The role of envelope statistics in detecting changes in interaural correlation

Matthew J. Goupell^{a,b)}

Waisman Center, University of Wisconsin, 1500 Highland Avenue, Madison, Wisconsin 53705

(Received 2 March 2011; revised 5 July 2012; accepted 10 July 2012)

The role of envelope statistics in binaural masking-level differences (BMLDs) and correlation change detection was investigated in normal-hearing listeners. Thresholds and just-noticeable differences (JNDs) were measured for different bandwidths and center frequencies (500, 2000, 4000, and 8000 Hz) using Gaussian noises (GNs) and low-fluctuation noises (LFNs). At a 500-Hz center frequency, GN NoSo thresholds were higher than, NoS π thresholds were lower than, and correlation change detection JNDs were the same as LFN thresholds and JNDs. At higher center frequencies, GN NoS π thresholds were the same or higher than LFN thresholds and GN correlation change detection JNDs were much smaller than LFN JNDs. Using a pulsed sine vocoder, a second experiment was performed to investigate if binaural adaptation might contribute to the difference in GN and LFN detection. There was no effect of pulse rate, thus providing no clear evidence that binaural adaptation plays a role in these tasks. Both a cross-correlation model and a model that utilized the fluctuations in the interaural differences could explain a majority of the variance in the LFN correlation change JNDs. © 2012 Acoustical Society of America. [http://dx.doi.org/10.1121/1.4740498]

PACS number(s): 43.66.Pn, 43.66.Ba, 43.66.Qp [TD]

Pages: 1561–1572

I. INTRODUCTION

Human listeners are typically very sensitive to differences in signals between the ears or interaural decorrelation (e.g., Gabriel and Colburn, 1981). Decorrelation can be introduced by adding an out-of-phase target (S π) to a diotic noise background (No), as is used in binaural masking-level difference (BMLD) experiments. The metric often used to measure the amount of interaural decorrelation in a stimulus is the normalized cross-correlation (Bernstein and Trahiotis, 1996). The purpose of this paper is to investigate the role of envelope statistics on detection of changes in interaural correlation, namely, the difference between Gaussian noise (GN) and low-fluctuation noise (LFN). LFN has an envelope that is particularly flat (few extreme peaks and dips) compared to GN (Pumplin, 1985). The distribution of the envelope amplitudes or the envelope fluctuation strength can be quantified by using the normalized envelope kurtosis:

$$Y = \frac{\langle E^4 \rangle}{\langle E^2 \rangle^2},$$

where E is the envelope and $\langle x \rangle$ denotes the expectation of x .

Several reports have compared tone detection in GN and LFN (Hartmann and Pumplin, 1988; Kohlrausch *et al.*, 1997; Eddins and Barber, 1998; Hall *et al.*, 1998). For NoSo (diotic noise, diotic target), it is *more difficult* to detect the presence of the tone in GN compared to LFN. One explanation for the difference in GN-LFN NoSo threshold is that adding an arbitrary phase sine tone to a LFN can alter Y and listeners use

the change in envelope fluctuations as a cue (Hartmann and Pumplin, 1988; Kohlrausch *et al.*, 1997). For NoS π , it is sometimes *easier* to detect the presence of the tone in GN compared to LFN. GN NoS π thresholds are about 2 dB lower than LFN NoS π thresholds for stimuli with a 500-Hz center frequency (CF) (Eddins and Barber, 1998; Hall *et al.*, 1998). The difference is about 1 dB for stimuli with a 1000-Hz CF (Kohlrausch *et al.*, 1997) and 0 dB for stimuli with a 4000-Hz CF (Eddins and Barber, 1998).

One would predict that LFN NoS π thresholds would be lower than the GN NoS π thresholds assuming detection at threshold is performed by detecting an equal change in the normalized cross-correlation for both masker types, which is opposite to the observed trend (Eddins and Barber, 1998). Bernstein *et al.* (1999) proposed calculating the normalized cross-correlation after a stage of compression so that the model would predict the correct order of GN and LFN NoS π thresholds. To the present day, the cross-correlation including compression and other physiologically relevant features can account for a myriad of binaural phenomena (e.g., Bernstein and Trahiotis, 2009, 2010).

Acoustic stimuli are compressed at many stages of the auditory system (van der Heijden and Joris, 2009); however, other factors besides compression might contribute to the GN-LFN difference in NoS π thresholds. The GN-LFN difference could be partially due to “binaural adaptation” (Haftner and Buell, 1990). Binaural adaptation is the phenomenon whereby sensitivity to interaural time differences (ITDs) and interaural level differences (ILDs) in constant-amplitude pulse trains decreases with increasing pulse rate (Haftner and Dye, 1983; Haftner *et al.*, 1983; Stecker and Haftner, 2002; Brown and Stecker, 2010; Stecker and Brown, 2010). The binaural system seems to recover from binaural adaptation if a novel element in the signal is introduced, such as a change in the interpulse interval or another sound like a sine tone

^{a)}Present address: Department of Hearing and Speech Sciences, University of Maryland, College Park, MD 20742.

^{b)}Author to whom correspondence should be addressed. Electronic mail: goupell@umd.edu.

(Haftner and Buell, 1990; Laback and Majdak, 2008; Goupell *et al.*, 2009b; Brown and Stecker, 2011). Such a change partially restores sensitivity to ITDs and ILDs at higher pulse rates. Recovery from binaural adaptation may be relevant to GN and LFN NoS π thresholds. GNs have intrinsic novel stimulus changes, namely, extreme peaks and dips in the envelope. LFNs have no appreciable peaks and dips, essentially being constant-amplitude stimuli, thus having minimal intrinsic novel stimulus changes. If binaural adaptation also occurs in the neural processing of interaurally decorrelated stimuli like NoS π , which have time-varying ITDs and ILDs, GN should have lower thresholds because the extreme peaks and dips in the envelope would provide a recovery from binaural adaptation. LFN should have higher thresholds because there are minimal peaks and dips in the envelope, thus providing little recovery from binaural adaptation.

LFNs are also interesting from the standpoint of hearing-impaired (HI) people and cochlear-implant (CI) users. Hearing-aid and CI processing schemes typically include compression and automatic gain control to limit the presentation of overly loud sounds. Thus, hearing-aid and CI users most likely often experience sounds with unusually flat envelopes, like a LFN. Although there are many reasons why binaural sensitivity is reduced in HI and CI listeners, a reduced dynamic range from compression and automatic gain control may limit utilization of binaural cues. Also, since CIs present only envelope information with high-rate electrical pulse trains and sensitivity to ITDs in constant-amplitude pulse trains are known to produce binaural adaptation (Laback and Majdak, 2008; van Hoesel *et al.*, 2009), it is important to investigate whether this might occur for correlation change detection using acoustic pulse trains as a CI simulation.

The purpose of this study was to investigate GN and LFN binaural detection to determine if the differences in them can be explained by signal compression, binaural adaptation, or some other facet of the stimuli. To create a data set that might best discriminate between binaural models (Domnitz and Colburn, 1976), GN and LFN NoSo and NoS π thresholds were measured at several CFs and BWs. In addition, GN and LFN correlation change detection just-noticeable differences (JNDs) were also measured, which have the advantage of the minimizing the monaural cues that may affect NoS π thresholds. In a second experiment, to determine if binaural adaptation possibly affects thresholds and JNDs, the stimuli envelopes were sampled using a pulsed sine vocoder at different pulse rates. A decrease in binaural sensitivity with increasing pulse rate would show binaural adaptation contributes to the GN-LFN detection difference. Finally, numerical models attempted to account for the variance in the detection data using several binaural stimulus metrics including the cross-correlation with compression and metrics that focused on time-varying interaural phase differences (IPDs) and ILDs. Investigating the ability of binaural metrics other than the normalized cross-correlation to predict LFN NoS π thresholds may elucidate the mechanism for the GN-LFN difference, including the form and parameters of binaural models to describe the detection data.

