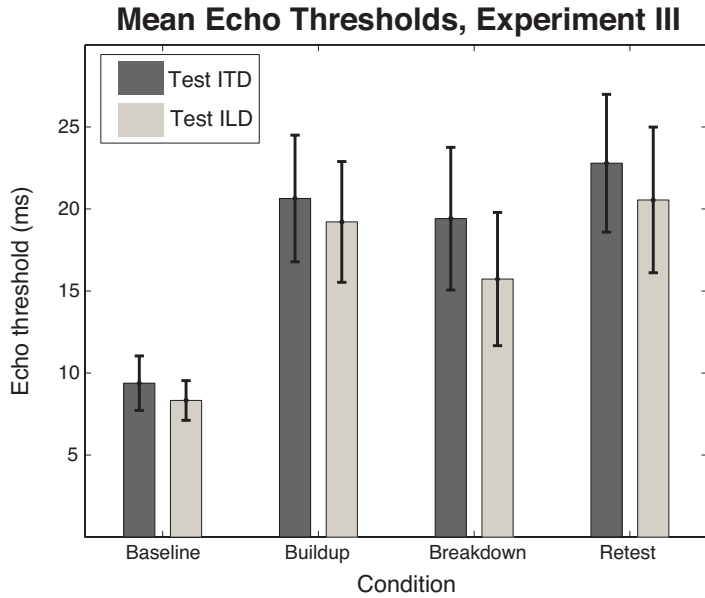


Figure 4.12: Mean echo thresholds for ITD and ILD conditions of Experiment III. Legend as in Figs. 4.4 and 4.8).

repeated-measures ANOVA demonstrated that the main effect of cue was significant, with higher ITD- than ILD-based echo thresholds across conditions ($F(1,9)=19.57$, $p<0.05$). The main effect of condition was also significant ($F(3,27)=12.34$, $p<0.05$), while the cue x condition interaction was not ($F(3,27)=1.09$, $p=0.36$). Follow-up paired t -tests demonstrated that the

breakdown effect remained significant although reduced in magnitude for ILD (Buildup vs. Breakdown, $t(9)=2.91$, $p<0.025$) and non-significant for ITD ($t(9)=0.56$, $p=0.591$). Individual data (Figure 4.13) generally support the mean data, although inter-subject variability was greater than in Experiments I and II, with several subjects exhibiting very high echo thresholds even in Baseline conditions. Finally, two post-hoc repeated-measures ANOVA demonstrated that

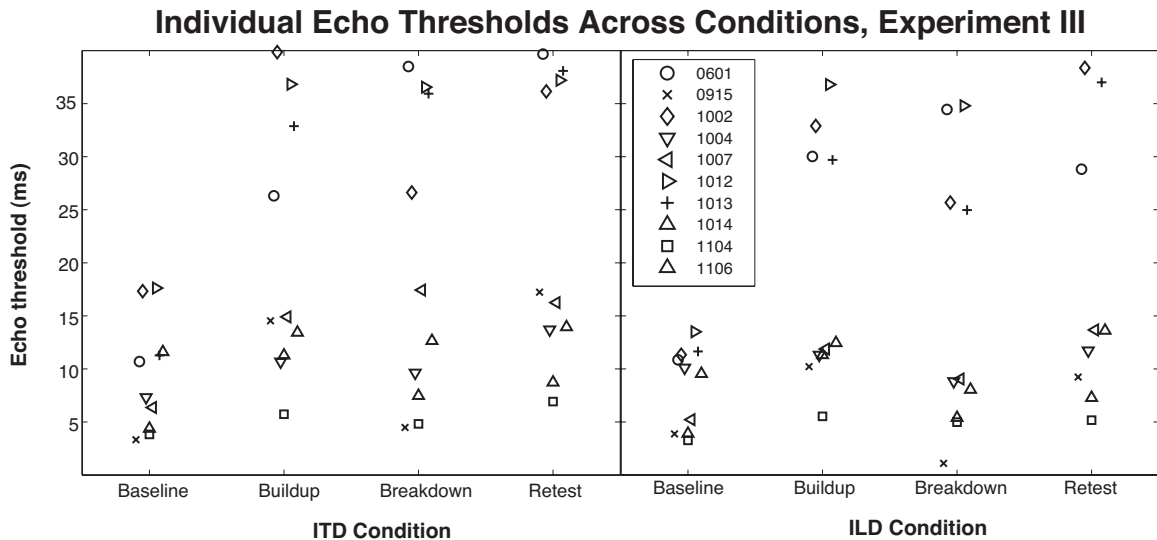


Figure 4.13. Individual echo thresholds for ITD (left) and ILD (right) conditions of Experiment III. Legend as in Figs. 4.5 and 4.9.

ILD-based echo thresholds were significantly higher in the “single-hemifield” conditions of the present experiment than in the otherwise identical conditions of Experiment I (effect of experiment, $F(1,9)=6.13$, $p<0.05$), while ITD-based echo thresholds were not ($F(1,9)=2.12$, $p=0.18$). The fact that ILD-based echo thresholds apparently depend on lead and lag locations while ITD-based echo thresholds do not is further indicative of discrete precedence effects for the cues.

4.5.2.2 Lateralization responses

Figure 4.14 gives lateralization data for Experiment III. As before, lateralization responses for “one location” trials at longer lead-lag delays tended to fall intermediate to “one location” responses at short delays and “two location” responses, giving negative slopes to lines fit to lateralization data (in 7 of 8 conditions, a significant cross-condition effect, ($t(7)=-3.25$, $p<0.05$). While this pattern was not as clear as in Experiments I and II, suggestive of somewhat stronger localization dominance in the present experiment, the present data support the notion that responses intermediate to the lead and lag typical at longer delays reflect weakened localization dominance rather than an artifact of stimulus design. Responses did *not* fall near 0 (the scaled equivalent of the interaural midline) as in Experiments I and II, suggesting that subjects were in all experiments responding to an “average” (or, more likely, a weighted average) of the manipulated lead and lag cues rather than the diotic unmanipulated cue. Finally, in support of earlier observations (section 4.4.1), “two locations” responses at *short* lead-lag delays (i.e., delays near or below the echo threshold) in conditions of Experiment III appeared to be “pulled” toward the lead, indicative of extant localization dominance (dominance of lead cues over lag cues) in the *absence* of fusion. This pattern was somewhat more evident in the present experiment than in Experiments I and II, giving rise to significantly negative slopes to lines fit to “two locations” responses as well ($t(7)=-6.28$, $p<0.001$), and further indicative that the precedence effect – in terms of both fusion and localization dominance – may be somewhat stronger when stimuli are confined to a single hemifield (see also Litovsky and Shinn-Cunningham, 2001).

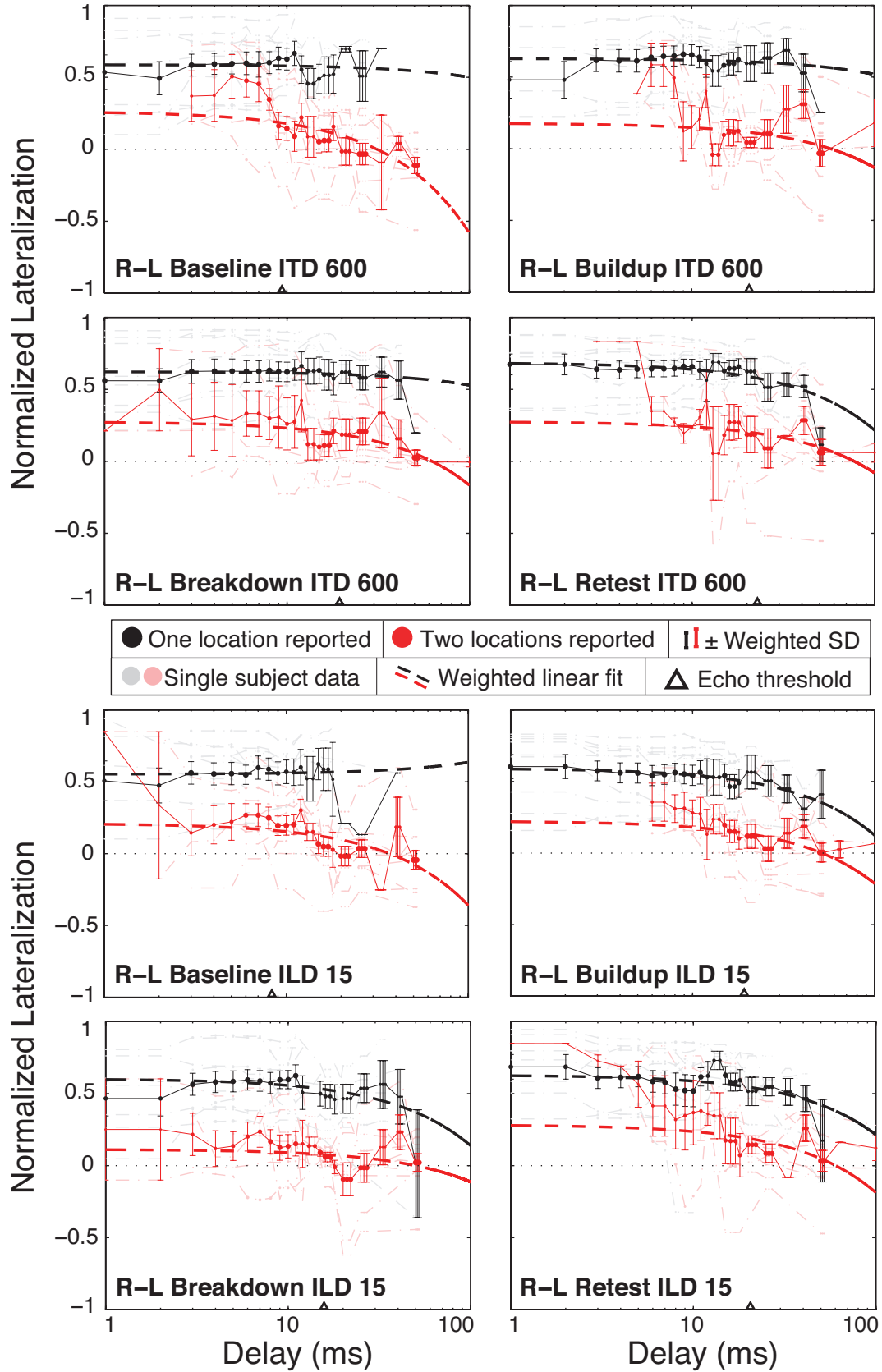


Figure 4.14: Lateralization data for Experiment III. Legend as in Figs. 4.6 and 4.10.

4.6 Summary and general discussion

Accurate sound localization in ordinary listening environments (e.g., rooms) requires that listeners respond to the early-arriving spatial cues carried by the direct sound rather than the spurious cues carried by reflections and reverberation arriving from a few to hundreds of milliseconds later. In general, this facility is believed to depend on (1) perceptual fusion of the direct signal and its reflections (fusion) and (2) dominance of the fused percept by the early-arriving spatial cues (localization dominance) (e.g., Wallach *et al.*, 1949, Thurlow and Parks, 1961; Djelani and Blauert, 2001). Clifton and colleagues (e.g., Clifton, 1987; Clifton *et al.*, 1994, Keen and Freyman, 2009) have further demonstrated that stimulus repetition leads to *enhancement* of the fusion effect, evidenced by a two- or three-fold increase in the “echo threshold” (although such enhancement is also known to be subject to disruption by interjecting novel stimuli, i.e., breakdown). By extension, it has been assumed that stimulus repetition must lead to similar enhancement of localization dominance, and therefore that the buildup effect must facilitate enhanced localization of repeated (or ongoing) sound sources in ordinary listening environments.

The fusion data of the present investigation indicate, consistent with the report of the Krumbholz and Nobbe (2002) that the fusion aspect of the precedence effect is more robust for stimuli lateralized by ITD than stimuli lateralized by ILD. Most notably, the data of Experiments I-III demonstrated, in close correspondence to the data of Krumbholz and Nobbe (2002), that “built-up” fusion for repeated lead-lag ITD stimuli persisted even when relatively drastic changes were made to the lead and lag ITD cues (relative to the conditioner ITD cues), while fusion for analogous ILD stimuli “broke down” with such changes. Further suggestive of discrete fusion effects for ITD and ILD, the data of Experiment II demonstrated minimal transfer of buildup between spatially *equivalent* conditioner and test stimuli carrying *different* spatial cues. Finally, the fusion data of Experiment III demonstrated that fusion of ILD stimuli was somewhat more robust when lead and lag cues were confined to a single hemifield (vs. interaurally opposed, Experiments I and II), while fusion of ITD

stimuli was statistically invariant across experiments. Taken together, these data suggest that the dynamic temporal limit of lead-lag fusion reported in free field and virtual auditory space studies (e.g., Clifton and Freyman, 1989; Djelani and Blauert, 2001; Keen and Freyman, 2009) depends on dual sensitivity to ITD and ILD: built-up fusion is robust to changes in ITD cues, but aberrant ILD cues readily disrupt fusion (Breakdown conditions of Experiments I and III, Buildup ITD, Test ILD condition of Experiment II). Interestingly, such disruption was apparently reduced when the ILD carried by the “aberrant” stimulus was 0 dB (Buildup ILD, Test ITD, Experiment II; Breakdown, Experiment III). We suggest that this dependence of the breakdown effect on aberrant (and especially nonzero) post-onset ILD may capitalize on the statistics of sound in natural environments: Assuming that ILD is computed with a 1-2 ms window of integration (e.g., Tollin, 2003), post-onset ILD in rooms are often near zero for brief stimuli (e.g., Rakerd and Hartmann, 1985; cf. Shinn-Cunningham *et al.*, 2005). Significantly *nonzero* post-onset ILD are thus likely to correspond to novel sound sources. A signal lateralized by ITD alone may not be similarly evaluated on the basis of post-onset cues, as post-onset ITD tends to vary erratically moment-to-moment (e.g., Rakerd and Hartmann, 1985).

Most interestingly, the *lateralization* data of the present investigation indicated that localization dominance and fusion aspects of the precedence effect were at least partially dissociated when the echo threshold was elevated by stimulus repetition. This observation is consistent with the data of Yang and Grantham (1997), which demonstrated a divergence of echo thresholds measured in two different buildup paradigms – (1) a subjective fusion paradigm (like that in the present experiment) and (2) an objective “discrimination suppression” paradigm, where subjects were required to discriminate changes in the location of the lag. In that study, thresholds were significantly more elevated by stimulus repetition in the fusion paradigm than in the discrimination suppression paradigm, suggesting that at extended lead-lag delays, listeners remained sensitive to spatial information carried by the lag despite their proclivity to report a single auditory event. In our paradigm, which assessed localization

more directly, perception of a fused image at lead-lag delays $> \sim 20$ ms appeared to entail an averaging of lead and lag information (i.e., weak or absent localization dominance). Critically, this finding suggests that the ecological utility of the buildup effect – and especially the notion of a high-level mechanism specialized for enhancement of sound localization in reverberant listening environments (e.g., Djelani and Blauert, 2001, Freyman and Keen, 2006; Keen and Freyman, 2009) – should be reconsidered.

