Asymmetric temporal envelope encoding: Implications for within- and across-ear envelope comparison

Sean R. Anderson, Alan Kan, and Ruth Y. Litovsky[a)]

Waisman Center, University of Wisconsin-Madison, Madison, Wisconsin 53705, USA

Separating sound sources in acoustic environments relies on making ongoing, highly accurate spectro-temporal comparisons. However, listeners with hearing impairment may have varying quality of temporal encoding within or across ears, which may limit the listeners' ability to make spectro-temporal comparisons between places-of-stimulation. In this study in normal hearing listeners, depth of amplitude modulation (AM) for sinusoidally amplitude modulated (SAM) tones was manipulated in an effort to reduce the coding of periodicity in the auditory nerve. The ability to judge differences in AM rates was studied for stimuli presented to different cochlear places-of-stimulation, within- or across-ears. It was hypothesized that if temporal encoding was poorer for one tone in a pair, then sensitivity to differences in AM rate of the pair would decrease. Results indicated that when the depth of AM was reduced from 50% to 20% for one SAM tone in a pair, sensitivity to differences in AM rate decreased. Sensitivity was greatest for AM rates near 90 Hz and depended upon the places-of-stimulation being compared. These results suggest that degraded temporal representations in the auditory nerve for one place-of-stimulation could lead to deficits comparing that temporal information with other places-of-stimulation. © 2019 Acoustical Society of America. https://doi.org/10.1121/1.5121423

I. INTRODUCTION

Individuals with hearing impairment who are fitted with hearing aids or cochlear implants (CIs) generally perform more poorly than normal-hearing (NH) listeners on speech in noise tasks (Festen and Plomp, 1990; Loizou *et al.*, 2009). However, performance varies widely across individuals with hearing impairment, where some listeners demonstrate performance in the NH range and others perform much more poorly. This variability may be related to how well listeners can segregate target speech from background noise, which involves the use of source segregation cues (Bregman, 1990; Shinn-Cunningham, 2008). It is likely that sensitivity to and efficacy of source segregation cues is intimately related to how well these cues are encoded at the auditory periphery. That is, if the signal is poorly encoded in early stages of the auditory system, then access to source segregation cues will be limited through the rest of the central auditory system.

To date, the ability to compare simultaneously presented temporal information across cochlear places-of-stimulation with asymmetries in the state of temporal encoding has not been directly addressed. Each stage of the auditory system has developed specialized mechanisms to extract information necessary for the spectro-temporal comparisons involved with source segregation. Thus, it may be that the ability to process auditory stimuli and extract information necessary for source segregation is limited by the stages of auditory processing with greatest decrement. We therefore hypothesized that the ability to compare temporal information across frequency and/or ears may be limited by the place-of-stimulation with the poorest state of encoding. In this experiment, we attempted to quantify the effect of a reduction of temporal information by reducing the amplitude modulation (AM) depth, for amplitude-modulated tones. Reducing AM depth can be viewed as a reduction in the stimulus dynamic range, which is suspected to lead to less phase locking of the auditory nerve (Joris and Yin, 1992; Zilany *et al.*, 2014). We predicted that reducing AM depth at one cochlear place would reduce the ability of the auditory system to compare temporal envelope fluctuations.

A. Hearing impairment leads to peripheral degradations

The fidelity of temporal input to the auditory system of individuals with hearing impairment can vary widely across the frequency spectrum and between right and left ears according to etiology of disease, duration of deafness, and the type of hearing device being used. Thus, access to temporal information used to segregate sound sources may be limited for some or many frequencies in one or both ears within an individual.

Long durations of deafness and etiology of disease have been linked to deterioration of the auditory nerve (for review, see Shepherd and Hardie, 2001). One consequence of long durations of deafness is demyelination of axons, reduced arborization of dendrites, and shrinkage of cell bodies (e.g., Leake and Hradek, 1988). Several studies suggest that long durations of deafness, which result in deterioration of dendrites and demyelination of axons, lead to poorer phase locking of the auditory nerve (e.g., Shepherd *et al.*, 2004; Zhou *et al.*, 1995).