II. EXPERIMENT 1: BINAURAL DETECTION THRESHOLDS AND JNDS AS A FUNCTION OF STIMULUS BANDWIDTH AND CENTER FREQUENCY

A. Listeners and equipment

Nine listeners (20 to 30 years) participated in this experiment. The listeners had pure tone thresholds within 20 dB of audiometrically normal hearing at octave interval frequencies between 250 and 8000 Hz, and the thresholds between the two ears differed by less than 10 dB at any tested frequency. Five listeners participated in the BMLD experiments (SMS, SNA, SNC, SND, and SYA), while all nine listeners participated in the correlation change detection experiments. Listener SYA was the author.

The stimuli were delivered by a personal computer to a Tucker-Davis System 3 real-time processor (RP2.1) and headphone driver (HB7), and finally a pair of headphones (Sennheiser HD580). Listeners performed the experiments in a double-walled sound attenuating booth (IAC).

B. Stimuli

Stimuli were dual-channel narrowband Gaussian or low-fluctuation noise-pairs, possibly with a sine tone embedded in the noise. A LFN was generated by extracting the Hilbert envelope of a GN, dividing the time waveform by the envelope, and filtering the stimulus to the desired BW in the frequency domain (Kohlrausch *et al.*, 1997). This process was repeated 10 times. Fifty LFNs were generated for each BW and CF condition tested. For the correlation change detection stimuli, dual-channel 10-Hz BW LFNs with various amounts of interaural decorrelation were made by randomly choosing two LFNs from the pool of 50 previously generated tokens ($Y = 1.02 - 1.06$) and performing Gram-Schmidt Orthogonalization (Culling *et al.*, 2001). After orthogonalization, Y was checked for each channel. A dual-channel LFN stimulus was accepted if both channels had Y less than 1.1, otherwise it was discarded.¹

Stimuli were 500 ms in duration and were temporally shaped by a Tukey window with a rise-fall time of 10 ms. The CF of the stimuli was either 500, 2000, 4000, or 8000 Hz. The BW of the NoSo and NoS π stimuli was 10 Hz, 50 Hz, or 1 equivalent rectangular BW (ERB; 78, 240, 456, or 888 Hz) according to Moore and Glasberg (1983). The correlation change detection stimuli had a 10-Hz BW because computation time limited the ability to make a large number of highly decorrelated LFNs with BWs greater than 10 Hz. The noises were arithmetically centered on the CF, had an A-weighted sound pressure level of 65 dB, were sampled at 48.848 kHz, and had 24-bit resolution.

The target stimuli for the BMLD tasks had a tone-in-noise configuration (NoSo and NoS π) with the tone at the CF of the noise with a varied signal-to-noise ratio (SNR) of the tone to the noise. The SNR is reported with respect to the overall level of the noise, not the spectrum level. The non-target stimuli for the BMLD tasks were diotic noises (No). The target stimuli for the correlation change detection tasks had a varied value of the cross-correlation (ρ), while the non-target stimuli were fixed at one of two reference correlations

($\rho_{ref} = 1$ or 0). In total, 64 conditions were tested in this experiment (BMLD: 4 CF $\times$ 3 BW $\times$ 2 noise types $\times$ 2 configurations; correlation change detection: 4 CF $\times$ 1 BW $\times$ 2 noise types $\times$ 2 references).

The stimuli were generated offline before the experiments because of the computation time to generate acceptable LFNs. There were 25 different noise tokens for each condition and each value of SNR or ρ used in the adaptive procedure. Stimuli were randomly drawn from the available tokens without replacement for a given trial.

C. Procedure

A 4-interval, 2-alternative forced-choice procedure was used. The listener initiated each trial, which consisted of four intervals grouped into two sequential pairs. The inter-interval duration was 250 ms between stimuli in the first and second pairs, and was 500 ms between the two pairs. Three intervals contained non-target stimuli. The target stimulus was in either the second or third interval, which was random from trial to trial. No feedback was provided.

A 2-down-1-up adaptive staircase procedure was used to measure thresholds or JNDs. For the BMLD configurations, the sine tone began at a +10-dB SNR. The step size was 8 dB until the first reversal, 4 dB until the second, and 2 dB for the rest of the staircase. For correlation change detection, the starting value was perfectly uncorrelated ($\rho = 0$) for a correlated reference ($\rho_{ref} = 1$), or perfectly correlated ($\rho = 1$) for an uncorrelated reference ($\rho_{ref} = 0$). The maximum allowable difference in correlation allowed in the task was $\Delta\rho = 1$. The step size of the staircase procedure was $\alpha = \sqrt{1 - \rho^2} = 0.4$ until the first reversal, 0.2 until the second, 0.1 until the third, and 0.05 for the rest of the staircase. Correlation change detection psychometric functions have been previously measured on an α scale (Goupell, 2010). For the correlation change detection tasks, if listeners could

not detect the change in correlation (i.e., had four incorrect answers at $\Delta\alpha = \Delta\rho = 1$), the run was stopped and the JND was recorded as “not determinable.” These JNDs translated to a numerical value of $\Delta\alpha = 1.1$ for that run when the data were analyzed. Ten reversals were measured for each staircase. The last six reversals were averaged to calculate the threshold or JND for a run. Three runs were performed per condition and the listener’s reported threshold or JND was the average threshold or JND over the three runs. In conditions that had a large standard deviation over the three runs (3 dB for the BMLD tasks and 33% of the JND for the correlation change detection tasks), two extra runs were performed and all five were used to calculate the average threshold or JND.

Training that included correct answer feedback was performed using 500-Hz CF, 10-Hz BW stimuli before data collection. Initially, training for correlation change detection of $\rho_{ref} = 1$, GN stimuli was provided and was approximately 2–4 h, which was until performance appeared to asymptote. Then there was training for six runs of $\rho_{ref} = 0$ GNs, three runs each of NoSo GNs and LFNs, and three runs each of NoS π GNs and LFNs.

During data collection, conditions were randomized across listeners for each task type, constrained that at least three runs were completed before moving on to a new condition. The task type order was $\rho_{ref} = 1$, $\rho_{ref} = 0$, NoSo, and NoS π . Testing was performed in this manner to minimize listener confusion from changing the pertinent cues of the task too often.

D. Results

The BMLD data averaged over listeners are shown in Fig. 1. For the NoSo data (Fig. 1, top row), LFN thresholds are lower than the GN thresholds, except LFN thresholds are nearly the same as the GN thresholds for the 1-ERB stimuli

FIG. 1. Average NoSo (top row) and NoS π (bottom row) thresholds for GN (open symbols) and LFN (closed symbols) as a function of bandwidth (BW). Panels are arranged in increasing center frequency (CF) from left to right. The data points represent the arithmetic mean; the error bars represent ± 1 standard deviation.

TABLE I. Differences between LFN and GN thresholds for the NoSo and NoS π conditions for individual listeners and average.