We suggest that the buildup effect may be more generally understood as an example of auditory object formation (for review, see Bregman, 1994). In the context of object formation, the utility of repetition-enhanced lead-lag fusion (i.e., buildup, a term already existing in the auditory scene analysis literature, Bregman, 1994) is clear – independent of localization, perceptual fusion of a signal and its reflected copies (e.g., speech) should facilitate improved *identification*. In support of this interpretation, a study by Brandewie and Zahorik (2010) demonstrated a significant improvement in speech intelligibility (a completely non-spatial measure) with several seconds of prior exposure to speech sounds in a simulated room (versus a control condition with no prior exposure). Indeed, from an ecological standpoint, the need for “enhancement” of our already excellent auditory localization capabilities may be minimal given the likelihood of concomitant and disambiguating visual cues (cf. Bishop et al., 2011). Nonetheless, localization dominance did appear to persist beyond 10 ms lead-lag delay in some conditions of the present investigation (Figs. 4, 8 and 12), and modest elevation of discrimination suppression thresholds has been measured in buildup paradigms previously (e.g., Freyman et al., 1991; Yang and Grantham, 1997). The buildup effect (and the precedence effect itself) may thus well depend on multiple mechanisms, one or more of which may modulate binaural sensitivity itself. Elucidation of the foregoing issues will require additional free field studies which measure localization in buildup paradigms explicitly and, ultimately, an understanding of the mechanism or mechanisms underlying fusion and localization dominance.

Textual footnotes

¹ Absolute sensitivity to ITD and ILD was not measured in subjects of the present investigation. We assume on the basis of past studies in our lab, accounts of individual differences in the literature (e.g., McFadden *et al.*, 1973), and variation in the lateral extent of touch screen responses across subjects in the present data (note error bars in Figs. 4.6, 4.10, and 4.14), that the employed ITD and ILD produced various degrees of lateralization across subjects. Nonetheless, ITD and ILD were calibrated (particularly in Experiments II and III) to produce equivalent lateralization *within* each subject. As absolute lateralization was not of particular interest, lateralization data were normalized across subjects.

² The term “localization dominance” is used throughout the present investigation to describe dominance of localization cues carried by the lead over those carried by the lag, consistent with the nomenclature delineated by Litovsky *et al.* (1999) and used across a majority of precedence effect studies since. We note, however, that our stimuli, like all “ITD only” and “ILD only” headphone stimuli, do not produce externalized perception of sound that is localized in space, but rather intracranial perception of sound that is *lateralized* in the direction consistent with the dominant cue. While the observed dominance of leading cues at brief delays in the present investigation might thus more accurately termed “lateralization dominance,” we have adopted the more common term already existing in the literature.

³ Lead-lag delays briefer than ~1 ms produce the perception of a strongly fused and compact image intermediate to “lead” and “lag” sources, a phenomenon qualitatively different from the usual precedence effect known as “summing localization” (Wallach *et al.*, 1949; see Litovsky *et al.*, 1999). Although summing localization was not under study in the present investigation, trials with delays briefer than 1 ms (~0.5 ms) were included as a variety of catch trial – a “two locations” response at 0.5 ms lead-lag delay would suggest the subject was not performing the task as instructed. Such responses were not observed; “summing localization” data were not further considered in the present report.

5

Consequences of peripheral auditory processing on the temporal dynamics of ITD and ILD sensitivity: A dual-display model

Models concerning the function of neurons and neural systems (e.g., James, 1890; Lopicque, 1907; Hebb, 1949; Hodgkin and Huxley, 1953) proliferated around the same time that pioneers in psychoacoustics such as Rayleigh and Stevens published the work that would influence the century of binaural psychophysical research to follow. Today neural modeling comprises an expansive field all its own, with a substantial degree of interest among investigators in the elaborate interconnections and temporal precision of the binaural system. The present chapter is but a brief foray into binaural auditory modeling in the context of the psychophysical studies described hereto. Although devoting a single chapter to the topic may seem almost cursory, the demonstrated explanatory (and predictive) power of relatively simple models of low-level auditory processing invites their continued development and refinement. Following a brief review of existing models of binaural processing, which have historically focused on ITD processing, a novel “dual-display” model that considers ITD and ILD processing *independently* is described. Finally, model outputs are considered for various signals of interest, including the signals employed in the foregoing psychophysical experiments.

5.1 Existing models of binaural processing

5.1.1 Peripheral components

Colburn (1973) published the first comprehensive model of binaural processing. In addition to binaural components to be detailed subsequently, his and other models since have included a peripheral processing stage, i.e. a stage at which each ear and/or auditory nerve transforms acoustic input into an “internal representation” of the signal. In general, the inclusion of a peripheral processing stage recognizes that the binaural system’s access to information

carried by an acoustic signal is constrained by the fidelity of auditory transduction (cf. Siebert, 1968). Colburn's (1973) model of peripheral processing was based on several assumptions – informed by the theoretical and empirical work of Siebert (e.g., 1965, 1968) and Kiang *et al.* (e.g., 1965) – that have persisted in more recent models, most notably that (1) the ears are identical but operate independently, (2) the nerve of each ear is comprised of many individual nerve fibers, each of which can be characterized at any moment as “firing” or “not firing,” (3) whether or not an individual nerve fiber is firing at a given moment depends on the acoustic input, its tuning characteristic, and internal noise (i.e., firing is always probabilistic), and (4) nerve fibers of a given tuning characteristic (i.e., characterizable by a narrowband filter centered on a given frequency) are bound to a corresponding impulse response. The significance of the final assumption is conveniently illustrated by considering the firing of several spectrally remote populations of model nerve fibers (i.e., sets of fibers innervating discrete regions of the cochlea) in response to a monophasic click (an approximate impulse) (see Figure 5.1): Although the input signal itself is exceedingly brief and broadband, the firing response of each set of fibers closely follows the periodicity and decay of the impulse response of the narrowband filter centered at the specified frequency (right margin). Importantly, empirical data suggest that the bandwidth of auditory nerve fiber tuning is approximately 1/3-octave across frequency (e.g., Kiang *et al.*, 1965, see Colburn, 1973), resulting in a *decrease* in absolute filter bandwidth with decreasing center frequency, and thus relatively slow dynamics in low-frequency channels (see Fig. 5.1).

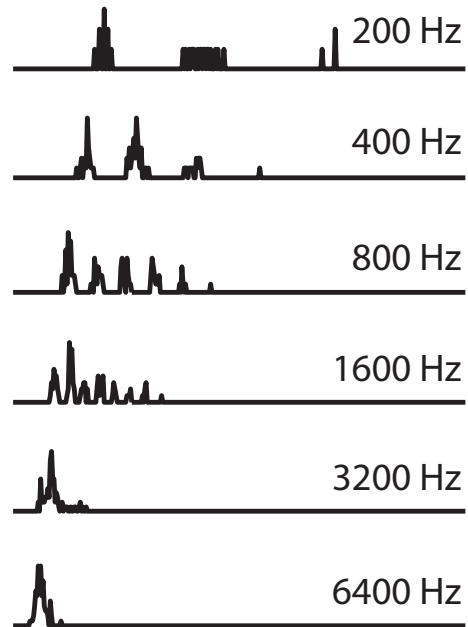


Figure 5.1: Post-stimulus time histograms (PSTHs) for model populations of auditory nerve fibers responding to an approximate impulse. The ringing response of the basilar membrane results in cyclic phase-locked responses in low-frequency channels. In high-frequency channels, greater bandwidth permits faster responses.

Binaural models over the past four decades have integrated refinements to the peripheral components (“front-end”) employed by Colburn (1973), but there has also been a curious tendency to ignore auditory nerve firing altogether and instead derive the inputs to the binaural processor from the simulated outputs of idealized cochlear filters. For example, an influential model popularized by Trahiotis and colleagues (e.g., see Bernstein and Trahiotis, 2002) employs a “bank” of gammatone filters (produced by multiplying a set of sinusoids by a gamma distribution) to mimic spectral decomposition of the signal by the cochlea. Filter outputs are then half-wave rectified to mimic the synaptic output of inner hair cells, and, in more recent models, passed through a compression algorithm to mimic the nonlinear amplification produced by outer hair cells (see, e.g., Bernstein and Trahiotis, 2011). Finally, outputs are low-pass filtered at ~ 150 Hz to account for observed limitations in temporal resolution (e.g., Bernstein and Trahiotis, 2002). While these half-wave rectified, compressed, low-passed gammatone filter outputs are comparable to the output generated in Colburn’s (1973) auditory nerve model assuming a sufficiently large number of memory-less nerve fibers, use of such idealized outputs ignores both the stochastic process underlying neural firing and, more significantly, the fact of neural adaptation. Although the reasons for a computationally tractable front-end are clear from a historic standpoint – the computational power necessary to simulate stochastic firing and biologically realistic adaptation did not exist four decades ago – consideration of adaptation effects seems critical to understanding binaural processing as it actually occurs, particularly for dynamic signals (e.g., sources in echoic environments). A model developed by Breebaart and colleagues (2001; section 5.1.3) emphasizes the importance of peripheral neural adaptation in binaural processing. The model we propose (section 5.2) follows that approach.

5.1.2 The Jeffress-Colburn model

The most influential paper in the field of binaural modeling – indeed, one of the most influential in the auditory field at large, having been cited almost 700 times at the time of this writing – was that by Jeffress (1948). The “Jeffress

model” (1948) describes a hypothetical neural network wherein excitatory bilateral input from the two ears converges at an array of target cells that fire most vigorously when inputs from the two sides are temporally coincident. According to the Jeffress model, the array of target cells is synapsed with both contralateral and ipsilateral fibers that vary in length, assembling a gradient that changes oppositely for the contra- and ipsilateral projections from one end of the array to the other. Since an ITD should result in a disparity of spike timing on the two sides equal to the ITD, a target binaural cell should receive maximally coincident excitation and thus respond most vigorously when the internal delay imposed by its inputs opposes the external ITD. The full array of “coincidence detectors” should thus effectively perform *cross-correlation* of signals from the two ears (i.e., interaural cross-correlation). Although Jeffress proposed his model in a time when little was known of binaural anatomy and neural recordings were rarely made, physiological studies in the decades since have confirmed essential features of his model in mammalian and avian systems, including the existence of brainstem coincidence detector neurons and gradients of “preferred ITD” across populations of such neurons (e.g., Carr and Konishi, 1990; Yin and Chan, 1990).

Sayers and Cherry (1957) were the first to employ a cross-correlation algorithm to account for listeners’ performance in binaural psychophysical tasks. Their model was based on cross-correlation of the *acoustic* inputs to the two ears. Colburn (1973, 1977), who appears to have developed his model independently, was the first to predict binaural perception quantitatively as a consequence of cross-correlation of “internal representations” of the signal at each ear (generated by his rather sophisticated auditory nerve model). The majority of binaural models have followed on Colburn’s formalization of the Jeffress model (thus the “Jeffress-Colburn” moniker). Improvements on Colburn’s original formulation by Colburn and colleagues (e.g., Colburn, 1977) have included the development of a “display” for visualization of interaural cross-correlation (the so-called “binaural display”), application of a “centrality” weighting function to reflect localization acuity across azimuth, and application of monaural

intensity weighting to better account for contributions to perception of intensive *differences* at the ears (i.e., ILD) (e.g., Stern et al., 1988). Toward this same end, Lindemann (1986) later added a “contralateral-inhibition” component to a Jeffress-Colburn type model better account for effects of ILD. Nonetheless, all Jeffress-Colburn-type models are dominated by the theme of excitatory coincidence detection. A separate class of models, considered below, depends explicitly on the convergence of antagonistic binaural inputs.

5.1.3 Models based on binaural “cancellation”

In 1963, Durlach proposed a model distinct from the Jeffress model (1948) to account specifically for observed psychophysical benefits in signal detection tasks arising from phase separation of target and masker signals (the “binaural masking level difference”). Durlach (1963) posited that masking noise arriving at the two ears is “equalized” by a central processor (i.e., set to 0 ITD, 0 ILD), after which the composite left and right signals (masker + target) are subtracted – a “cancellation” process. Excepting remnants of the masker attributed to imperfect cancellation (i.e., noise in the system), the remaining signal should thus be comprised of only target components, assuming a phase (or level) difference existed between target and masker signals in the first place (e.g., noise in 0 phase, target signal in π phase). Although the “Equalization-Cancellation Model” (EC model), like the Jeffress Model, was devised primarily to explain perception related to interaural *temporal* disparities, the EC model has been highly influential toward the use of *subtraction* in models of binaural processing – a process by which ILD can theoretically also be encoded. Nonetheless, Bekesy (1930) and, following on his work, van Bergeijk (1962) were actually the first to propose a model of binaural processing based on the interaction of antagonistic inputs; van Bergeijk (1962) went so far as to specify that both ITD and ILD could be processed by a nucleus receiving excitatory input from one ear and inhibitory input from the other ear (an “EI” nucleus). Neither paper ever gained much notoriety, perhaps because neither featured the quantitative precision nor predictive utility of the EC model (Durlach, 1963).

Among other investigators, Breebaart and colleagues (e.g., Breebaart *et al.*, 2001a, 2001b) have developed a particularly sophisticated “cancellation”-based model. The Breebaart (2001a) model front-end includes middle and inner ear filtering, compression, and simulation of neural adaptation with five different time constants. These peripheral inputs are passed (via branching collaterals) to a matrix of “EI” binaural comparators, which, as in the Jeffress model, are “tuned” to specific ITD, resulting from naturally opposed delays in input conduction along the horizontal dimension of the matrix. The critical addition of the Breebaart (2001a) model is that arrays of *attenuators* are interleaved with the matrix of coincidence detectors in the vertical dimension (i.e., along each collateral), such that that input is attenuated as it progresses vertically along each collateral. Thus, an ITD biases the activity of EI comparators along the horizontal dimension, while an ILD biases activity in the vertical dimension, such that each EI element has both a preferred ITD and a preferred ILD. While the Breebaart (2001a) model thus seems an improvement on previous models that have failed to treat ILD as a primary cue, the model assumes that ITD and ILD are necessarily combined. Our psychophysical data (e.g., section 4.4) suggest discrete processing (or at least discrete contributions to perception) of ITD and ILD, consistent with studies by Hafter and Jeffress (e.g., 1968) concerning “two-image” lateralization and physiological data demonstrating partially discrete (though certainly not entirely discrete) populations of ITD- vs. ILD-sensitive neurons in the medial superior (ITD, e.g., Goldberg and Brown, 1969) and lateral superior olives (ILD, e.g., Boudreau and Tsuchitani, 1968), the inferior colliculus (e.g., Chase and Young, 2005) and perhaps even the cortex (Edmonds and Krumbholz, 2008). In the following section, we therefore propose a model in which the cues are processed in parallel rather than in series.