CH (Hz)		500			2000			4000			8000			Average	
BW (Hz)		10	50	78	10	50	240	10	50	456	10	50	888	10&50	All
NoSo	SMS	-11.4	-9.4	-11.5	-5.5	-5.7	-4.3	-11.6	-13.8	-2.9	-12.5	-9.3	-0.2	-9.9	-8.2
	SNA	-10.9	-4.7	-7.0	-1.2	-5.6	-2.2	-4.9	-8.2	-0.8	-3.7	-11.0	-0.6	-6.3	-5.1
	SNC	-9.9	-10.9	-4.7	-9.3	-10.7	-2.4	-8.3	-13.4	-2.1	-12.2	-10.9	-0.9	-10.7	-8.0
	SND	-13.8	-7.7	-6.1	-13.0	-12.7	-16.0	-17.1	-8.8	-5.2	-12.1	-13.6	-6.3	-12.4	-11.0
	SYA	-10.7	-5.1	0.4	-6.8	-7.6	-3.3	-9.7	-14.6	-4.2	-9.6	-17.8	-4.0	-10.2	-7.7
	Average	-11.3	-7.5	-5.8	-7.2	-8.4	-5.7	-10.3	-11.8	-3.0	-10.0	-12.5	-2.4	-9.9	-8.0
NoS π	SMS	2.1	3.9	1.4	3.7	3.8	-0.7	-0.1	-1.1	1.4	-0.1	-7.0	-2.2	0.7	0.4
	SNA	6.3	1.2	1.1	-9.6	-3.9	-2.8	-4.8	-6.3	-4.6	-7.2	-6.9	0.3	-3.9	-3.1
	SNC	5.3	-0.6	-3.1	1.5	6.8	0.1	5.7	0.7	-3.2	-7.9	-8.7	5.7	0.4	0.2
	SND	1.9	0.6	4.2	0.0	-1.2	0.1	-5.0	1.9	-1.8	-2.8	-4.2	-4.7	-1.1	-0.9
	SYA	4.7	6.1	5.4	4.8	0.8	5.8	2.3	0.9	0.4	-1.6	-2.1	0.7	2.0	2.3
	Average	4.1	2.2	1.8	0.1	1.2	0.5	-0.4	-0.8	-1.5	-3.9	-5.8	0.0	-0.4	-0.2

at the 4- and 8-kHz CFs. Table I shows the difference between the LFN and GN thresholds are indeed smaller for the 1-ERB stimuli.

For the NoS π data, Fig. 1 and Table I show the LFN thresholds were higher than the GN thresholds at the 500-Hz CF (difference of 4.1 dB, 2.2 dB, and 1.8 dB for BWs of 10 Hz, 50 Hz, and 78 Hz, respectively). The LFN thresholds were essentially the same as the GN thresholds at the 2000- and 4000-Hz CFs. The LFN thresholds were lower than the GN thresholds for two of the BWs at the 8000-Hz CF (difference of 3.9 dB, 5.8 dB, and 0.0 dB for BWs of 10 Hz, 50 Hz, and 888 Hz, respectively).

Since listeners probably use different detection cues to perform NoSo and NoS π tasks (monaural vs binaural, respectively), separate statistical analyses were used. A three-way repeated-measure analysis of variance (RM ANOVA) was performed on the NoSo data with factors BW, CF, and noise type. All three main effects were significant [BW: $F(2,48) = 9.62$, $p < 0.0001$; CF: $F(3,48) = 29.1$, $p < 0.0001$; noise: $F(1,48) = 737.4$, $p = 0.0003$]. The interactions

BW $\times$ CF, BW $\times$ noise, and BW $\times$ CF $\times$ noise were significant [$F(6,48) = 3.65$, $p = 0.005$; $F(2,48) = 41.3$, $p < 0.0001$; $F(6,48) = 5.33$, $p = 0.0003$, respectively]. The interaction CF $\times$ noise was not significant [$F(2,48) = 1.07$, $p = 0.37$].

A three-way RM ANOVA was performed on the NoS π data with factors BW, CF, and noise type. Two main effects were significant [BW: $F(2,48) = 13.7$, $p < 0.0001$; CF: $F(3,48) = 87.2$, $p < 0.0001$] and one was not significant [noise: $F(1,48) = 0.25$, $p = 0.62$]. The interactions BW $\times$ CF and CF $\times$ noise were significant [$F(6,48) = 5.37$, $p = 0.0003$; $F(2,48) = 8.80$, $p < 0.0001$, respectively]. The interactions BW $\times$ noise and BW $\times$ CF $\times$ noise were not significant [$F(2,48) = 0.48$, $p = 0.63$; $F(6,48) = 1.50$, $p = 0.20$, respectively].

The correlation change detection data are shown in Fig. 2. Because of the large inter-individual variability observed, the data were also broken up into a group of “good” listeners (SMS, SNC, SND, and SYA) and “poor” listeners (SNA, SNB, SNG, SYI, and SYJ), who were separated by their correlation change detection JNDs at higher frequencies, particularly at the 4000- and 8000-Hz CFs.

FIG. 2. Correlation change detection JNDs for GN (open symbols) and LFN (closed symbols) for $\rho_{ref} = 1$ (top row) and $\rho_{ref} = 0$ (bottom row). The individual listener data are shown in the left two panels. The overall averages and averages over “good” and “poor” listeners are shown in the right three columns. The data points in the right three columns represent the arithmetic mean; the error bars represent ± 1 standard deviation. The symbols connected by dashed lines represent the best model predictions for the data.

Figure 2 shows JNDs increased with increasing CF. LFN correlation change detection yielded higher JNDs than GN at all CFs for both reference correlations except at the 500-Hz CF. At the 500-Hz CF, the GN and LFN JNDs were essentially the same for both $\rho_{ref} = 1$ and 0. At CFs greater than 500 Hz for $\rho_{ref} = 1$, the good listeners had determinable JNDs, whereas the poor listeners had JNDs that were mostly near $\Delta\alpha = 1$ (which was the maximum allowable difference in the task) or were not determinable. For $\rho_{ref} = 0$, all of the listeners had JNDs were near $\Delta\alpha = 1$ or were not determinable.

A three-way RM ANOVA was performed on the correlation change detection data with factors CF, noise type, and reference ($\rho_{ref} = 1$ or 0). All three main effects were significant [CF: $F(3,32) = 139.3$, $p < 0.0001$; noise: $F(1,32) = 78.2$, $p < 0.0001$; reference: $F(1,32) = 69.5$, $p < 0.0001$]. The interactions CF $\times$ noise, CF $\times$ reference, and CF $\times$ noise $\times$ reference were significant [$F(3,32) = 11.0$, $p < 0.0001$; $F(3,32) = 3.39$, $p = 0.030$; $F(3,32) = 3.26$, $p = 0.034$, respectively]. Note that it is possible that ceiling effects contributed to the significant interactions. The interaction noise $\times$ reference was not significant [$F(1,32) = 0.035$, $p = 0.85$]. A separate two-way RM ANOVA was performed on just the 500-Hz CF data to confirm that the 500-Hz CF conditions were in fact not statistically different. There was a significant effect of reference after correction for multiple comparisons [$F(1,8) = 50.2$, $p < 0.0001$], but there was no significant effect of noise type [$F(1,8) = 0.15$, $p = 0.71$] and the interaction noise $\times$ reference was not significant [$F(1,8) = 3.50$, $p = 0.098$]. Thus, the GN and LFN data at the 500-Hz CF were not significantly different.

E. Discussion

Previous studies measured LFN NoSo thresholds that were approximately 5–10 dB lower than GN NoSo thresholds (Hartmann and Pumplin, 1988; Kohlrausch *et al.*, 1997; Eddins and Barber, 1998; Hall *et al.*, 1998). In this study, LFN NoSo thresholds were 9.9 dB lower than the GN NoSo thresholds for BWs smaller than a critical band when averaged over CF (Fig. 1 and Table I). The GN-LFN difference in NoSo thresholds was smaller for the 1-ERB stimuli compared to the others, particularly at the 4- and 8-kHz CF. The higher thresholds for the LFN 1-ERB stimuli is consistent with the notion that when the noise has a BW that comparable or larger than the BW of the auditory filter, this would cause the amplitude and phase spectrum to be modified and increase Y of the internal representation of the sound on-frequency (Hartmann and Pumplin, 1988) and off-frequency (Kohlrausch *et al.*, 1997), thus removing the potential envelope fluctuation cue when the tone was added to the noise.

For the NoS π data in Fig. 1 and Table I, the LFN thresholds were 1.8–4.4 dB higher than the GN thresholds at a 500-Hz CF. The differences between the LFN and GN thresholds were negligible at the 2- and 4-kHz CFs for all of the measured BWs. Therefore, the measurements are roughly consistent with the previous measurements; any differences between the studies may be due to different methods

(Kohlrausch *et al.*, 1997; Eddins and Barber, 1998; Hall *et al.*, 1998). At the 8-kHz CF, the LFN thresholds were 3.9 and 5.8 dB lower for BWs of 10 and 50 Hz, respectively. Although the LFN NoS π thresholds are lower than the NoSo thresholds, one explanation for the LFN NoS π thresholds being lower than the GN NoS π thresholds at the subcritical bandwidths is that the monaural envelope fluctuation cues may have contributed to detection at low SNRs (Hartmann and Pumplin, 1988; Kohlrausch *et al.*, 1997), particularly for listeners with relatively high NoS π thresholds.

A control experiment was run to determine if listeners used monaural cues in the NoS π detection task. Detection thresholds were measured for NoS π targets vs No non-targets and NoS π targets vs NoSo non-targets. Using NoSo non-targets should remove the potential monaural cues and force listeners to rely on the binaural cues. The 4- and 8-kHz CFs were tested using 10-Hz BW stimuli. Four listeners were tested in this experiment, three of which were new but performed the same training as the listeners in the main experiment. The data in Table II show that all of the No non-target thresholds are within a few dB of the NoSo non-target thresholds for the individual listeners and the average. Thus, it appears that listeners were using solely binaural cues in the control experiment. However, the control experiment data did not reproduce a key difference in the data in main experiment (the GN NoS π threshold is 3.9 dB higher than LFN NoS π threshold, see Fig. 1 and Table I). The control and main experiments may have produced different data patterns because of the use of different listeners (with the exception of SYA). The listeners in the main experiment could have been more sensitive to monaural cues as they performed the entire battery of NoSo conditions before the NoS π conditions. Therefore, the control experiment provided inconclusive evidence whether the listeners in the main experiment were partially utilizing monaural cues.

For the correlation change detection data in Fig. 2, JNDs increased with increasing CF, and the $\rho_{ref} = 1$ JNDs were lower than the $\rho_{ref} = 0$. It was found that detection of changes in the correlation was *more difficult* for the LFN than the GN stimuli at higher CFs, but was *the same* at the 500-Hz CF. Listeners may rely upon the change in ρ for both correlation change detection and NoS π detection for GN stimuli (Koehnke *et al.*, 1986; Jain *et al.*, 1991; Bernstein and Trahiotis, 1997). Therefore, it may be surprising that the detection of NoS π was *the same* for the LFN and GN stimuli at higher CFs, but was *more difficult* for the LFN

TABLE II. NoS π thresholds for No and NoSo non-targets.

CF	Noise	Non-target	SYA	SZW	SZX	SZY	Average
4	GN	No	-22.9	-12.7	-9.7	-18.0	-15.8
4	LFN	No	-21.4	-14.3	-14.2	-17.6	-16.9
4	GN	NoSo	-21.6	-11.0	-10.9	-17.6	-15.3
4	LFN	NoSo	-21.1	-12.7	-12.2	-12.9	-14.7
8	GN	No	-17.1	-9.4	-14.2	-13.9	-13.7
8	LFN	No	-16.3	-8.1	-15.5	-15.2	-13.8
8	GN	NoSo	-16.3	-5.3	-14.2	-14.5	-12.6
8	LFN	NoSo	-17.2	-7.2	-14.1	-11.9	-12.6

compared to the GN at the 500-Hz CF, the opposite trend that was found for correlation change detection. Such a contradiction can be understood by assuming that binaural cues for LFN are negligible compared to the binaural cues for GN at CFs higher than 500 Hz. If this is the case, then listeners would have difficulty detecting the target signal in the correlation change detection tasks because there are only binaural cues, translating to a large GN-LFN JND difference for CFs higher than 500 Hz. However, the listeners in the NoS π detection task could utilize monaural envelope fluctuation cues in the near absence of the binaural cues, translating to no GN-LFN threshold difference at the 2- and 4-kHz CFs.

In summary, this experiment replicated previous BMLD findings in the literature and expanded upon the CFs and BWs previously tested. Correlation change detection conditions that focused on reducing any potential monaural cues were also tested. It was found that listeners' ability to use binaural cues decreased with increasing CF and this reduction was more severe for LFN compared to GN. One possible explanation for the correlation change detection results is that binaural adaptation occurs at high frequencies, which has been observed in experiments using pulse trains with high-frequency carrier signals (e.g., [Haft et al., 1983](#); [Haft et al., 1983](#)). Therefore, experiment 2 was performed to test whether binaural adaptation contributes to the GN-LFN detection difference.

III. EXPERIMENT 2: EFFECTS OF REPETITION RATE ON BINAURAL PERFORMANCE USING PULSE-TRAIN VOCODED STIMULI

Binaural adaptation, the reduction in sensitivity to ITDs and ILDs in constant-amplitude pulse trains when the inter-pulse interval is shortened, occurs similarly for both normal-hearing (NH) and CI listeners ([Goupell et al., 2009a](#)). [Haft et al. \(1983\)](#) showed that at 100 pulses per second (pps), NH listeners optimally used the ITD information in all 16 clicks presented in a constant-amplitude click train. As the pulse rate increased to 200 pps, 500 pps, and 1000 pps, listeners relied more on the ITD in the initial click and could not completely utilize the ITD information in the subsequent clicks. Likewise, CI listeners showed similar rate limitations in binaural processing such that ITD sensitivity decreases for pulse rates above 100 pps ([van Hoesel et al., 2009](#)).

It is possible that binaural adaptation occurs in detection of interaural decorrelation in LFNs because they have envelopes that are very similar to constant-amplitude stimuli. GN has more extensive envelope modulations and therefore will suffer less from binaural adaptation because of restarting ([Haft et al., 1990](#)). The envelopes of the stimuli from experiment 1 were sampled with band-limited pulses that encoded the instantaneous amplitude at regular time intervals. It was hypothesized that at a pulse rate of 100 pps, LFN correlation change detection JNDs would be equivalent to that measured for GN because there is minimal binaural adaptation at low rates for ITD and ILD sensitivity. As the pulse rate increases, the LFN JNDs would increase while GN JNDs would remain constant. NoSo and NoS π conditions were also tested.

A. Method

GNs and LFNs from experiment 1 with a 500-Hz CF and 10-Hz BW were used to generate the vocoded stimuli in this experiment. Vocoded stimuli were created by extracting the Hilbert envelope from both channels of the unprocessed stimuli. These envelopes were sampled with periodic trains of Gaussian envelope 4-kHz tone pulses ([Goupell et al., 2010](#)). The rate of the pulse trains was 100, 200, 400, or 800 pps. Note that all of these rates provided adequate sampling of the temporal envelope and no aliasing occurred. The final -3-dB BW of the stimuli after vocoding was 2 kHz and the CF was 4 kHz. The waveform and spectrum of a 100-pps vocoded GN is shown in Fig. 3. The modulation depth of the waveforms was greater than 99% up to the 800-pps pulse rate. The spectrum shows that the frequency content is -50 dB by 1500 Hz. It is possible that energy that is -50 dB could contribute to the ability to utilize ITD cues if they occur at low frequencies ([Bernstein and Trahiotis, 1982](#)). However for the stimuli in this experiment, the binaural differences are only in the peak amplitude of the pulses. The fine-structure is interaurally synchronized. Since the signals have interaural decorrelation in the envelope, there would be no consistent envelope ITD cue present. Therefore, listeners can only utilize the changing ILDs, which is not expected to be better at low compared to high frequencies ([Yost and Dye, 1988](#)). For these reasons, it was deemed not necessary to include low-frequency masking noise in this experiment.

Four conditions were tested: NoSo, NoS π , $\rho_{ref} = 1$, $\rho_{ref} = 0$. The first two conditions tested 100- and 800-pps pulse rates and five listeners from experiment 1. The latter two conditions were tested at all four pulse rates with the nine listeners from experiment 1. The methods were the same as in experiment 1, except that there was no training.