5.2 A dual-display model of binaural processing

5.2.1 Auditory nerve model: Zilany *et al.*, 2009

The front-end of our model is taken directly from the work of Carney and colleagues (see Zilany *et al.*, 2009). The Zilany *et al.* (2009) auditory nerve model

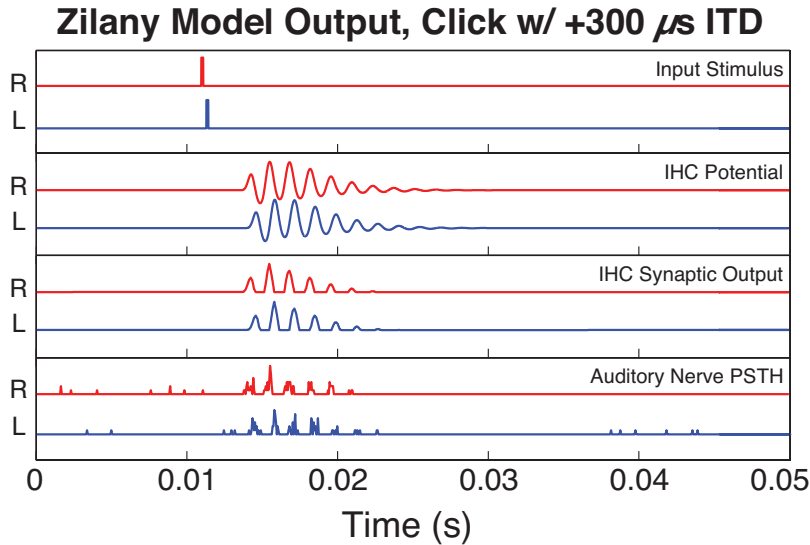


Figure 5.2: Zilany model output for “right ear” (red) and “left ear” (blue) populations of 100 800-Hz center-frequency auditory nerve fibers given a 120 μ s dichotic impulse carrying +300 μ s ITD (see Appendix B for model parameter details).

is a comprehensive model of peripheral auditory processing which includes external, middle, and inner ear mechanical filtering with compression, inner hair cell synaptic output, and simulated auditory nerve firing (the ultimate output) as regulated by a

sophisticated “power-law” adaptation algorithm. Power-law adaptation is adaptation *without fixed time constants*, where the time courses of adaptation and recovery from adaptation depend on the temporal and intensive profile of the input stimulus (e.g., Chapman and Smith, 1963). Although power-law-type adaptation is observable in a variety of auditory nerve data (e.g., Kiang, 1965, Young and Sachs, 1973), implementation of a power-law adaptation algorithm is computationally intensive, substantially more-so than implementation of traditional adaptation algorithms based on fixed exponential decay (e.g. Breebaart, 2001). Model code (run in MATLAB, Mathworks, Natick, NJ) is freely available from L. Carney’s laboratory website (2012). Selected model parameters (e.g., number of nerve fibers, fiber type, fiber center frequencies) as implemented in the present investigation are given in Appendix B.

Figure 5.2 illustrates a binaural stimulus (120 μ s dichotic click with a 300 μ s right-favoring ITD) and the three stages of Zilany model output for “right ear” (red) and “left ear” (blue) populations of 100 800-Hz center-frequency (in this example) hair cells and auditory nerve fibers. The first stage of output gives changes in inner hair cell potential as a function of time, reflecting the oscillatory rise and decay of local basilar membrane motion. The second stage gives the

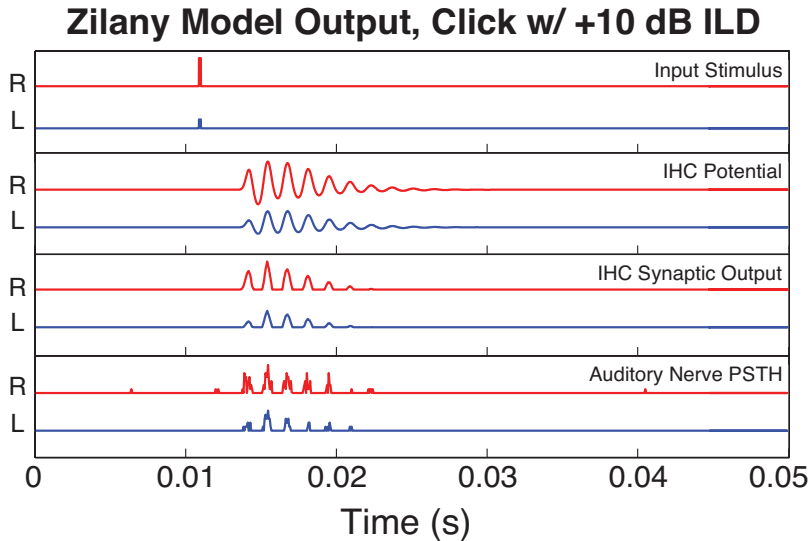


Figure 5.3: Zilany model output given a $120\ \mu\text{s}$ dichotic impulse carrying +10 dB ILD. Legend as in Fig. 5.2.

resultant changes in inner hair cell synaptic output as a function of time; the synaptic output function is much like a half-wave rectified version of the inner hair cell potential function, with added compression. The third and final stage gives the simulated post-stimulus time histogram (PSTH) of a population of 100 auditory nerve fibers innervating the population of hair cells. Given the exceedingly brief input stimulus, limited adaptation is evident in either channel, but “noise” in firing is evident, giving rise to slight perturbations in the temporal profile of each channel. Nonetheless, the right-favoring ITD in the input stimulus is clearly preserved at each stage of model output. Likewise, the approximately equal amplitude of left and right outputs reflects the 0 dB signal ILD. For the purposes of further illustration, Figure 5.3 displays another binaural stimulus passed through the Zilany model, in this case a $120\ \mu\text{s}$ dichotic click with a 10 dB right-favoring ILD and 0 ITD. The greater amplitude in the right channel than in the left channel is evident at all stages of model output; thus, the model also effectively preserves nonzero ILD.

5.2.2 ITD display: Frequency-weighted interaural cross-correlation

A plot of interaural cross-correlation across internal delay (μs , abscissa) and frequency (Hz, ordinate) is conventionally referred to as the “binaural display” (e.g., see Akeroyd, 2001). We suggest that the “binaural display” may be more accurately termed the *ITD display*, as even when the cross-correlation function is weighted by interaural intensity (e.g., Stern and Colburn, 1978), the

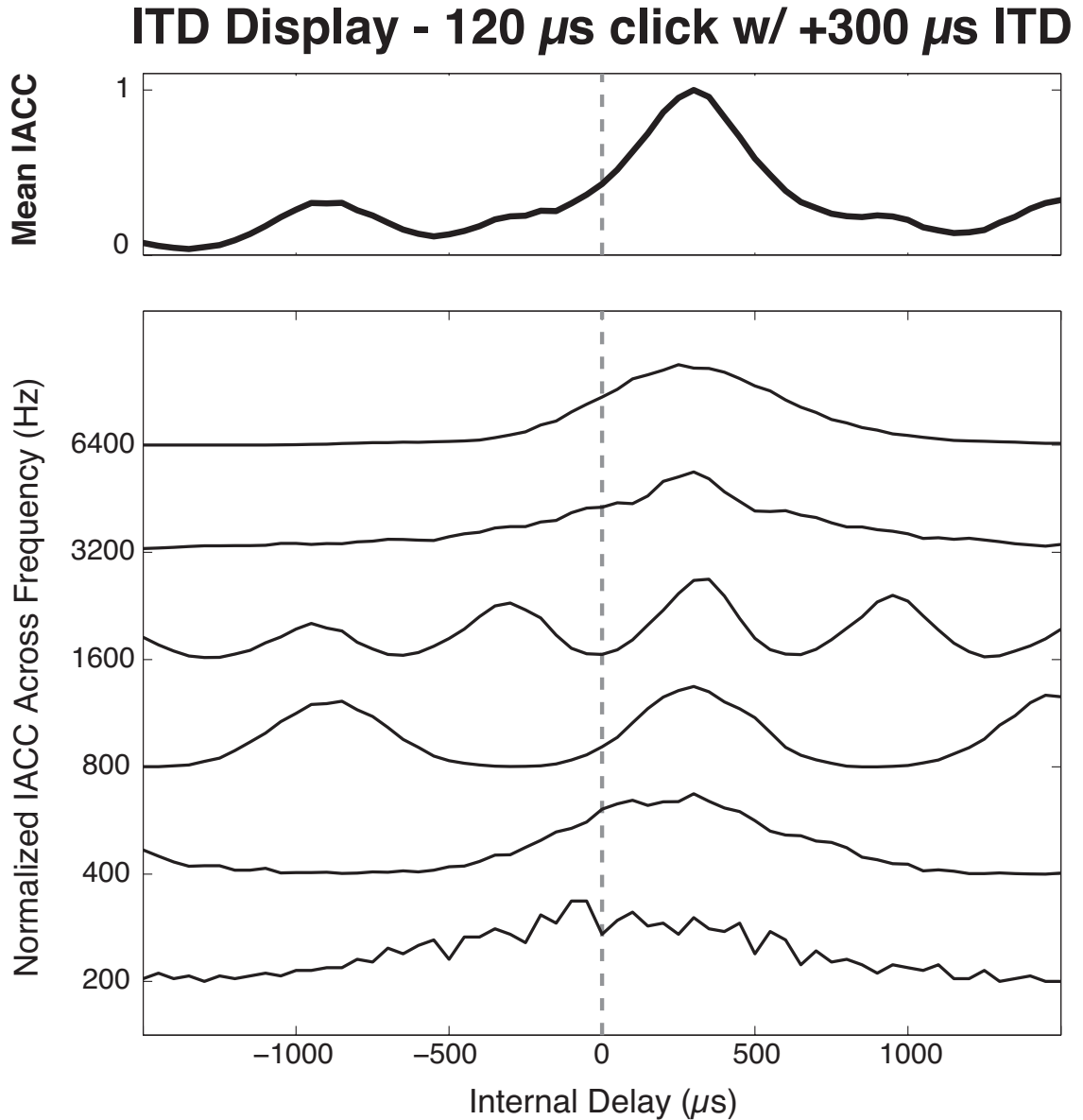


Figure 5.4: ITD display: Post-stimulus time histograms for 6 populations of auditory nerve fibers with center-frequencies spaced logarithmically from 200-6400 Hz are cross-correlated over a range of delays from -1500 to +1500 μs in 50 μs increments; the normalized interaural cross-correlation is plotted for each frequency, and the weighted cross-frequency mean (see text) is given at the top of the display. Here, the input signal was the dichotic right-leading click illustrated in Fig. 5.2. A peak in the cross-correlation function is evident in all frequency channels near the veridical +300 μs ITD, most pronouncedly in the cross-frequency mean.

predicted percept is determined by peaks across internal delay, i.e., the peaks along the “ITD” axis. Our model employs a conventional approach for computation of the ITD display. Specifically, the cross-products of left and right ear PSTHs (generated for a given input stimulus by the Zilany *et al.*, 2009 auditory nerve model) are computed at each 50 μs increment across the range of

physiologically possible delays and slightly beyond ($-1500 \mu\text{s}$ to $1500 \mu\text{s}$ “internal delay”) using MATLAB’s `xcorr` function. In order to obtain a final estimate of the ITD carried by the composite auditory nerve PSTH, a weighted cross-frequency average is computed according to Stern *et al.*’s (1988) frequency-weighting function, which gives the greatest weight to correlations in channels between approximately 400 and 1000 Hz, reflecting the demonstrated perceptual dominance of ITD within that spectral region (e.g., Zwislocki and Feldman, 1956; Wightman and Kistler, 1992, Tollin and Henning, 1999). The present version of the ITD display integrates 6 frequency channels spaced logarithmically from 200 to 6400 Hz; thus, the 400 and 800 Hz channels contribute most to the cross-frequency average. Figure 5.4 gives an example ITD display for the right-favoring dichotic click illustrated in Fig. 5.2. Although multiple peaks are evident within the lower-frequency channels, resulting from redundant peaks in cross-correlation at “slipped cycles” of the periodic PSTHs, the normalized cross-frequency average (“Mean IACC,” upper panel) reveals a single primary peak at the internal delay corresponding to the $300 \mu\text{s}$ signal ITD. For clarity, in subsequent plots, the ITD display is given as the cross-frequency mean only.

5.2.3 ILD Display: Count-comparison of left- and right-ear PSTHs

Physiological ILD sensitivity depends on integration of competitive excitatory and inhibitory inputs from the two ears within a running integration window 1-2 ms in duration (for review, see Tollin, 2003; see also van Bergeijk, 1962). If a significantly greater number of excitatory inputs arrive at a target binaural “EI” cell within the window, the cell is likely to fire, while a dominance of inhibitory inputs is likely to suppress firing. Such EI integration is known to occur in both the lateral superior olive (Boudreau and Tsuchitani, 1968) and inferior colliculus (Burger and Pollak, 2001) independent of ITD processing, which requires temporal *differentiation* rather than integration of input (although EI neurons sensitive to both ITD and ILD also certainly exist, e.g. Tollin and Yin, 2005). Thus, our goal was to develop an “ILD display” independent of interaural cross-correlation (the ITD display). A sensible first approach was to simply

ILD Display - 120 μ s click w/ +10 dB ILD

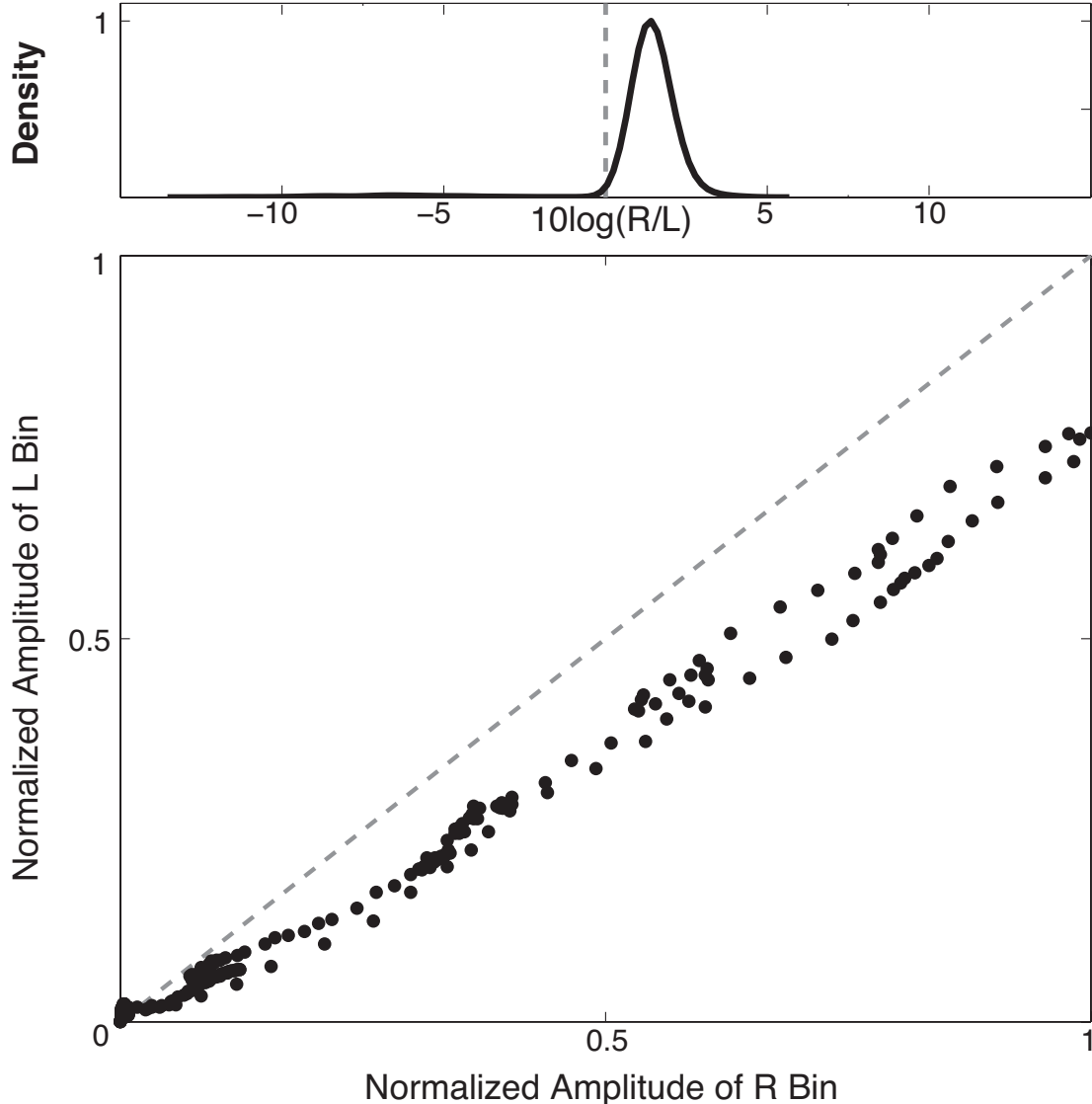


Figure 5.5: ILD display: Post-stimulus time histograms (PSTHs) for six populations of auditory nerve fibers with center-frequencies spaced logarithmically from 200-6400 Hz are summed into a composite PSTH. Normalized amplitudes of composite right versus left PSTHs within a 1.5 ms sliding window of integration are then scatter-plotted over the full PSTH duration (lower panel). Finally, a smoothed estimate of the density of right-to-left ratios represented in the scatter-plot (taken as $10 \times \log(R/L)$ is computed (by kernel density estimation). In this example, a prominent right-favoring peak is evident, consistent with the +10 dB ILD carried by the input.