B. Results

The NoSo and NoS π data for the vocoded stimuli are shown in Fig. 4. As in experiment 1, NoSo detection was better for LFN compared to GN. Again, NoS π detection was similar for GN and LFN stimuli. NoS π thresholds were lower than NoSo thresholds. There appears to be no effect of pulse rate for both NoSo and NoS π conditions. A two-way RM ANOVA was performed on the NoSo data with factors noise type and pulse rate. The effect of noise type was significant [$F(1,8) = 55.2$, $p < 0.0001$]. The effect of pulse rate [$F(1,8) = 0.68$, $p = 0.43$] and the interaction noise $\times$ rate

FIG. 3. The waveform (left panel) and spectrum (right panel) of a 100-pps GN used in experiment 2. The spectrum was normalized to have a peak amplitude at 0 dB.

FIG. 4. Average NoSo (left panel) and NoS π (right panel) thresholds for GN (open symbols) and LFN (closed symbols) as a function of pulse rate.

[$F(1,8)=0.40$, $p=0.55$] were not significant. A two-way RM ANOVA was performed on the NoS π data with factors noise type and pulse rate. The effect of noise type [$F(1,8)=1.51$, $p=0.26$], the effect of pulse rate [$F(1,8)=4.23$, $p=0.074$], and the interaction noise $\times$ rate [$F(1,8)=0.33$, $p=0.58$] were not significant.

The correlation change detection data for the vocoded stimuli are shown in Fig. 5. The LFN JNDs were much higher than the GN JNDs. The $\rho_{ref}=1$ JNDs were lower than the $\rho_{ref}=0$ JNDs. There was no apparent effect of pulse rate for both references. A three-way RM ANOVA was performed on the correlation change detection data for the vocoded stimuli with factors noise, rate, and reference. The effects of noise [$F(1,32)=264.3$, $p<0.0001$] and reference [$F(1,32)=117.1$, $p<0.0001$] were significant. The effect of rate was not significant [$F(3,32)=2.09$, $p=0.12$]. All of the interactions were not significant ($p>0.05$ for all). Two listeners (SMS and SYA) had some noticeably lower LFN JNDs than the other listeners in Fig. 4. These two

FIG. 5. Individual (top row) and average (bottom row) correlation change detection JNDs for GN (open symbols) and LFN (closed symbols) as a function of pulse rate. The left column shows $\rho_{ref}=1$ data; the right column shows $\rho_{ref}=0$ data. The individual listener symbols are the same as in Fig. 3.

listeners had the lowest JNDs from experiment 1 and also had extensive previous experience in correlation change detection. For these two listeners, there appeared to be an effect of pulse rate for the LFN stimuli. Therefore, these two listeners performed additional runs (10 in total) for the LFN conditions of this experiment in an attempt to decrease the variability in the data and show a significant effect of pulse rate. No significant differences were found to suggest that higher pulse rates yielded larger JNDs.

C. Discussion

It was hypothesized that binaural adaptation, the decrease in binaural sensitivity with increasing pulse rate, may affect NoS π and correlation change detection similarly to how it affects static interaural difference detection, especially for the LFN conditions. The effect of pulse rate was not significant for either the GN or LFN conditions shown in Figs. 4 and 5. Therefore, it seems that binaural adaptation does not affect dynamic binaural detection like it affects static binaural detection. Perhaps even the slightest variations in the envelope amplitude over time can cause recovery from binaural adaptation.

Although there was no effect of pulse rate in this experiment, the GN NoSo thresholds were worse than LFN thresholds, GN NoS π thresholds were the same as LFN thresholds, and GN correlation change detection JNDs were much better than LFN JNDs. The NoSo results of this experiment are consistent with the results of experiment 1. The NoS π and correlation change results for this experiment are consistent with the results of experiment 1 for CFs greater than 500 Hz. The vocoding process provided only ILDs for the NoS π and correlation change stimuli. Therefore, listeners could use only dynamic ILDs to perform the task in experiment 2. Even if fine-structure IPDs were presented in the vocoded stimuli, they would probably not be perceived because IPDs are not salient for CFs greater than 2 kHz (Yost, 1974). Comparing thresholds and JNDs from this experiment with those in experiment 1, the thresholds and JNDs for the vocoded stimuli are lower than the unprocessed stimuli in experiment 1 with a 4-kHz CF. This can be attributed to the fact that the vocoded stimuli had a larger area of tonotopic excitation than the unprocessed stimuli, which improved the binaural detection thresholds and JNDs.

IV. NUMERICAL MODELING

Since binaural adaptation appears not to contribute to the difference in GN-LFN JNDs, modeling was done to try to describe the variance in the data. Correlation change detection JNDs were predicted from the GN JNDs. The predictions were made using the same method as Bernstein *et al.* (1999), who assumed that the detectable change in an interaural metric was the same for both the GN and LFN stimuli. The NoS π thresholds were not predicted because they are already well predicted by the cross-correlation model at the 500-Hz CF, the thresholds did not vary as a function of BW at the 500-Hz CF, and listeners had the potential to utilize monaural cues.

A. Interaural metrics

1. Normalized cross-correlation

The normalized cross-correlation is calculated by

$$\rho = \frac{\int x_R(t) \cdot x_L(t) dt}{\sqrt{\int x_R^2(t) dt} \sqrt{\int x_L^2(t) dt}},$$

where x_L is the left channel and x_R is the right channel. This interaural metric is often used to describe binaural processing of stimuli that are interaurally decorrelated or have an ITD. However, as noted by [Bernstein et al. \(1999\)](#), the GN-LFN difference in NoS π thresholds at 500 Hz cannot be accounted for with such an interaural metric because ρ predicts the incorrect ordering of thresholds. Therefore, envelope compression was applied to the waveforms before the calculation of ρ . This new metric could predict the GN-LFN difference in thresholds.

The calculation of the interaural cross-correlation used in several studies includes preprocessing stages of filtering by a fourth-order Gammatone filter centered at the CF of the stimulus, envelope compression with a compressive factor of 0.46, half-wave rectification, fourth-order low-pass filtering with a 425-Hz cutoff, and an additional second-order low-pass filtering with a 150-Hz cutoff before the calculation of the normalized cross-correlation. The interaural cross-correlation with peripheral processing stages is denoted ρ_{PP} . Several forms of this model were used in modeling the data, which used or did not use Gammatone filtering and/or envelope extraction (rectification and two stages of low-pass filtering). The envelope compression was also varied between 0 and 1 in 0.05 steps.

2. Fluctuations in interaural differences

The fluctuations in the IPD were calculated by finding the standard deviation over time:

$$s_t(\Delta\Phi) = \frac{1}{T} \sqrt{\int_0^T [\Delta\Phi(t) - \overline{\Delta\Phi}]^2 dt},$$

where $\Delta\Phi(t)$ is the IPD as a function of time calculated from the difference of the instantaneous phases of the analytic signal, $\overline{\Delta\Phi}$ is the average value over time of $\Delta\Phi(t)$, and T is the duration of the stimulus. The fluctuations in the ILD, $s_t(\Delta L)$, were calculated similarly by using the ILDs in the instantaneous envelope of the analytic signal. These two statistics have been used in several studies to account for binaural detection sensitivity (e.g., [Goupell and Hartmann, 2007](#); [Davidson et al., 2009](#); [van der Heijden and Joris, 2009](#)).

3. Fluctuations in interaural differences with transformations

Psychological transformations were also included before the calculation of the interaural differences ([Goupell and Hartmann, 2007](#); [Goupell, 2010](#)). The first transformation

was laterality compression, which compresses the physical interaural difference to a perceptual scale as measured for static IPDs and ILDs by [Yost \(1981\)](#). This is done by using exponentially saturating functions ([Goupell and Hartmann, 2007](#)). Since the functional form of the perceived laterality appeared to be fairly consistent across CF (up to 5 kHz for ILDs) in the data from [Yost \(1981\)](#), the same laterality compression was used for all four CFs in the calculations. The second transformation was envelope weighting, which included IPDs only if the signals in the left and right channels were above 20% of the root-mean-square envelope amplitude. The rationale behind this transformation was to discount IPDs when the instantaneous signal energy was negligible, which helped describe the most variance in the data of [Goupell and Hartmann \(2007\)](#). After these transformations, the IPD and ILD are denoted $\Psi_{\Delta\Phi}$ and $\Psi_{\Delta L}$, respectively. Because the transformed interaural differences are on the same scale of perceived laterality from -10 to 10 , they can be easily combined. The sum of interaural standard deviations was calculated as:

$$SSD(a) = a \cdot s_t(\Psi_{\Delta\Phi}) + (1 - a) \cdot s_t(\Psi_{\Delta L}),$$

where the weighting of IPD and ILD is varied by a . Several forms of this model were used in modeling the data, which used or did not use envelope weighting. Also, the weighting factor a was varied between 0 and 1 in 0.1 steps.