scatter-plot the amplitude of the composite (cross-frequency) right versus left PSTHs (i.e., spike count comparison) for a sliding 1.5 ms window of integration over the complete PSTH duration. The lower panel of Figure 5.5 gives such a scatter-plot for right and left PSTHs generated using a dichotic click with a 10 dB right-favoring ILD (see Fig. 5.3). All points in the scatter-plot fall to the right of the

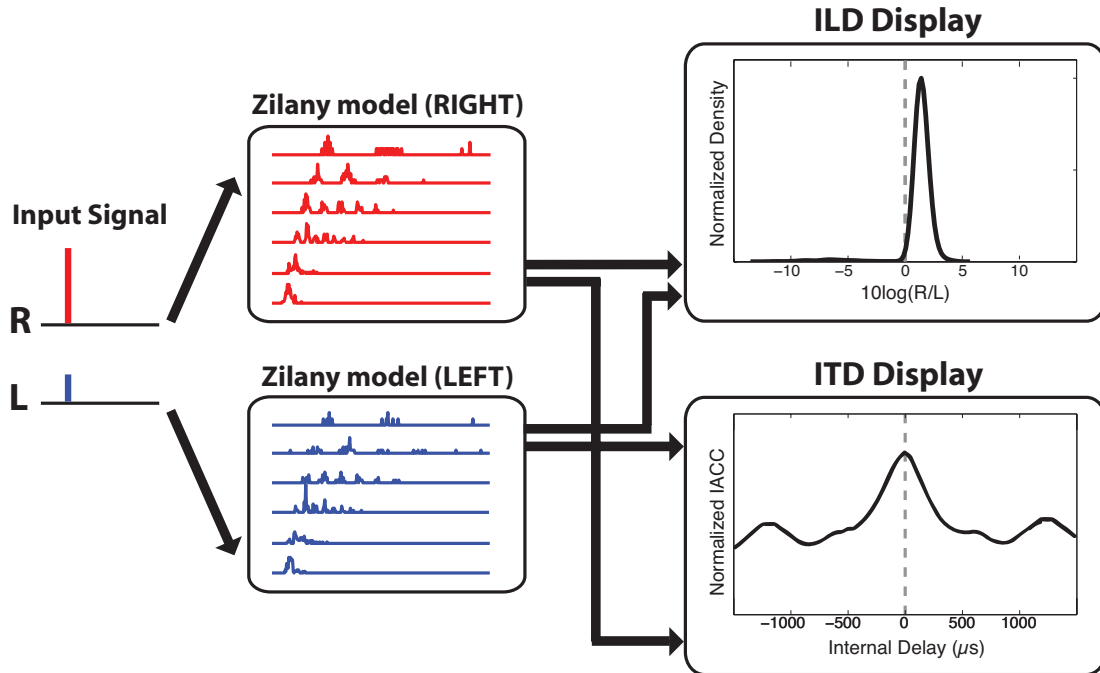


Figure 5.6: Summary of complete model output for a 120 μs dichotic click carrying a 10 dB right-favoring ILD (see also Fig. 5.3). The ITD display gives a peak at 0 μs , consistent with the 0 μs signal ITD. The ILD display gives a right-shifted peak, consistent with the right-favoring signal ILD. The ILD display output is proportional to, but not equal to, the signal ILD (see text).

“0 ILD” diagonal, reflecting the greater amplitude of the right PSTH. To obtain a final estimate of the ILD conveyed over the signal duration similar to that given in the ITD display (Fig. 5.4, upper panel), the MATLAB function `ksdensity` is used to obtain a smoothed estimate (by weighted kernel density estimation) of the distribution of right-to-left ratios represented in the scatter-plot (taken as $10 \times \log(R/L)$, Figure 5.5, upper panel). To minimize the influence of spontaneous firing, each ratio is weighted by the number of spikes within the integration window it is drawn from. Right-favoring ILD are thus given by a right-shifted peak in the density function (Re: vertical dashed line at 0 ILD), while left-favoring ILD are evidenced by a left-shifted peak. An important caveat is that the magnitude of “internal ILD” given by the ILD display is proportional, but not equal, to the ILD carried by the input signal. Although a “correction” factor could be applied to yield the signal ILD, there is no reason to suppose that the internal representation of ILD corresponds directly to the physical cue (indeed, the actual distribution of “internal ITD” in the mammalian system remains a matter of considerable debate, e.g., McAlpine *et al.*, 2001). This detail notwithstanding, the ILD display gives a

representation of the signal ILD in parallel with, but independent of, the ITD display (see Figure 5.6). In the following section, we consider ITD and ILD displays produced by a variety of signals. Like the ITD display, the ILD display will be given subsequently as the smoothed cross-frequency average only (Fig. 5.5, upper panel). Both displays will only be given when both cues were manipulated (see section 5.3.3).

5.3 Model demonstrations

Here, we consider the potential contributions of peripheral auditory processing and basic binaural comparison (as implemented in our model) to the psychophysical observations reported in Chapters 2-4. While precedence-like changes in binaural sensitivity are frequently attributed to synaptic inhibition or active “suppression” of post-onset information (e.g., Hafter *et al.*, 1988; Wickesburg and Oertel, 1990; Litovsky *et al.*, 1999; Xia *et al.*, 2010), Tollin and Henning (1999) and others (e.g., Hartung and Trahiotis, 2001; see section 5.3.2) have argued that some such effects can be explained by normal peripheral processing. This suggestion is based on the notion that encoding of the post-onset signal may be compromised by interaction of the basilar membrane’s response to the post-onset signal and its ongoing response to the signal onset. Here, we pursue this notion further, with the important addition of auditory nerve adaptation as provided by the Zilany *et al.* (2009) model.

5.3.1 Gabor click trains

Stimuli employed in Chapters 2 and 3 were trains of high-frequency (4 kHz) filtered impulses (“Gabor click trains”). Trains were presented with one of four different average inter-click intervals (1.25, 2.5, 5, or 10 ms ICI), and were lateralized by either ITD or ILD. Individual cues for each click were drawn randomly from a uniform distribution spanning $\pm 100 \mu\text{s}$ and $\pm 2 \text{ dB}$ for ITD and ILD conditions, respectively. The major findings of Chapters 2 and 3 were that (1) for brief inter-click intervals ($\leq 2.5 \text{ ms ICI}$), the ITD or ILD carried by the first click in a train exerted more influence on subjects’ perception than the ITD or ILD of

any subsequent click but also that (2) for ILD conditions, at ICI >1.25 ms, the average ILD over the stimulus duration was more salient than the ILD of any single click, while for ITD conditions, the ITD of the first click (onset ITD) was more salient than the average ITD for ICI <5 ms. The data of Chapter 3 additionally provided that the salience of onset cues (i.e., onset dominance) was reduced by temporal irregularity in the signal envelope (“jitter” of the ICI).

To assess the extent to which peripheral processing effects might account for these data, “dual display” model outputs were generated for a variety of Gabor click train stimuli. All stimuli were 16-click trains of 4-kHz filtered impulses with an average ICI of 2.5 ms. For both ITD and ILD, three different stimulus types were considered: (1) binaurally *static* trains (that were also temporally isochronous) carrying 100 μ s right-favoring ITD or 2 dB right-favoring ILD in all clicks, (2) temporally *isochronous* trains carrying a 100 μ s or 2 dB right-favoring onset cue and random post-onset cues (Chapter 2 stimuli) and (3) temporally *jittered* trains carrying a 100 μ s or 2 dB right-favoring onset cue and random post-onset cues (Chapter 3 stimuli). ITD and ILD displays for “Static” (type 1) stimuli give the maximum rightward peak shifts possible for the range of ITD and ILD cue values employed, against which “Isochronous” (type 2) and “Jittered” (type 3) displays may be compared (note that *onset* cues were identical across stimulus types). Chapter 2 data predict that the peak of the ITD display for the Isochronous ITD stimulus should remain significantly right-shifted, while the peak of the ILD display for the corresponding ILD stimulus should fall near 0, reflecting the near-zero average stimulus ILD. Chapter 3 data predict that the peaks of both ITD and ILD displays for Jittered stimuli should fall still nearer 0, reflecting an increased contribution of post-onset cues. Figure 5.7 gives ITD and ILD displays for all three ITD and ILD stimulus types. Data are consistent with the foregoing predictions: Relative to Static displays, the Isochronous ITD peak remains somewhat right-shifted (reflecting the strong contribution of the onset ITD), while the Isochronous ILD peak falls near 0 (consistent with the average stimulus ILD). Both ITD and ILD peaks fall near 0 for Jittered stimuli, in agreement with the reduced contribution of onset cues for Jittered stimuli evident in Chapter 3 data.

Model output for 2.5 ms ICI Gabor click trains

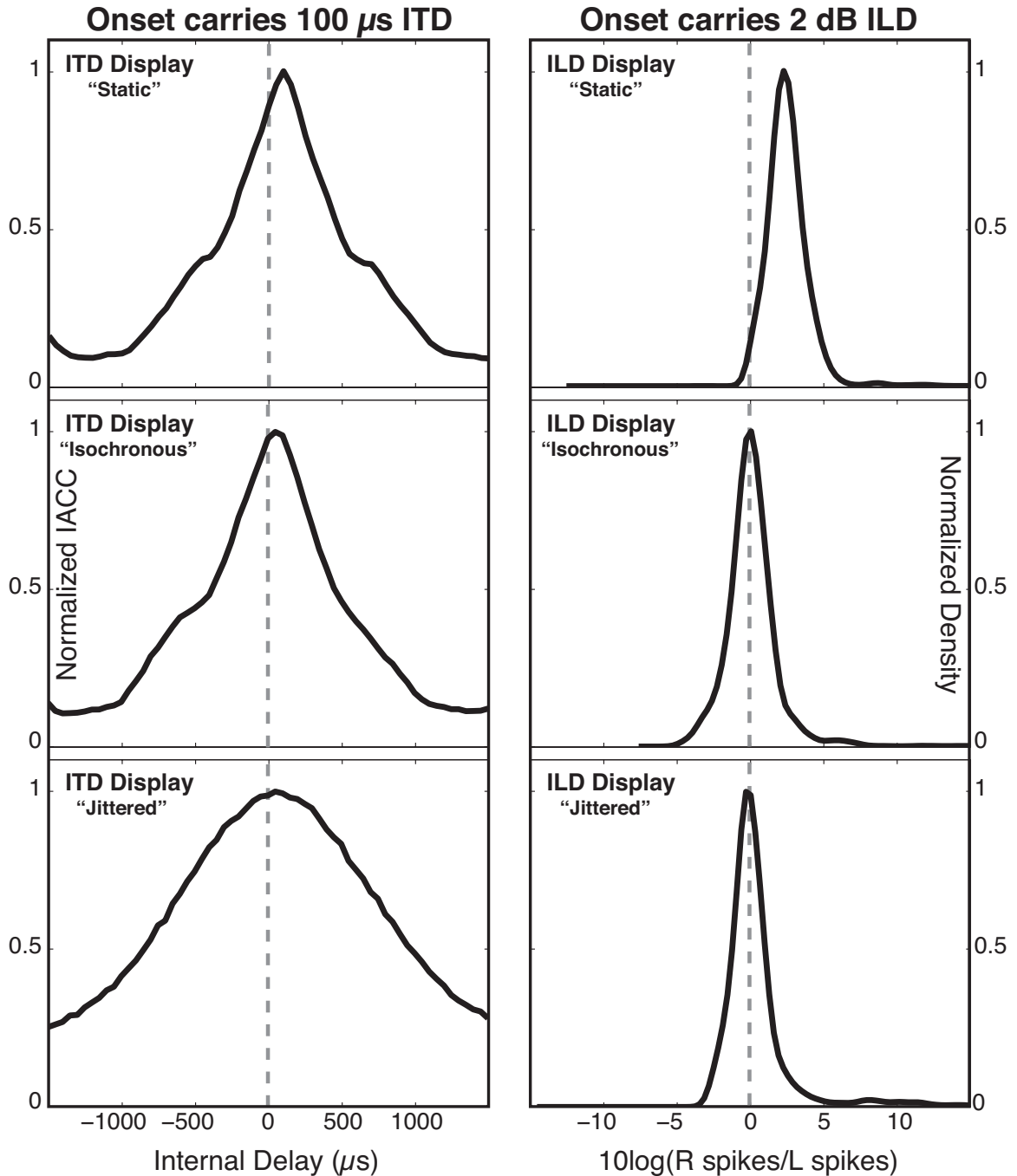


Figure 5.7: Model outputs for six different 2.5 ms ICI 16-click Gabor click train stimuli. Left column: ITD displays for click trains carrying 100 μ s ITD in the onset and (top) 100 μ s ITD in 15 subsequent clicks, (middle) random ITD from the range $\pm 100 \mu$ s in 15 subsequent clicks, or (bottom) random ITD from the range $\pm 100 \mu$ s in 15 subsequent clicks, with random inter-click intervals from the range 0.25 ms – 4.75 ms. Right column: ILD displays for click trains carrying 2 dB ILD in the onset and (top) 2 dB ILD in 15 subsequent clicks, (middle) random ILD from the range ± 2 dB in 15 subsequent clicks, or (bottom) random ITD from the range $\pm 100 \mu$ s in 15 subsequent clicks, with random inter-click intervals from the range 0.25 ms – 4.75 ms. The contribution of the onset ITD and ILD to each composite display may be assessed by comparing middle and bottom panels to the top panel within each column (see text).

5.3.2 Lead-lag click pairs: Baseline and Buildup conditions

Following on the work of Tollin and Henning (e.g., 1999), Hartung and Trahiotis (2001) argued that the precedence effect as classically demonstrated with successive transient “lead-lag” stimuli (Wallach *et al.*, 1949; see Chapter 4) could be explained by interaction of the lead and lag on the basilar membrane, leading to reduced peripheral and (thus) binaural encoding of the lag. Their report, however, only considered such effects for lead-lag delays expected to produce significant peripheral overlap (up to 3 ms), and then only for lead-lag stimuli lateralized by ITD. The authors did not offer an explanation for typical ITD-based echo thresholds over the range 5-10 ms (see Blauert, 1997, Litovsky *et al.*, 1999; Chapter 4), or for ILD-based echo thresholds of any degree, other than to conclude that the precedence effect at “long interstimulus intervals... cannot be explained adequately by considering only peripheral processing.” Although we do not wish to discredit Hartung and Trahiotis (2001), or the several investigators who have suggested more central mechanisms of precedence involving synaptic inhibition (e.g., Litovsky and Yin, 1998; Pecka *et al.*, 2007; Xia, 2010), we believe it is critical to consider, in addition to cochlear mechanics, the potential contributions of peripheral auditory nerve adaptation to the precedence effect before more elaborate origins should be assumed. The present section thus considers model outputs for precedence stimuli employed in Chapter 4.

5.3.2.1 Baseline ITD stimuli

The Baseline ITD stimulus employed in Experiment 1 of Chapter 4 has been employed in dozens of other investigations, and typically produces an echo threshold between 5 and 10 ms. Across the 10 subjects tested in Experiment 1, the average echo threshold was just below 9 ms. Thus, the present question is whether our model predicts a reduced representation of the lag ITD for a lead-lag stimulus with a lead-lag delay of ~9 ms. To address this question, lead-lag stimuli of Experiment 1 (120 μ s dichotic clicks, +300 μ s ITD in lead, -300 μ s in lag) of six different lead-lag delays (spaced logarithmically from 1 to 32 ms) were passed through the complete model. A cross-delay ITD display is given in Figure 5.8.

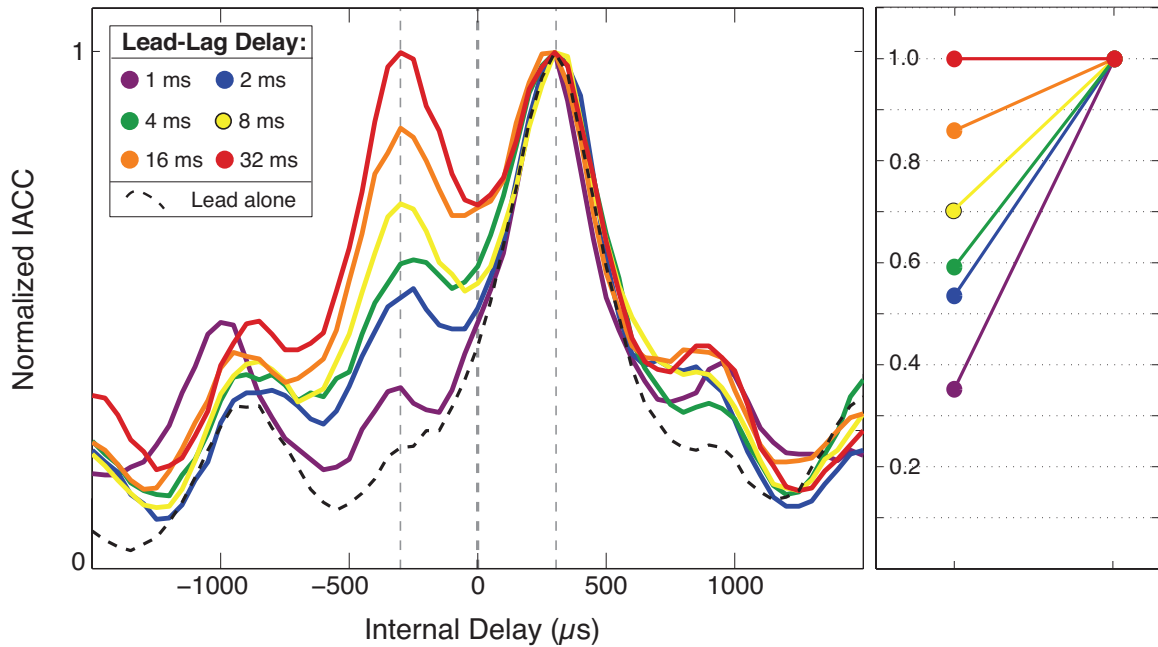


Figure 5.8: Left panel: ITD display for Baseline ITD lead-lag stimuli employed in Experiment 1 of Chapter 4. Model outputs are illustrated for six lead-lag delays (1-32 ms, see legend). The lead ITD (+300 μ s) is represented with unchanged fidelity across lead-lag delays (compare against display for lead alone, dashed black line). The representation of the lag ITD (-300 μ s), in contrast, is systematically degraded as a function of lead-lag delay. Only at the 32 ms lead-lag delay are the lead and lag represented with *equal* fidelity. Right panel: Normalized amplitudes of primary and secondary peaks in the cross-correlation function (the 1000 μ s secondary peak at 1 ms lead-lag delay is well outside the range of physiologically plausible ITD and was thus excluded).

Model output yields a clear precedence effect: While the lead ITD (+300 μ s) is represented equally at all lead-lag delays (compare to lead alone, dashed black line), the representation of the lag ITD (-300 μ s) is systematically reduced as a function of lead-lag delay. At 8 ms lead-lag delay, near the behavioral echo threshold, the secondary IACC peak in the ITD display (at -300 ms) is only 70% of the primary peak height. Although it is not clear what the correspondence between relative “peak heights” and perception should be, the model output is qualitatively consistent with the behavioral data. Considering auditory nerve adaptation thus appears to significantly extend the range of lead-lag delays over which peripheral processing can account for ITD-based precedence.

5.3.2.2 Baseline ILD stimuli

The peripheral model of Hartung and Trahiotis (2001) did not address precedence effects for stimuli lateralized by ILD (e.g., Zurek, 1980, Krumbholz and Nobbe, 2002; Chapter 4), assumedly because the binaural processor

employed in that study produced estimates of lateralization based on ITD alone. Nonetheless, a precedence effect for stimuli lateralized by ILD alone clearly exists. Interestingly, the effect is significantly weaker (mean echo threshold just below 5 ms) than that for analogous ITD stimuli (a primary conclusion of this dissertation). A physiological explanation for this disparity is not obvious, especially in the context of models of precedence processing based on synaptic inhibition, which is ubiquitous at all levels of the central auditory pathway (e.g., Wickesburg and Oertel, 1990; Pecka *et al.*, 2007; Xia *et al.*, 2010). Moreover, ILD processing should not be particularly sensitive to phase distortions resulting from lead-lag interactions on the basilar membrane (a certain contributor to peripheral precedence for ITD). Nonetheless, auditory nerve adaptation as implemented in the present model could theoretically produce a precedence effect for ILD. To probe for such an effect, Baseline ILD stimuli from Experiment 1 of Chapter 4 (of the same 6 lead-lag delays illustrated in Fig. 5.8) were passed through the model. A cross-delay ILD display is given in Figure 5.9. Model outputs are

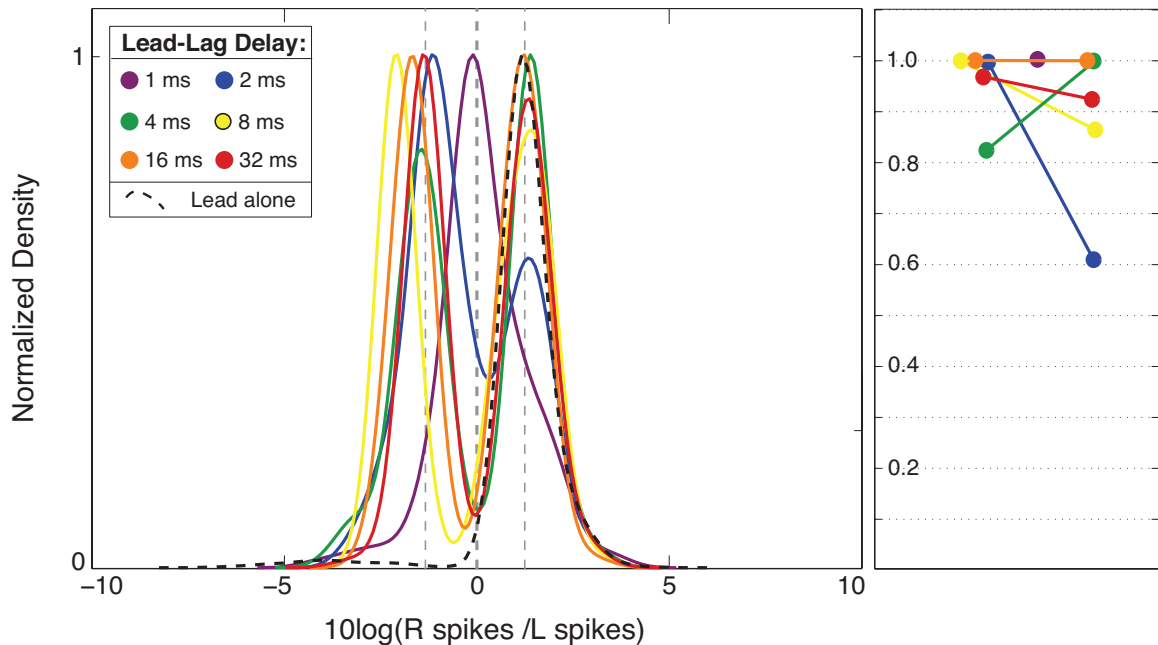


Figure 5.9: Left panel: ILD display for Baseline ILD lead-lag stimuli employed in Experiment 1 of Chapter 4. Model outputs are illustrated for six lead-lag delays (1-32 ms, see legend). Model outputs are considerably less systematic than ITD display outputs for analogous ITD stimuli (Fig. 5.8). Nonetheless, at lead-lag delays ≥ 4 ms, two prominent peaks (one right-shifted, one left-shifted) are evident in the ILD display (compare with dashed black line, lead alone, and dashed vertical lines representing “lead alone” and “lag alone”). Right panel: Normalized amplitudes of primary and secondary peaks in the cross-correlation function.

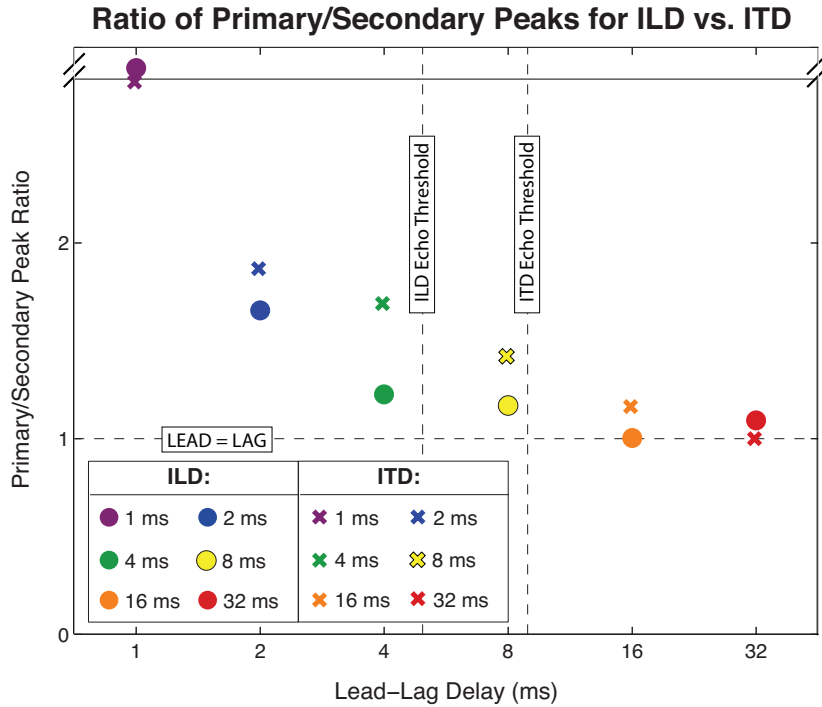


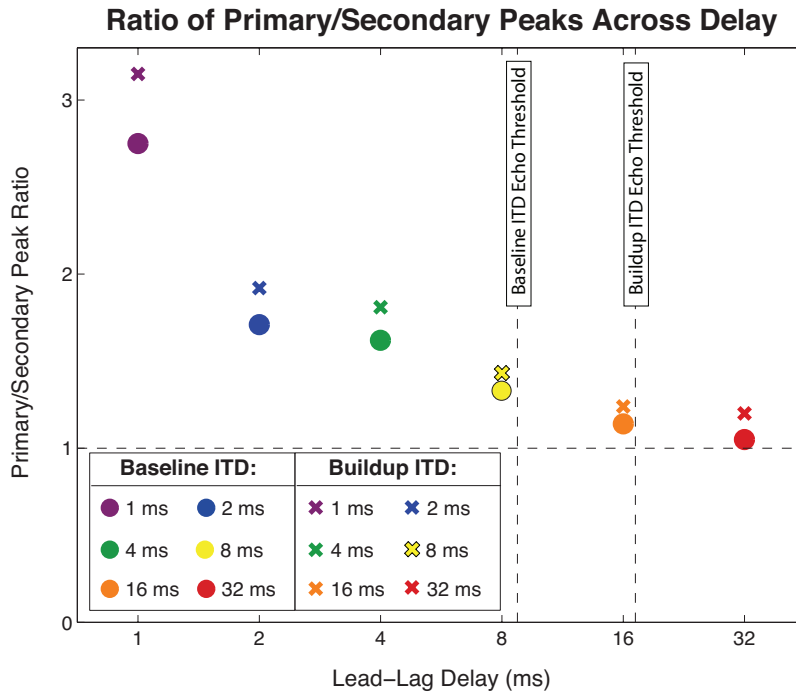
Figure 5.10: Primary/Secondary peak ratios taken as the normalized height of the largest peak in the ITD (crosses) or ILD (circles) displays divided by the height of the second largest peak. The ratio for the Baseline ITD condition (crosses) exceeds that for the Baseline ILD condition (circles) at lead lag delays over the range 2-16 ms, most notably at 4 and 8 ms. Behavioral echo thresholds are plotted for visual comparison. The break in the ordinate denotes the undefined ratio for ILD at 1 ms lead-lag delay, resulting from the absence of a secondary peak in the display.

considerably less systematic than outputs for Baseline ITD stimuli (previous section; Fig. 5.8). Nonetheless, at lead-lag delays ≥ 4 ms, two prominent peaks (one right-shifted, one left-shifted) are evident. Significantly, there appears to be less disparity between the “primary” and “secondary” peaks than in the ITD display for Baseline

ITD stimuli. Primary/secondary peak ratios – an ostensibly reasonable proxy for the magnitude of precedence – are plotted as a function of lead-lag delay for Baseline ITD and ILD conditions in Figure 5.10. The ratio is generally higher for ITD than ILD, particularly at 4 and 8 ms lead-lag delay – the range over which the behavioral echo thresholds for both cues emerge. The present model thus predicts “baseline” precedence effects for both ITD and ILD of an order consistent with behavioral measurements, and on the basis of peripheral processing alone (i.e., without synaptic inhibition).