B. Method

The interaural statistics were calculated over the actual GNs and LFNs used in experiment 1, each level of α having 25 tokens. The resulting metrics for the 500-Hz CF correlation change detection stimuli can be seen in Fig. 6 for ρ_{PP} (with Gammatone filtering, 0.46 compression, and envelope extraction), $s_t(\Delta\Phi)$, $s_t(\Delta L)$, and $SSD(0.5)$ (with envelope weighting). The data points represent the arithmetic average of the interaural statistic over the 25 tokens, and the error bars are two standard deviations in overall length. As was shown in previous reports (e.g., [Goupell and Hartmann, 2006](#); [Goupell, 2010](#)), stimuli with the same value of the interaural cross-correlation ρ can have a wide range of values in ρ_{PP} , $s_t(\Delta\Phi)$, and $s_t(\Delta L)$ when the BW is only 10 Hz.

Because only 25 tokens were used per signal level, the interaural metric functions were sometimes non-monotonic. To avoid problems with calculating the value of an interaural metric at the GN or LFN JND due to non-monotonicity, a third-order polynomial fit of an interaural metric vs α function was used. The LFN JND was predicted by finding the value at which the LFN stimuli could produce the same *change* in the interaural metric as the GN stimuli at the GN JND.

C. Results

To compare the ability of the models to describe the detection data, the percentage of variance described by each model was calculated ([Bernstein and Trahiotis, 1996](#)). For the ρ_{PP} and $SSD(a)$ models, the free parameters were allowed

FIG. 6. Values of four different interaural metrics (see text) for the 500-Hz CF stimuli used in experiment 1. Data points represent the average of the 25 tokens used in experiment 1 per level of α ; the error bars show ± 1 standard deviation.

to vary across CF and the parameters that had the best predicted JND were used. However, it was assumed that the same mechanism was used for detection of changes in correlation for both references. Thus, the free parameters had to describe both the reference conditions for a given CF. The criterion for the $\rho_{ref} = 1$ and 0 predictions were not required to be the same. The differences in the observed and predicted LFN JNDs are shown in Fig. 7. A perfect prediction would be a straight line through zero.

The cross-correlation without any processing stages (ρ) could only describe 34% and 23% of the variance in the average $\rho_{ref} = 1$ and 0 JNDs, respectively. This is better than the $s_t(\Delta L)$ metric that could describe none of the variance in the JNDs. The reason this occurred is that the GN JNDs needed a change in the value of $s_t(\Delta L)$ that was larger than could be produced by the LFN stimuli [compare open and closed symbols in Fig. 6(c)]. In other words, the LFN stimuli did not have large enough ILD fluctuations to predict the LFN JND from the GN JND using a constant criterion change. The $s_t(\Delta\Phi)$ metric could describe 90 and 38% of the variance in the average $\rho_{ref} = 1$ and 0 JNDs, respectively. The ρ_{PP} model could describe 63% and 92% of the variance in the average $\rho_{ref} = 1$ and 0 JNDs, respectively. The $SSD(a)$ model could describe 94% and 62% of the variance in the $\rho_{ref} = 1$ and 0 data, respectively. Therefore, the ρ_{PP} and $SSD(a)$ models are the best models for describing these data.

The major shortcoming of the ρ_{PP} model is that it over predicts LFN JNDs at and above 2000 Hz for $\rho_{ref} = 1$. The major shortcoming of the $SSD(a)$ model is that it over predicts the LFN JNDs at 500 Hz for $\rho_{ref} = 0$.

The parameters of the best models at each CF are listed in Table III. For the average LFN JND prediction, the ρ_{PP} model had no compression at 500 Hz and much more compression than 0.46 at higher CFs. Gammatone filtering and envelope extraction was for the most part not necessary to maximize the variance described. The $SSD(a)$ model used only IPDs at 500 Hz and shifted to an equal weight of IPD and ILD at higher CFs. The best model included envelope weighting at 500 and 2000 Hz but not at higher frequencies.

Predictions for the individual listeners and the good and bad listener groups are also shown on Fig. 7. The model effectiveness varied over the individual listeners and the good and bad listener groups. However, predictions are roughly consistent with the prediction for the average LFN JND. The best ρ_{PP} and $SSD(a)$ model parameters were different for individual listeners, and the good and poor listener groups, some of which can be seen in Table III.

It is possible to compare the best ρ_{PP} and $SSD(a)$ model predictions to the average, good, and bad data in Fig. 2. Most predictions for these two models are within the error bars of the LFN JNDs, which were large because of inter-listener variability.

FIG. 7. The difference in the observations and predictions for the LFN correlation change JNDs for the normalized cross-correlation, cross-correlation with peripheral processing, fluctuation in IPDs, fluctuations in ILDs, and sum of interaural fluctuations. Conditions for $\rho_{ref} = 1$ are in the top row. Conditions for $\rho_{ref} = 0$ are in the bottom row. Each dashed line represents predictions for an individual listener. The upward pointing triangles show predictions for the good listener group. The downward pointing triangles show predictions for the poor listener group. The circles show predictions for the average of all listeners. The percentage of variance described by a model for the average data is reported in each panel.

TABLE III. Model parameters that maximized the percentage of variance described in the correlation change data. The model components are compression (p), Gammatone filtering (GT), envelope extraction (Env. Ext.), IPD/ILD weighting (a), and envelope weighting (Env. Wt.). A “Y” denotes that the component was used in the model; an “N” denotes that it was not.

CF	ρ_{PP}				SSD(a)			
	Parameter	Good	Poor	Average	Parameter	Good	Poor	Average
500	p	0.25	0.6	1	a	1	0.9	1
	GT	Y	N	N	Env.Wt.	Y	Y	Y
	Env. Ext.	Y	N	N				
2000	p	0.75	0.25	0.25	a	0.6	0.5	0.6
	GT	Y	N	N	Env.Wt.	Y	Y	Y
	Env. Ext.	Y	Y	Y				
4000	p	0.6	0.25	0.1	a	0.7	0.7	0.8
	GT	N	N	N	Env. Wt.	Y	Y	N
	Env. Ext.	N	Y	N				
8000	p	0.1	0.25	0.1	a	0.6	0.3	0.7
	GT	N	N	N	Env. Wt.	N	Y	N
	Env. Ext.	N	Y	N				

D. Discussion

The modeling shows that both a cross-correlation model with compression (ρ_{PP}) and one that utilizes fluctuations in the interaural differences [SSD(a)] can be used to describe a majority of the variance in the correlation change detection data from experiment 1. These two models were shown to produce nearly isomorphic statistics in Goupell (2010), therefore it may not be surprising that both models could describe a majority of the variance in the data.

The cross-correlation model typically used to describe binaural detection data needed to be altered to describe the data in experiment 1. In Fig. 6, the cross-correlation model with 0.46 compression was inappropriate to describe the 500-Hz LFN JND; in fact the best form of this model utilized no compression at this low frequency for the average data. However, more compression was necessary at higher frequencies, up to compression values of 0.1. This result is consistent with the physiological finding that there is less compression in the apex of the cochlea and more in the base (Robles and Ruggero, 2001). Gammatone filtering and envelope extraction was not necessary in many of the conditions tested, which is likely a consequence of the narrow bandwidth and slow envelope fluctuations of the stimuli.