5.3.2.3 Buildup ITD stimuli

As detailed in section 5.2.1, power-law adaptation has no fixed time constants. The degree of adaptation thus generally increases with stimulus duration for durations up to several seconds (and even beyond, see Zilany *et al.*,



5.11: Primary/secondary peak ratios drawn from ITD displays for Baseline and Buildup ITD stimuli (see text). Note the slightly different scale on the ordinate and slightly different Baseline ITD data Re: Fig. 5.10. The model produces slightly different data on each run (much like a human subject).

2009). Recognizing this attribute of our model's front-end, we were curious to evaluate the degree of "precedence" in model outputs for Buildup versus Baseline test stimuli (see Chapter 4). Although it has been suggested that the mechanism of buildup is likely to be cortical (e.g., Blauert, 1997; Sanders *et al.*, 2011), multi-second

adaptation in the auditory nerve could theoretically also contribute to buildup. Thus, to test for a buildup-like effect resulting from peripheral processing, we computed the primary/secondary peak ratios (cf. Fig. 5.10) from the ITD displays for Baseline and Buildup ITD stimuli across lead-lag delays 1-32 ms. Ratios are plotted for Baseline (circles) and Buildup (crosses) ITD conditions in Figure 5.11. A slight increase in the ratio is evident for the Buildup condition at all lead-lag delays. A paired *t*-test indicated that Buildup ratios were significantly higher than Baseline ratios ($t(5) = 3.98, p < 0.05$). While these data thus predict a significant buildup-like effect resulting from *peripheral processing alone*, we note that the magnitude of the effect appears to be relatively small. Toward this point, we reiterate that the buildup effect apparently consists of a minor enhancement of the precedence effect *per se* (fusion + localization dominance), and a more substantial enhancement of fusion alone, which may be consistent with a minor contribution of peripheral processing and a more substantial contribution of more central processing.

5.3.3 Binaural acoustic recordings

In the preceding sections, we have applied our “dual display” binaural model to the synthesized headphone stimuli used in Chapter 2-4 investigations, primarily for qualitative comparison of model outputs and observed patterns in the psychophysical data. In this final section, we consider the accuracy with which the model is able to localize an actual sound source in a real echoic environment. While normal-hearing human listeners are quite good at localization in echoic environments, hearing-impaired listeners and artificial intelligence systems are not (see Chapter 1, and Blauert 1997). To assess the degree to which *normal* peripheral processing (including auditory nerve adaptation) and “dual display” binaural comparison model might enable extraction of veridical ITD and ILD cues, a balloon pop (approximate impulse) was recorded through the ears of a binaural manikin positioned at the center of a 2.8 x 1.4 x 2 m sound booth. The balloon was positioned 1 m from the manikin at +90° azimuth. Resultant ITD and ILD displays are illustrated in Figure 5.12. The ITD display gives a prominent peak at +670 μ s internal delay, corresponding to the veridical ITD. The ILD display, in contrast, gives a prominent peak at 0 (and a much lesser right-favoring peak), reflecting the dominant influence of binaurally diffuse reflected sound. Thus, a listener (or computer system) would ideally depend on the ITD (and not the ILD) for localization of the signal. The implications of this and the foregoing results are discussed in the final (concluding) Chapter 6.

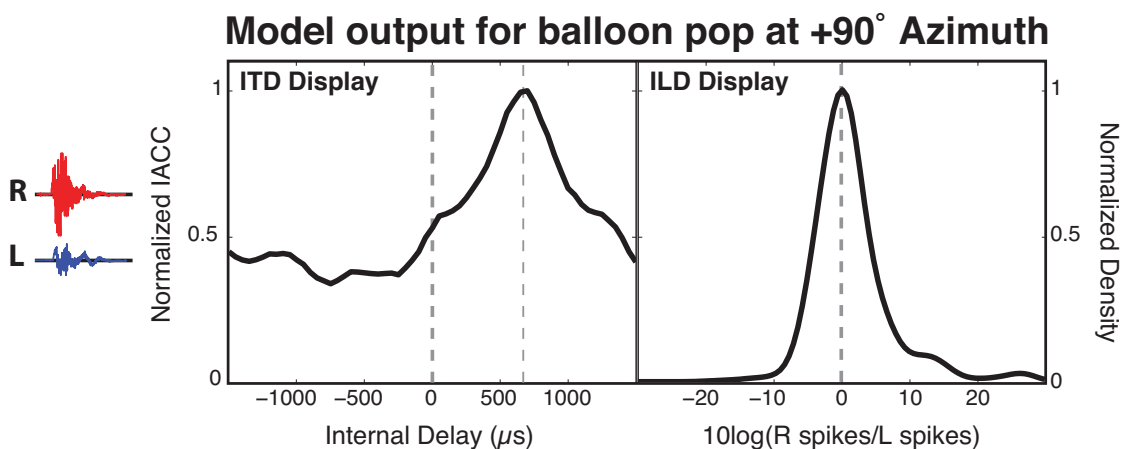


Figure 5.12: ITD and ILD displays for a binaural recording of a balloon pop (illustrated at left; see text).

6

Concluding remarks

Sound localization is one of several evolved faculties for spatial orientation that is ubiquitous among vertebrates. While vision is the most critical sense for navigation in many species, including humans (Schone, 1986), sound localization is critical for both environmental awareness and communication. Although the basic physical cues sub-serving sound localization have been identified for over a century (Rayleigh, 1907; Chapter 1) and the basic neural substrates of sound localization have been identified for nearly half a century (e.g., Boudreau and Tsuchitani, 1968; Goldberg and Brown, 1969), there remains much to learn about the psychological and physiological underpinnings of our ability to localize (and segregate) sound sources across the effectively endless variety of acoustic environments we may encounter. The investigations described in the foregoing chapters were designed to elucidate aspects of sensitivity to *binaural* cues for sound localization in ordinary listening environments (e.g., rooms). Key findings and their significance are considered below.

In Chapters 2-4, a variety of stimuli and listening paradigms were employed to assess the temporal dynamics of listeners' sensitivity to ITD and ILD. These studies were conducted in part because the cues themselves become dynamic – even for stationary sources – shortly after the signal onset in ordinary listening environments (due to signal reflections; see Chapter 1). Reliable detection of a cue (or a *change* in a cue) corresponding to an actual sound source thus depends critically on the temporal dynamics of binaural sensitivity: If sensitivity were constant over the signal duration, then veridical onset cues and spurious post-onset cues (corresponding to signal reflections) would contribute equally to spatial judgments, leading to persistent errors in localization (and source segregation). On the other hand, if listeners were sensitive only to the veridical onset cues, detection of changes in the post-onset cues – potentially including cue changes corresponding to real changes in the environment – would become impossible.

In general, the data of Chapters 2, 3, and particularly 4 suggest that the binaural system “solves” the problem of accurate localization versus adaptive monitoring of the acoustic environment via a two-cue mechanism. Specifically, for rapidly modulated signals like those typical of environments that produce signal reflections, data suggest that the ITD at stimulus onset determines the initial localization (note that at onset, prior to the arrival of signal reflections, the ILD should naturally agree with the ITD): This percept tends to persist despite drastic fluctuations in the ongoing ITD, which are also typical of signals in environments that produce signal reflections (e.g., Rakerd and Hartmann, 1985; cf. random post-onset ITD of Chapter 2 and 3 stimuli, “Breakdown ITD” stimuli of Chapter 4). In contrast, significant fluctuations in post-onset ILD, which are *not* typical of signals in environments that produce signal reflections, readily influence perception (e.g., Rakerd and Hartmann, 1985; “cue-averaging” data of Chapters 2 and 3; “Breakdown ILD” data of Chapter 4), consistent with the likelihood that such fluctuations correspond to real changes in the environment (e.g., a new source within the same spectral band). Such discrete contributions to perception by ITD and ILD are not obvious in free field data from studies of, for example, buildup and breakdown effects (e.g., Clifton and Freyman, 1989) or temporal weighting in sound localization (Stecker and Hafter, 2002); nonetheless, headphone data presented in this dissertation and in a number of other reports seem to *require* that these free field phenomena depend on (at least) two temporally distinct channels of sensitivity.

The modeling data of Chapter 5 suggest that the temporal dynamics of processing discussed hereto may be largely explained by peripheral auditory processing (external, middle, and inner ear mechanical filtering, and adaptation in auditory transduction). Although the model employed (particularly its front-end, which yields the majority of interesting effects reported) was built on a foundation of reputable physiological data (see Zilany *et al.*, 2009), the task remains to verify model predictions empirically in a living system. One approach would be to collect data from a sufficiently large number of neurons in an appropriate animal model to approximate the “population” response generated by the model. A quite

different but exciting alternative would be to conduct experiments in a system *lacking* the components incorporated in the peripheral model.

Bilateral CI users, discussed in several previous sections of this dissertation, hear through direct electrical stimulation of the auditory nerve by an electrode array inserted into the cochlea. Peripheral processing in CI users is thus drastically different than in normal hearing: Cochlear mechanics (including limits on temporal resolution related to cochlear mechanics) and hair cell synaptic fatigue (a major contributor to adaptation observed in the auditory nerve, see Zilany *et al.*, 2009) are irrelevant, as those systems are bypassed entirely. Instead, auditory nerve firing is driven by the externally worn CI processor using one of several different processing strategies (not reviewed here). Although in bilateral CI users, the processors for the two ears are typically not synchronized, the devices can be plugged into research processors in the laboratory that effectively synchronize inputs to the two ears, restoring true binaural stimulation (e.g., Laback and Majdak, 2008; Litovsky *et al.*, 2010). Bilateral CI users thus offer an unusual “control group” for the study of peripheral versus central (e.g. synaptic inhibitory) effects on the temporal dynamics of binaural sensitivity. While many bilateral CI users, particularly those with early onset of deafness or prolonged deafness leading to neural atrophy, are insensitive to ITD and thus unable to participate in ITD experiments, a subset of listeners, typically with adult-onset deafness, demonstrate ITD sensitivity of the same order as normal-hearing listeners when inputs are synchronized in the laboratory (most all bilateral CI users are sensitive to ILD even without processor synchronization, e.g., Grantham *et al.*, 2008; Litovsky *et al.*, 2010).

Therefore, future studies should test bilateral CI users in a psychophysical paradigm using stimulus conditions similar to those tested in Chapter 4 – conditions known to produce the precedence effect in normal-hearing listeners, and hypothesized to do so on the basis of *peripheral* processing. If the normal precedence effect indeed depends on peripheral processing and particularly on its consequences for the encoding of ITD, then bilateral CI users should be expected to essentially lack a precedence effect. If, alternatively, the precedence

effect depends on central mechanisms, bilateral CI patients with otherwise “normal” binaural sensitivity should be expected to experience a normal or near-normal precedence effect. A combination of mechanisms could of course produce data intermediate to these two extremes. A preliminary report of precedence experiments conducted in the free field with bilateral CI users suggested a large range of variability in lead-lag fusion (Seeber and Hafter, 2008). The outcomes of experiments such as these offer to provide key insight on the contributions of peripheral processing to the temporal dynamics of binaural sensitivity in the normal hearing system (present report) and, also, to elucidate challenges to the improvement of CI patient outcomes. Although bilateral CI users perform somewhat better than monaural users in difficult environments, bilateral and monolateral CI users alike persistently report difficulties in complex listening environments (e.g., social gatherings, Loizou *et al.*, 2009; or classrooms, Neuman *et al.*, 2012; see section 1.5). While the inevitable lack of ITD sensitivity without processor synchronization and listeners’ consequent reliance on ILD – which we have shown to be a relatively poor localization cue in echoic environments – must contribute to the reported difficulties, it may be that even once a clinical strategy for processor synchronization is devised, the absence of normal peripheral processing will preclude a normal precedence effect. The task would then become to engineer a processing strategy that directly attenuates spurious (post-onset) cues while preserving veridical (onset) cues.

Appendix A

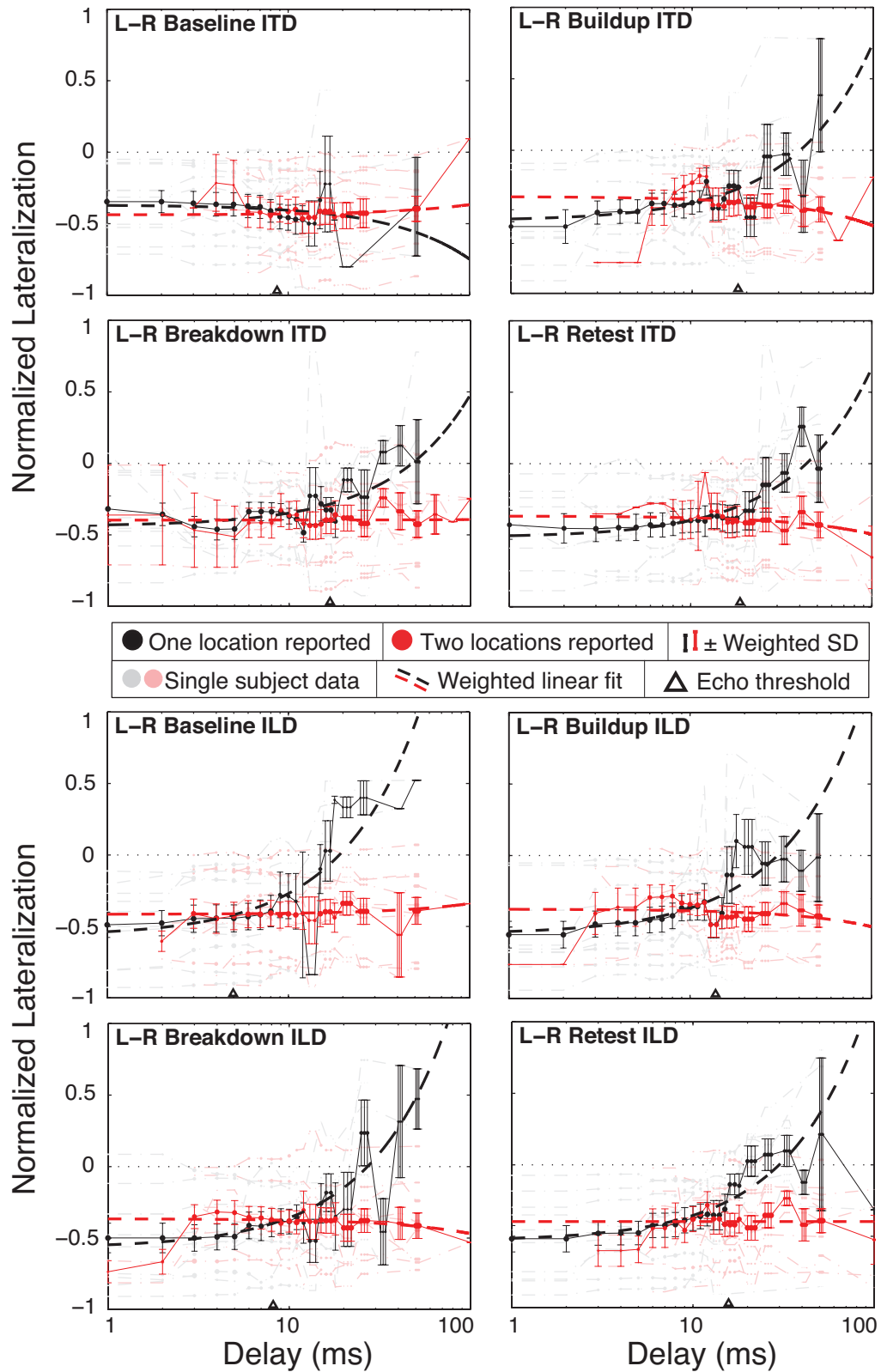


Figure A.1: Lateralization responses for left-lead, right-lag trials from Chapter 4, Experiment I (see text).

Figures A.1, A.2, and A.3 give lateralization responses for left-lead, right-lag trials across ITD and ILD conditions within in the three experiments of Chapter 4. The “One Location” versus “Two Locations” dimension of responses on these trials contributed to the measurement of echo thresholds reported in the main text, but the left-right dimension (lateralization dominance aspect) of these responses was not considered. One might presume that with a left-favoring lead stimulus and with the instruction to point to the left-most location perceived in the event two locations were perceived, subjects would point to the left-favoring lead across all lead-lag delays. Nonetheless, the data Figs. A.1-3 demonstrate that lateralization for “One Location” trials was reduced at long lead-lag delays across experiments, suggesting influence of the right-favoring lag on responses and consistent with responses in right-lead, left-lag trials (see Chapter 4).

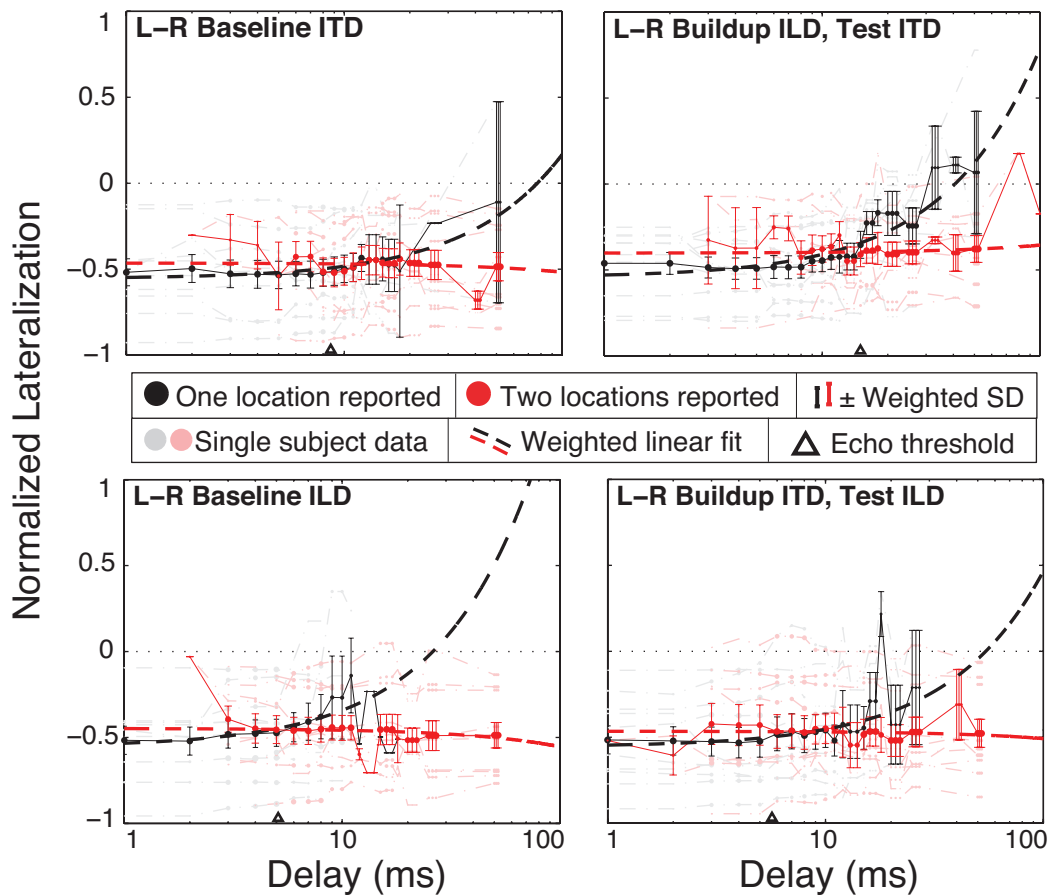


Figure A.2: Lateralization responses for left-lead, right-lag trials of Experiment II, Chapter 4 (see text).

Appendix B

Details of Zilany *et al.* (2009) model implementation

-For original description of model, see:

Zilany, M. S. A., Bruce, I. C., Nelson, P.C., and Carney, L.H. (2009). "A phenomenological model of the synapse between the inner hair cell and auditory nerve: Long-term adaptation with power-law dynamics," *J. Acoust. Soc. Am.* **126**, 2390-2412.

-For model code, see:

Laurel Carney's laboratory website (as of July 2012).
<http://www.urmc.rochester.edu/bme/people/faculty/bio/project.cfm?id=229>

-Model implementation in present study:

>>"Approximate" implementation (versus "actual" implementation) was used at the recommendation of M. Zilany (personal communication). The two options only produce substantively different output for very long-duration signals, and the "actual" option takes several times longer to run (the "approximate" option can still take several days).

>>Input signals were passed through outer and middle ear filters into the inner hair cell model component.

>>Outputs of the inner hair cell model were used as the inputs to the hair cell synapse model component (in which the adaptation algorithm is also implemented).

>>Nerve fibers at synapse were selected to be medium spontaneous rate fibers (model option "type 2").

>>Fiber center frequencies were set at one of six different values (200, 400, 800, 1600, 3200, 6400 Hz).

>>The model was repeated 100 times at each center frequency to give an effective population of 100 nerve fibers per center frequency.

>>The model was run for left and right ears independently; the ultimate outputs, the composite PSTHs for each frequency channel in each ear, were saved and later passed through the "dual-display" model as described in Chapter 5.

Bibliography

- Akeroyd, M.A. and Bernstein, L.R. (2001). "The variation across time of sensitivity to interaural disparities: Behavioral measurements and quantitative analyses," *J. Acoust. Soc. Am.* **110**, 2516-2526.
- Akeroyd, M.A. (2001). "A Binaural Cross-correlogram Toolbox for MATLAB," <http://www.ihr.mrc.ac.uk/products/display/matlab>. Accessed Dec. 2011.
- von Békésy, G. (1930). "Zur theorie des horens. Über das richtunschoren bei einer zeitdifferenz oder lautstarkenungleichhet der beiderseitigen schalleinwirkungen," *Physik. Z.* **31**, 824-835, 857-868.
- Berg, B.G. (1989). "Analysis of weights in multiple observation tasks," *J. Acoust. Soc. Am.* **86**, 1743-1746.
- van Bergeijk, W.A. (1962). "Variation on a theme of Békésy: A model of binaural interaction," *J. Acoust. Soc. Am.* **34**, 1431-1437.
- Bernstein, L.R. and Trahiotis, C. (1994). "Detection of interaural delay in high-frequency sinusoidally amplitude-modulated tones, two-tone complexes, and bands of noise," *J. Acoust. Soc. Am.* **95**, 3561-3567.
- Bernstein, L.R. and Trahiotis, C. (2002). "Enhancing sensitivity to interaural delays at high frequencies by using 'transposed stimuli'," *J. Acoust. Soc. Am.* **112**, 1026-1036.
- Bishop, C.W., London, S., Miller, L.M. (2011). "Visual influences on echo suppression," *Curr. Biol.* **21**, 221-225.
- Blauert, J. (1997). *Spatial Hearing: The Psychophysics of Human Sound Localization, Revised Edition* (The MIT Press, Cambridge, MA).
- Boudreau, J.C. and Tsuchitani, C. (1968). "Binaural interaction in the cat superior olive S segment," *J. Neurophysiol.* **31**, 442-454.
- Brandewie, E. and Zahorik, P. (2010). "Prior listening in rooms improves speech intelligibility," **128**, 291-299.
- Breebaart, J., van de Par, S., and Kohlrausch, A. (2001a). "Binaural processing model based on contralateral inhibition. I. Model structure," *J. Acoust. Soc. Am. J. Acoust. Soc. Am.* **110**, 1074-1088.
- Breebaart, J., van de Par, S., and Kohlrausch, A. (2001b). "Binaural processing model based on contralateral inhibition. II. Dependence on spectral parameters," *J. Acoust. Soc. Am.* **110**, 1089-1104.
- Bregman, A.S. (1994). *Auditory Scene Analysis: The Perceptual Organization of Sound*. (The MIT Press, Cambridge, MA).
- Brown, A. D. and Stecker, G. C. (2010). "Temporal weighting of interaural time and level differences in high-rate click trains," *J. Acoust. Soc. Am.* **128**, 332-341.
- Buell, T.N., Trahiotis, C., and Bernstein, L.R. (1994). "Lateralization of bands of noise as a function of combinations of interaural intensive differences, interaural temporal differences, and bandwidth," *J. Acoust. Soc. Am.* **95**, 1482-1489.
- Buell, T.N., Griffin, S.J. and Bernstein, L.R. (2008). "Listeners' sensitivity to 'onset/offset' and 'ongoing' interaural delays in high-frequency,

- sinusoidally amplitude-modulated tones," *J. Acoust. Soc. Am.* **123**, 279-294.
- Burger, R.M. and Pollak, G.D. (2001). "Reversible inactivation of the dorsal nucleus of the lateral lemniscus reveals its role in the processing of multiple sound sources in the inferior colliculus of bats," *J. Neurosci.* **21**, 4830-4843.
- Carney, L. (2012). Laboratory website. <http://www.urmc.rochester.edu/bme/people/faculty/bio/project.cfm?id=229>
- Carr, C.E. and Konishi, M. (1990). "A circuit for detection of interaural time differences in the brain stem of the barn owl," *J. Neurosci.* **10**, 3227-3246.
- Chase, S.M. and Young, E.D. (2005). "Limited segregation of different types of sound localization information among classes of units in the inferior colliculus," *J. Neurosci.* **25**, 7575-7585.
- Clifton, R.K. (1987). "Breakdown of echo suppression of the precedence effect," *J. Acoust. Soc. Am.* **82**, 1834-1835.
- Clifton, R.K., and Freyman, R.L. (1989). "Effect of click rate and delay on breakdown of the precedence effect," *Percept. Psychophys.* **46**, 139-145.
- Clifton, R.K., Freyman, R.L., Litovsky, R.Y., and McCall, D. (1994). "Listeners' expectations about echoes can raise or lower echo threshold," *J. Acoust. Soc. Am.* **95**, 1525-1533.
- Colburn, H.S. (1973). "Theory of binaural interaction based on auditory-nerve data. I. General strategy and preliminary results on interaural discrimination," *J. Acoust. Soc. Am.* **54**, 1458-1470.
- Colburn, H.S. (1977). "Theory of binaural interaction based on auditory-nerve data. II. Detection of tones in noise," *J. Acoust. Soc. Am.* **61**, 523-533.
- Djelani, T. and Blauert, J. (2001). "Investigations into the build-up and breakdown of the precedence effect," *Acta Acustica*, **87**, 253-261.
- Dunai, L. and Hartmann, W.M. (2011). "Frequency dependence of interaural time difference thresholds in human listeners," *J. Acoust. Soc. Am.* **129**, 2485.
- Formby, C., and Muir, K. (1988). "Modulation and gap detection for broadband and filtered noise signals," *J. Acoust. Soc. Am.* **84**, 545-550.
- Freyman, R.L., Zurek, P.M., Balakrishnan, U., and Chiang, Y.C. (1997). "Onset dominance in lateralization," *J. Acoust. Soc. Am.* **101**, 1649-1659.
- Getzmann, S. (2004). Spatial discrimination of sound sources in the horizontal plane following an adapter sound. *Hear. Res.* **191**, 14-20.
- Goldberg, J.M. and Brown, P.B. (1969). "Response of binaural neurons of dog superior olivary complex to dichotic tonal stimuli: Some physiological mechanisms of sound localization," *J. Neurophysiol.* **32**, 613-636.
- Goupell, M.J., Laback, B., and Majdak, P. (2009). "Enhancing sensitivity to interaural time differences at high modulation rates by introducing temporal jitter," *J. Acoust. Soc. Am.* **126**, 2511-21.
- Grantham, W.D. (1996). "Left-right asymmetry in the buildup of echo suppression in normal-hearing adults," *J. Acoust. Soc. Am.* **99**, 1118-1123.
- Grantham, W.D., Ashmead, D.H., Ricketts, T.A., Haynes, D.S., and Labadie, R.F. (2008). "Interaural time and level difference thresholds for

- acoustically presented signals in post-lingually deafened adults fitted with bilateral cochlear implants using CIS+ processing," *Ear Hearing* **29**, 33-44.
- Gulick, W.L., Gescheider, G.A., and Frisina, R.D. (1989). *Hearing: Physiological Acoustics, Neural Coding, and Psychoacoustics* (Oxford University Press, New York, NY).
- Haftner, E.R. and Jeffress, L.A. (1968). "Two-image lateralization of tones and clicks," *J. Acoust. Soc. Am.* **44**, 563-569.
- Haftner, E.R., and Dye, R.H. (1983). "Detection of interaural differences of time in trains of high-frequency clicks as a function of interclick interval and number," *J. Acoust. Soc. Am.* **73**, 644-651.
- Haftner, E.R., Dye, R.H., and Wenzel, E. (1983). "Detection of interaural differences of intensity in trains of high-frequency clicks as a function of interclick interval and number," *J. Acoust. Soc. Am.* **73**, 1708-1713.
- Haftner, E.R., and Buell, T.N. (1990). "Restarting the adapted binaural system," *J. Acoust. Soc. Am.* **88**, 806-812.
- Haftner, E.R., Dye Jr, R.H., Wenzel, E., and Knecht, K. (1990). "The combination of interaural time and intensity in the lateralization of high-frequency complex signals," *J. Acoust. Soc. Am.* **87**, 1702-1708.
- Hartmann, W.M., and Constan, (2002). "Interaural level differences and the level-meter model," *J. Acoust. Soc. Am.* **112**, 1037-1043.
- Hartung, K. and Trahiotis, C. (2001). "Peripheral auditory processing and investigations of the 'precedence effect' which utilize successive transient stimuli," *J. Acoust. Soc. Am.* **110**, 1505-1513.
- Hancock, K.E., Noel, V., Delgutte, B. (2010). "Neural coding of ITD with bilateral cochlear implants: effect of auditory experience," *Assoc. Res. Otolaryngol. Abstr.* **33**, 795.
- Henning, G.B. (1974). "Detectability of interaural delay in high-frequency complex waveforms," **55**, 84-90.
- van Hoesel, R.J.M. (2004). "Exploring the benefits of bilateral cochlear implants," *Audio. Neurotol.* **9**, 234-246.
- van Hoesel, R.J.M. (2008(a)). "Observer weighting of level and timing cues in bilateral cochlear implant users," *J. Acoust. Soc. Am.* **124**, 3861-3872.
- van Hoesel, R.J.M. (2008(b)). "Binaural jitter with cochlear implants, improved interaural time-delay sensitivity, and normal hearing," *Proc. Nat. Acad. Sci.* **105**, E51.
- Keen, R. and Freyman, R.L. (2009). "Release and re-buildup of listeners' models of auditory space," *J. Acoust. Soc. Am.* **125**, 3243-3252.
- Kiang, N.Y.S., Watanabe, T., Thomas, E.C., and Clark, L.F. (1965). *Discharge Patterns of Single Fibers in the Cat's Auditory Nerve* (MIT Press, Cambridge, MA), Research Monograph 35.
- Knudsen, E.I. and Konishi, M. (1978). "A neural map of auditory space in the barn owl," *Science* **200**, 795-797.
- Krumbholz, K. and Nobbe, A. (2002). "Buildup and breakdown on echo suppression for stimuli presented over headphones—the effects of interaural time and level differences," *J. Acoust. Soc. Am.* **112**, 654-663.