For the interaural fluctuation model, utilizing only IPDs provided the best prediction of the data at 500-Hz CF, consistent with van der Heijden and Joris (2009) but inconsistent with the reproducible noise experiments of Goupell and Hartmann (2007). At higher CFs, IPDs and ILDs were more equally weighted. This is interesting because one might expect listeners to rely solely on ILDs at high frequencies where phase locking cannot follow the fine-structure IPDs in the stimuli. An explanation for this is that the fluctuations in the IPDs and ILDs are not independent for randomly generated noises (Goupell, 2010) and IPDs could be transformed to ILDs at high CFs (van de Par and Kohlrausch, 1998). Envelope weighting was beneficial in the model predictions at 500 and 2000 Hz. It is unclear why discounting IPD fluctua-

tions at higher frequencies might be advantageous to the auditory system. This result suggests future modeling efforts should consider an envelope weighting threshold function that varies with frequency.

One shortcoming of modeling the correlation change detection JNDs was that there was a high inter-listener variability in the task, which can be seen in Fig. 2. Less variable data might have provided more convincing difference between model performances.

V. GENERAL DISCUSSION

A. Summary

Over the years, the normalized cross-correlation has been adapted by physiologically motivated transformations to account for many types of binaural data (e.g., Bernstein and Trahiotis, 1996; Bernstein *et al.*, 1999). One of these transformations is envelope compression. This addition was motivated psychophysically by the difference in GN and LFN NoS π detection thresholds (Eddins and Barber, 1998; Hall *et al.*, 1998) and can also explain so-called multiplied-noise masker binaural thresholds (van de Par and Kohlrausch, 1998). To investigate the GN-LFN threshold differences further, a tone in noise detection experiment was performed with both GN and LFN maskers at several CFs and BWs. The data showed that NoS π thresholds were 2–4 dB lower for GN compared to LFN at the 500-Hz CF but were the same or higher at the higher CFs. It is possible that listeners used monaural envelope cues for the high CF LFN NoS π conditions (Kohlrausch *et al.*, 1997), particularly if they had relatively high NoS π thresholds; however, a control experiment could not conclusively show this was the indeed the case.

Correlation change detection conditions were also tested, which should have minimized listeners’ potential to utilize monaural envelope fluctuation cues to perform the task. GN and LFN correlation change JNDs were the same at a 500-Hz CF for both correlated and uncorrelated references. There was a larger GN-LFN difference at higher CFs, where it was much more difficult to detect correlation changes in the LFN.

It was hypothesized that binaural adaptation may cause the relatively larger increase in LFN JNDs compared to GN JNDs at higher CFs. Using a pulsed-sine vocoder, the stimuli envelopes from experiment 1 were sampled at different pulse rates. There was no effect of pulse rate in the measured JNDs. Hence, the binaural adaptation hypothesis cannot explain the GN-LFN difference in JNDs in correlation change detection tasks.

In an attempt to describe the correlation change JNDs, numerical calculations on several interaural metrics were performed, including the normalized cross-correlation with physiologically-inspired transformations. The LFN JNDs were predicted from GN JNDs by assuming that listeners used the same change in the value of a binaural metric for detecting changes in correlation in both noise types. The cross-correlation model and a model that utilized the fluctuations in the IPDs and ILDs of the stimulus described a majority of the variance in the LFN JNDs. However, neither

model clearly outperformed the other. For the cross-correlation model, it was necessary to have a compression function that changes from no compression at a 500-Hz CF to highly compressive at and above a 2000-Hz CF to describe the average data. For the interaural difference model, it was necessary to use only IPDs at a 500-Hz CF and a more equal weighting of IPDs and ILDs at and above a 2000-Hz CF to describe the data, which might be expected from the duplex theory of sound localization (Rayleigh, 1907).

All of the binaural data and modeling results are consistent with strategy of “listening in the dips” (Hall *et al.*, 1998). Tones added to GN are likely to create relatively larger IPDs and ILDs during low level epochs of the noise. Thus, the signal is the most interaurally decorrelated in the low level epochs or the dips in the envelope. LFNs have no appreciable dips in the waveform, which results in the relatively smaller values of the interaural metrics (Fig. 6) compared to GNs. This ultimately produces higher thresholds and JNDs.

B. Implications for cochlear implants

High pulse rates, on the order of 1000 pps per electrode, are used to sample the envelope of incoming acoustic signals that have been bandpass filtered into several frequency channels in CIs. Stimulation strategies that use high pulse rates yield better speech understanding than lower pulse rates (Wilson *et al.*, 1991). Although CIs were not originally designed for bilateral use (e.g., the processors are not synchronized across the two ears, which results in random ITDs), more and more people are becoming bilaterally implanted. Studies suggest that bilateral CI users can understand speech slightly better than unilateral CI users, especially in noise (e.g., Schleich *et al.*, 2004; Litovsky *et al.*, 2009).

There might be the potential worry that binaural sensitivity would be stunted by binaural adaptation because of the high rate of stimulation combined with any signal compression from the stimulus encoding in bilateral CI users. If the results for the acoustic simulation of CI processing in experiment 2 apply to CI users, then binaural adaptation does not seem to occur for NoS π and correlation change detection for stimuli with only dynamic ILD cues. Stecker and Brown (2010) showed little to no binaural adaptation for dynamic ILDs using Gaussian pulses at a 4-kHz CF in NH listeners, consistent with the results of this study. It is possible that a different type of CI signal processing that introduces dynamic ITDs might show binaural adaptation (Carlyon and Deeks, 2002; van Hoesel, 2007).

Extrapolating the results of experiment 2 further to speech signals, the results of this study seem to imply that speech understanding in noisy environments would not be affected by binaural adaptation and thus not detrimentally affect CI users for this task. Since binaural adaptation seems to occur only for high-rate, constant-amplitude pulse trains, only the localization of unmodulated sounds using ITDs would suffer in CI users with present processing strategies (assuming bilateral synchronization). Nevertheless, bilateral CI users could still localize in the horizontal plane using ILD and envelope ITD information.

ACKNOWLEDGMENTS

The author would like to thank Corey Stoelb, Mitch Mostardi, Nick Liimatta, Garrison Draves, and Abbey Baus for help collecting data. Corey Stoelb was also vital in organization and plotting of the data. Antje Ihlefeld, Ruth Litovsky, Steve Colburn, associate editor Torsten Dau, and two anonymous reviewers provided helpful feedback on previous versions of this work. This work was supported by NIH Grant K99/R00 DC010206 (M. J. Goupell), R01 DC003083 (R. Y. Litovsky), and P30 HD03352 (Waisman Center core grant).

¹The orthogonalization process in the generation of interaurally decorrelated LFNs created stimuli that had values of Y in the right channel that were between 1.02 and 1.1. When compared to the left channel range of 1.02 to 1.06, it might have been possible to use the slight average change in Y in only the right ear to perform a $\rho_{ref} = 1$ condition. A monaural control experiment using only the right channel of 500-Hz CF, 10-Hz BW LFNs was performed to check if this occurred. The percentage of correct responses for four listeners was 57%, 56%, 56%, and 52% over 100 trials. Since this is below the 70.7% that is targeted in the adaptive procedure, the monaural cue would not be a reliable cue for the task. However, since three of these values are statistically above chance performance, it is possible that the cue from the change in Y in the right channel, which was introduced from the stimulus generation, could be used to decrease JNDs slightly for the $\rho_{ref} = 1$ conditions.