- Laback, B., and Majdak, P. (2008). "Binaural jitter improves interaural time-difference sensitivity of cochlear implantees at high pulse rates," *Proc. Nat. Acad. Sci.* **105**, 814-817.
- Levitt, H. and Rabiner, L.R. (1967). "Binaural release from masking for speech and gain in intelligibility," *J. Acoust. Soc. Am.* **42**, 601-608.
- Lindemann, W. (1986). "Extension of a binaural cross-correlation model by contralateral inhibition. I. Simulation of lateralization for stationary signals," *J. Acoust. Soc. Am.* **80**, 1608-1622.
- Litovsky, R.Y., Colburn, H.S., Yost, W.A., and Guzman, S.J. (1999). "The precedence effect," *J. Acoust. Soc. Am.* **106**, 1633-1654.
- Litovsky, R.Y. and Shinn-Cunningham, B.G. (2001). "Investigation of the relationship among three common measures of precedence: fusion, localization dominance, and discrimination suppression," *J. Acoust. Soc. Am.* **109**, 346-358.
- Litovsky, R.Y., Jones, G.L., Agrawal, S., and van Hoesel, R.J.M. (2010). "Effect of age at onset of deafness on binaural sensitivity in electric hearing in humans," *J. Acoust. Soc. Am.* **127**, 400-414.
- Loizou, P.C., Hu, Y., Litovsky, R.Y., Yu, G., Peters, R., Lake, J., and Roland, P. (2009). "Speech recognition by bilateral cochlear implant users in a cocktail-party setting," **125**, 372-383.
- Lutfi, R.A. (1995). "Correlation coefficients and correlation ratios as estimates of observer weights in multiple-observation tasks," *J. Acoust. Soc. Am.* **97**, 1333-1334.
- Macauley, E.J., Hartmann, W.M., and Rakerd, B. (2010). "The acoustical bright spot and mislocalization of tones by human listeners," *J. Acoust. Soc. Am.* **127**, 1440-1449.
- Macpherson, E.A., and Wagner, M.L. (2008). "Temporal weighting of cues for vertical-plane sound localization," *Assoc. Res. Otolaryngol. Abstr.* **31**, 882.
- Maier, J.K., McAlpine, D., Klump, G.M., and Pressnitzer, D. (2010). Context effects in the discriminability of spatial cues. *JARO* **11**, 319-328.
- Majdak, P. and Laback B. (2009). "Effects of center frequency and rate on the sensitivity to interaural delay in high-frequency click trains," *J. Acoust. Soc. Am.* **125**, 3903-3913.
- McAlpine, D., Jiang, D., and Palmer, A.R. (2001). "A neural code for low-frequency sound localization in mammals," *Nat. Neurosci.* **4**, 396-401.
- McCall, D.D., Freyman, R.L., and Clifton, R.K. (1998). "Sudden changes in spectrum of an echo cause a breakdown of the precedence effect," *Percept. Psychophys.* **60**, 593-601.
- McFadden, D., Jeffress, L.A., and Russel, W.E. (1973). "Individual differences in sensitivity to interaural differences in time and level," *Percept. Mot. Skills.* **37**, 755-761.
- Mills, A.W. (1958). "On the minimum audible angle," *J. Acoust. Soc. Am.* **30**, 237-246.
- Mills, A.W. (1960). "Lateralization of high-frequency tones," *J Acoust Soc Am* **32**, 132-134.

- Moore, D.S. and McCabe, G.P. (2005). *Introduction to the practice of statistics, 5th Edition*. (W.H. Freeman, New York, NY).
- Neuman, A.C., Wroblewski, M., Hajicek, J., and Rubinstein, A. (2012). "Measuring speech recognition in children with cochlear implants in a virtual classroom," JSHLR doi: **10.1044/1092-4388(2011/11-0058)**.
- Patynen, J., Katz, B.F.G., Lokki, T. (2011). "Investigations on the balloon as an impulse source," J. Acoust. Soc. Am. **129**, EL27-EL33.
- Pecka, M., Zahn, T.P., Saunier-Rebori, B., Siveke, I., Wiegrebe, L, Klug, A., Pollak, G.D., and Grothe, B. (2007). "Inhibiting the inhibition: A neuronal network for sound localization in reverberant environments." J. Neurosci. **267**, 1782-1790.
- Phillips, D.P. and Hall, S.E., 2005. Psychophysical evidence for adaptation of central auditory processors for interaural differences in time and level. *Hear. Res.* **202**, 188-199.
- Plenge, G. (1974). "On the differences between localization and lateralization," J. Acoust. Soc. Am. **56**, 944-951.
- Rakerd, B. and Hartmann, W.M. (1985). "Localization of sound in rooms, II: The effects of a single reflecting surface." J. Acoust. Soc. Am. **78**, 524-533.
- Rayleigh, Lord (1907). "On our perception of sound direction," *Phil. Mag.* **13**, 214-232.
- Richards, V.M. and Zhu, S. (1994). "Relative estimates of combination weights, decision criteria, and internal noise based on correlation coefficients," J. Acoust. Soc. Am. **95**, 423-434.
- Saberi, K (1996). "Observer weighting of interaural delays in filtered impulses," *Percept. Psychophys.* **58**, 1037-1046.
- Saberi, K. and Perrot, D.R. (1995). "Lateralization of click-trains with opposing onset and ongoing interaural delays," *Acustica* **81**, 272-275.
- Saberi, K. and Antonio, J.V. (2003). "Precedence-effect thresholds for a population of untrained listeners as a function of stimulus intensity and interclick interval," J. Acoust. Soc. Am. **114**, 420-429.
- Saberi, K., Antonio, J.V., and Petrosyan, A. (2004). "A population study of the precedence effect," *Hear. Res.* **191**, 1-13.
- Sanders, L.D., Zobel, B., Keen, R. and Freyman, R.L. (2011). "Manipulations of listeners' echo perception are reflected in event-related potentials. J. Acoust. Soc. Am. **129**, 301-309.
- Sayers, B., and Cherry, E.C. (1957). "Mechanisms of binaural fusion in the hearing of speech," J. Acoust. Soc. Am. **29**, 973-987.
- Schone, H. (1986). *Spatial Orientation*. In *Princeton Series in Neurobiology and Behavior*, Caprancia, R.R., Marler, P., and Adler, N.T. [series Eds.] (Princeton University Press, Princeton, N.J.).
- Schnupp, J.W.H. and Carr, C.E. (2009). "On hearing with more than one ear: Lessons from evolution," *Nat. Neurosci.* **12**, 692-697.
- Seeber, B.U. and Hafter, E.R. (2008). "Parameters affecting the precedence-effect with cochlear implants," J. Acoust. Soc. Am. **123**, 3055.
- Shannon, R.V. (2012). "Advances in auditory prostheses," *Curr. Opin. Neurol.* **25**, 61-66.

- Shen, J., Feng, A.S., Xu, Z., Yu, Z., Arch, V.S., Yu, X., and Narins, P.M. (2008). "Ultrasonic frogs show hyperacute phonotaxis to female courtship calls," *Nature* **453**, 914-916.
- Shinn-Cunningham, B.G., Zurek, P.M., and Durlach, N.I. (1993). "Adjustment and discrimination measurements of the precedence effect," *J. Acoust. Soc. Am.* **98**, 164-171.
- Shinn-Cunningham, B.G., Santarelli, S., and Kopco, N. (2000). "Tori of confusion: Binaural localization cues for sources within reach of a listener," **107**, 1627-1636.
- Shinn-Cunningham, B.G., Kopco, N., Martin, T.J. (2005). "Localizing nearby sound sources in a classroom: binaural room impulse responses," **117**, 3100-3015.
- Siebert, W.M. (1968). "Stimulus transformations in the peripheral auditory system," in *Recognizing Patterns*, eds. P. Kolars and M. Eden (MIT Press, Cambridge, MA).
- Skottun, B.C., Shackleton, T.M., Arnott, R.H., and Palmer, A.R. (2002). "The ability of inferior colliculus neurons to signal differences in interaural delay," *Proc. Nat. Acad. Sci.* **98**, 14050-14054.
- Stecker, G.C. (2000). Observer weighting in sound localization. Doctoral Dissertation. University of California, Berkeley.
- Stecker, G.C. and Hafter, E.R. (2002). "Temporal weighting in sound localization," *J. Acoust. Soc. Am.* **112**, 1046-1057.
- Stecker, G.C. and Hafter, E.R. (2009). "A recency effect in sound localization?" *J. Acoust. Soc. Am.* **125**, 3914-3924.
- Stecker, G.C. and Brown, A.D. (2009). "Modeling temporal weighting of interaural time and level differences in high-rate click trains," *J. Acoust. Soc. Am.* **125**, 2523.
- Stecker, G. C. and Brown, A. D. (2010). "Temporal weighting of binaural cues revealed by detection of dynamic interaural differences in high-rate Gabor click trains," *J. Acoust. Soc. Am.* **127**, 3092-3103.
- Stern, R.M., and Colburn, H.S. (1978). "Theory of binaural interaction based on auditory-nerve data. IV. A model for subjective lateral position," *J. Acoust. Soc. Am.* **64**, 127-140.
- Stern, R.M., Zeilberg, A.S., and Trahiotis, C. (1988). "Lateralization of complex binaural stimuli: A weighted-image model." *J. Acoust. Soc. Am.* **84**, 952-964.
- Stevens, S.S. and Newman, E.B. (1936). "The localization of actual sources of sound," *Am. J. Psychol.* **48**, 297-306.
- Tollin, D.J. and Henning, G.B. (1998). "Some aspects of the lateralization of echoed sound in man. I. The classical interaural-delay based precedence effect," *J. Acoust. Soc. Am.* **104**, 3030-3038.
- Tollin, D.J. and Henning, G.B. (1999). "Some aspects of the lateralization of echoed sound in man. II. The role of the stimulus spectrum," *J. Acoust. Soc. Am.* **105**, 838-849.
- Tollin, D.J. (2003). "The lateral superior olive: A function role in sound source localization," *Neuroscientist*, **9**, 127-143.

- Tollin, D. J. and Yin, T.C.T. (2003). "Psychophysical investigation of an auditory spatial illusion in cats: the precedence effect," *J. Neurophysiol.* **90**, 2149-2162.
- Tollin, D. J. and Yin, T.C.T. (2003). "Interaural phase and level difference sensitivity in low-frequency neurons in the lateral superior olive. *J. Neurosci.* **25**, 10648-10657.
- Viemeister, N.F. and Wakefield, G.H. (1991). "Temporal integration and multiple looks," *J. Acoust. Soc. Am.* **90**, 858-865.
- Wallach, H., Newman, E.B., and Rosenzweig, R. (1949). "The precedence effect in sound localization," *Am. J. Psych.*, **62**, 315-336.
- Wickesburg, R.E. and Oertel, D. (1990). "Delayed, frequency-specific inhibition in the cochlear nuclei of mice: A mechanism for monaural echo suppression," *J. Neurosci.* **10**, 1762-1768.
- Wightman, F.L. and Kistler, D.J. (1992). "The dominant role of low-frequency interaural time differences in sound localization," *J. Acoust. Soc. Am.* **91**, 1648-1661.
- Xia, J., Brughera, A., Colburn, H.S., Shinn-Cunningham, B.G. (2010). "Physiological and psychophysical modeling of the precedence effect," *JARO* **11**, 495-513.
- Yang, X. and Grantham, D.W. (1997). "Echo suppression and discrimination suppression aspects of the precedence effect," *Percept. Psychophys.* **59**, 1108-1117.
- Yin, T.C.T. and Chan, J.K.C. (1990). "Interaural time sensitivity in medial superior olive of cat," *J. Neurophysiol.* **64**, 465-488.
- Zilany, M.S.A., Bruce, I.C., Nelson, P.C., and Carney, L.H. (2009). "A phenomenological model of the synapse between the inner hair cell and auditory nerve: Long-term adaptation with power-law dynamics," *J. Acoust. Soc. Am.* **126**, 2390-2412.
- Zurek, P.M. (1980). "The precedence effect and its possible role in the avoidance of interaural ambiguities," *J. Acoust. Soc. Am.* **67**, 952-964.
- Zwislocki, J. and Feldman, R.S. (1956). "Just noticeable differences in dichotic phase," *J. Acoust. Soc. Am.* **28**, 860-864.

Vita

Personal information

Born: January 27, 1987, in Kent, Washington

Education

Ph.D.: Speech and Hearing Sciences, 2012. University of Washington.

B.S.: Speech and Hearing Sciences, 2007, University of Washington. President's Medal Nominee. *Summa Cum Laude*. Departmental Honors. **Minor in Music**.

A.A.: Associate of Arts, 2005. With Highest Honors. Green River Community College, Auburn, WA.

Papers

Brown, A.D., Kuznetsova, M.S., Spain, W.J., and Stecker, G.C. (2012). "Frequency-specific, location-nonspecific adaptation of interaural time difference sensitivity," *Hear. Res.* **291**: 52-56.

Stecker, G.C. and **Brown, A.D.** (2012). "Onset- and offset-specific effects in interaural level difference discrimination," *J. Acoust. Soc. Am.* (**in press**).

Brown, A.D. and Stecker, G.C. (2012). "The precedence effect in sound localization: Fusion and lateralization measures for pairs and trains of clicks lateralized by interaural time and level differences," *J. Acoust. Soc. Am.* (**submitted**).

Stecker, G.C., Ostreicher, J.O. and **Brown, A.D.** (2012). "Temporal weighting functions for interaural time and level differences. III. Measurements using open-loop lateralization tasks," *J. Acoust. Soc. Am.* (**submitted**).

Brown, A.D., Mussen, T.D., Sisneros, J.A., and Coffin, A.B. (2011). "Reevaluating the use of aminoglycoside antibiotics in behavioral studies of the lateral line," *Hear. Res.* **272**, 1-4.

Brown, A.D. and Stecker, G.C. (2011). "Temporal weighting of interaural time and level differences in high-rate click trains. II. The effect of binaurally synchronous temporal jitter," *J. Acoust. Soc. Am.* **129**, 293-300.

Brown, A.D. and Stecker, G.C. (2010). "Temporal weighting of interaural time and level differences in high-rate click trains," *J. Acoust. Soc. Am.* **128**, 332-341.

Stecker, G.C. and **Brown, A.D.** (2010). "Temporal weighting of binaural cues revealed by detection of dynamic interaural differences in high-rate Gabor click trains," *J. Acoust. Soc. Am.* **127**, 3092-3103.