- Bernstein, L. R., and Trahiotis, C. (1982). “Detection of interaural delay in high-frequency noise,” *J. Acoust. Soc. Am.* **71**, 147–152.
- Bernstein, L. R., and Trahiotis, C. (1996). “The normalized correlation: Accounting for binaural detection across center frequency,” *J. Acoust. Soc. Am.* **100**, 3774–3784.
- Bernstein, L. R., and Trahiotis, C. (1997). “The effects of randomizing values of interaural disparities on binaural detection and on discrimination of interaural correlation,” *J. Acoust. Soc. Am.* **102**, 1113–1120.
- Bernstein, L. R., and Trahiotis, C. (2009). “How sensitivity to ongoing interaural temporal disparities is affected by manipulations of temporal features of the envelopes of high-frequency stimuli,” *J. Acoust. Soc. Am.* **125**, 3234–3242.
- Bernstein, L. R., and Trahiotis, C. (2010). “Accounting quantitatively for sensitivity to envelope-based interaural temporal disparities at high frequencies,” *J. Acoust. Soc. Am.* **128**, 1224–1234.
- Bernstein, L. R., van de Par, S., and Trahiotis, C. (1999). “The normalized interaural correlation: Accounting for NoS π thresholds obtained with Gaussian and ‘low-noise’ masking noise,” *J. Acoust. Soc. Am.* **106**, 870–876.
- 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.
- Brown, A. D., and Stecker, G. C. (2011). “Temporal weighting functions for interaural time and level differences. II. The effect of binaurally synchronous temporal jitter,” *J. Acoust. Soc. Am.* **129**, 293.
- Carlyon, R. P., and Deeks, J. M. (2002). “Limitations on rate discrimination,” *J. Acoust. Soc. Am.* **112**, 1009–1025.
- Culling, J. F., Colburn, H. S., and Spurchise, M. (2001). “Interaural correlation sensitivity,” *J. Acoust. Soc. Am.* **110**, 1020–1029.
- Davidson, S. A., Gilkey, R. H., Colburn, H. S., and Carney, E. (2009). “An evaluation of models for diotic and dichotic detection in reproducible noises,” *J. Acoust. Soc. Am.* **126**, 1906–1925.
- Domnitz, R. H., and Colburn, H. S. (1976). “Analysis of binaural detection models for dependence on interaural target parameters,” *J. Acoust. Soc. Am.* **59**, 598–601.
- Eddins, D. A., and Barber, L. E. (1998). “The influence of stimulus envelope and fine structure on the binaural masking level difference,” *J. Acoust. Soc. Am.* **103**, 2578–2589.
- Gabriel, K. J., and Colburn, H. S. (1981). “Interaural correlation discrimination: I. Bandwidth and level dependence,” *J. Acoust. Soc. Am.* **69**, 1394–1401.
- Goupell, M. J. (2010). “Interaural fluctuations and the detection of interaural incoherence. IV. The effect of compression on stimulus statistics,” *J. Acoust. Soc. Am.* **128**, 3691–3702.

- Goupell, M. J., Hancock, K. E., Majdak, P., Laback, B., and Delgutte, B. (2009a). "Binaurally-coherent jitter improves neural and perceptual ITD sensitivity in normal and electric hearing," in *The Neurophysiological Bases of Auditory Perception*, edited by E. A. Lopez-Poveda, A. R. Palmer, and R. Meddis (Springer, London), pp. 303–313.
- Goupell, M. J., and Hartmann, W. M. (2006). "Interaural fluctuations and the detection of interaural incoherence: Bandwidth effects," *J. Acoust. Soc. Am.* **119**, 3971–3986.
- Goupell, M. J., and Hartmann, W. M. (2007). "Interaural fluctuations and the detection of interaural incoherence. III. Narrowband experiments and binaural models," *J. Acoust. Soc. Am.* **122**, 1029–1045.
- Goupell, M. J., Laback, B., and Majdak, P. (2009b). "Enhancing sensitivity to interaural time differences at high modulation rates by introducing temporal jitter," *J. Acoust. Soc. Am.* **126**, 2511–2521.
- Goupell, M. J., Majdak, P., and Laback, B. (2010). "Median-plane sound localization as a function of the number of spectral channels using a channel vocoder," *J. Acoust. Soc. Am.* **127**, 990–1001.
- Haft, E. R., and Buell, T. N. (1990). "Restarting the adapted binaural system," *J. Acoust. Soc. Am.* **88**, 806–812.
- Haft, E. R., and Dye, R. H., Jr. (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.
- Haft, E. R., Dye, R. H., Jr., 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.
- Hall, J. W. III, Grose, J. H., and Hartmann, W. M. (1998). "The masking-level difference in low-noise noise," *J. Acoust. Soc. Am.* **103**, 2573–2577.
- Hartmann, W. M., and Pumplin, J. (1988). "Noise power fluctuations and the masking of sine signals," *J. Acoust. Soc. Am.* **83**, 2277–2289.
- Jain, M., Gallagher, D. T., Koehnke, J., and Colburn, H. S. (1991). "Fringed correlation discrimination and binaural detection," *J. Acoust. Soc. Am.* **90**, 1918–1926.
- Koehnke, J., Colburn, H. S., and Durlach, N. I. (1986). "Performance in several binaural-interaction experiments," *J. Acoust. Soc. Am.* **79**, 1558–1562.
- Kohlrausch, A., Fassel, R., van der Heijden, M., Kortekaas, R., van de Par, S., and Oxenham, A. J. (1997). "Detection of tones in low-noise noise: Further evidence for the role of envelope fluctuations," *Acta Acust. Acust.* **83**, 659–669.
- Laback, B., and Majdak, P. (2008). "Binaural jitter improves interaural time-difference sensitivity of cochlear implant users at high pulse rates," *Proc. Natl. Acad. Sci. U.S.A.* **105**, 814–817.
- Litovsky, R. Y., Parkinson, A., and Arcaroli, J. (2009). "Spatial hearing and speech intelligibility in bilateral cochlear implant users," *Ear Hear.* **30**, 419–431.
- Moore, B. C., and Glasberg, B. R. (1983). "Suggested formulae for calculating auditory-filter bandwidths and excitation patterns," *J. Acoust. Soc. Am.* **74**, 750–753.
- Pumplin, J. (1985). "Low-noise noise," *J. Acoust. Soc. Am.* **78**, 100–104.
- Rayleigh, L. (1907). "On our perception of sound direction," *Philos. Mag.* **13**, 214–232.
- Robles, L., and Ruggero, M. A. (2001). "Mechanics of the mammalian cochlea," *Physiol. Rev.* **81**, 1305–1352.
- Schleich, P., Nopp, P., and D'Haese, P. (2004). "Head shadow, squelch, and summation effects in bilateral users of the MED-EL COMBI 40/40+ cochlear implant," *Ear Hear.* **25**, 197–204.
- 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.
- Stecker, G. C., and Haft, E. R. (2002). "Temporal weighting in sound localization," *J. Acoust. Soc. Am.* **112**, 1046–1057.
- van de Par, S., and Kohlrausch, A. (1998). "Diotic and dichotic detection using multiplied-noise maskers," *J. Acoust. Soc. Am.* **103**, 2100–2110.
- van der Heijden, M., and Joris, P. X. (2009). "Interaural correlation fails to account for detection in a classic binaural task: Dynamic ITDs dominate N0S π detection," *J. Assoc. Res. Otolaryngol.* **11**, 113–131.
- van Hoesel, R. J. M. (2007). "Sensitivity to binaural timing in bilateral cochlear implant users," *J. Acoust. Soc. Am.* **121**, 2192–2206.
- van Hoesel, R. J. M., Jones, G. L., and Litovsky, R. Y. (2009). "Interaural time-delay sensitivity in bilateral cochlear implant users: Effects of pulse rate, modulation rate, and place of stimulation," *J. Assoc. Res. Otolaryngol.* **10**, 557–567.
- Wilson, B. S., Finley, C. C., Lawson, D. T., Wolford, R. D., Eddington, D. K., and Rabinowitz, W. M. (1991). "New levels of speech recognition with cochlear implants," *Nature* **352**, 236–238.
- Yost, W. A. (1974). "Discriminations of interaural phase differences," *J. Acoust. Soc. Am.* **55**, 1299–1303.
- Yost, W. A. (1981). "Lateral position of sinusoids presented with interaural intensive and temporal differences," *J. Acoust. Soc. Am.* **70**, 397–409.
- Yost, W. A., and Dye, R. H., Jr. (1988). "Discrimination of interaural differences of level as a function of frequency," *J. Acoust. Soc. Am.* **83**, 1846–1851.