[a)]Electronic mail: litovsky@waisman.wisc.edu

Additional studies in rats indicate that long durations of deafness (>6 months) resulted in a reduction in the number of spiral ganglion cells in the auditory nerve (Shepherd et al., 2004). Human temporal bone studies confirmed that spiral ganglion cells tended to decrease as durations of deafness increased (Nadol et al., 1989) and depended on etiology (Spoendlin and Schrott, 1989). Spectral holes, or non-uniform losses in spiral ganglion cells in the auditory nerve, have also been demonstrated in populations with sensorineural hearing impairment (Nadol, 1997). Together, results from studies in animals and humans suggest that the ability to convey temporal information can vary highly across or within individuals due to etiology of disease or long durations of deafness.

Losses of auditory nerve fibers lead to an increase in threshold for the compound action potential and deterioration of peripheral processes that are associated with increased threshold for individual auditory nerve fibers (Goldwyn et al., 2010; Shepherd et al., 2004; Zhou et al., 1995). Sensorineural hearing loss due to damage of outer hair cells results in a change in loudness growth, increases in threshold, and reduced dynamic range (for review, see Oxenham and Bacon, 2003). Similarly, poor interface between CI electrodes and auditory nerve fibers is associated with reduced dynamic range in listeners with CIs (Bierer, 2010; Bierer and Nye, 2014). Thus, dynamic range of auditory nerve fibers is reduced in listeners with hearing impairment and may also vary with place-of-stimulation, depending upon the status of the auditory nerve. This suggests that the same temporal information (rate and depth of temporal fluctuations) must be represented in a smaller stimulus space. Accordingly, the central auditory system may lose access to information in the stimulus (i.e., via changes in spiking with the rate and depth of envelope fluctuations) because of the limitations in the peripheral auditory system.

Compared with acoustical stimulation, electrical stimulation leads to different neural representations for the same temporal information in stimuli. Modeling studies and extracellular recordings from the auditory nerve show that electrical stimulation of auditory nerve fibers gives rise to very little temporal jitter compared to normal, acoustic stimulation (Litvak et al., 2003). This difference in temporal envelope encoding due to acoustical vs electrical stimulation suggests that listeners that receive both methods of stimulation must reconcile differences in the representation of temporal information during sound source segregation. For example, the introduction of hybrid CIs has led to electric and acoustic hearing within the same ear. Thus, for patients with a hybrid CI, the same temporal envelope in the acoustic stimulus may have very different representations in the auditory nerve within the same ear. Similarly, for listeners with bimodal hearing or single-sided deafness that are fitted with a cochlear implant in the deaf ear (SSD-CI), acoustic and electric inputs must be integrated across the ears. Differences in temporal envelope encoding due to hearing device may be compounded by interactions with the previous factors, like duration of deafness and etiology of disease. It has even been suggested that introducing stochasticity to electric hearing via high rates of stimulation may improve hearing outcomes in listeners with CIs (e.g., Rubinstein et al., 1999), thereby making electrical stimulation more similar to acoustical stimulus representations. While it is well known that listeners with bimodal hearing, hybrid CIs (e.g., Gifford et al., 2013), SSD-CI, or bilateral CIs (e.g., Bernstein et al., 2016) perform worse than NH listeners in the same speech understanding tasks, the functional implications of differences in temporal encoding have never been studied systematically.

Hearing impairment can lead to differences in the physiology in each ear, and by extension the representation of incoming signals. How these differences impact real world listening where multiple sound sources are often present is unclear. However, temporal information in the envelope contributes to sound source segregation in several different ways and differences in fidelity of temporal encoding within and/or across ears may have a substantial impact on one's ability to segregate sound sources. In the next section, the role of temporal envelope cues in sound source grouping and segregation will be discussed briefly.

B. Temporal envelope information facilitates grouping and segregation of sounds

Temporal information in a signal can be decomposed into two categories: Envelope and fine structure. Broadly, the envelope refers to slow, gradual changes in the stimulus over time, whereas the fine structure refers to faster changes that fall within the envelope. From a signal processing perspective, signals can be decomposed as the product of the temporal envelope (akin to amplitude modulation) and fine structure (akin to carrier). The temporal envelope is preserved in CI processing (Loizou, 2006) and the role of the temporal envelope in speech perception has been studied in depth (e.g., Brungart, 2001; Drullman et al., 1994; Shannon et al., 1995). Temporal envelope fluctuations make up two aspects of information in speech, corresponding to slower and faster fluctuations (for review, see Rosen, 1992). For low AM rates, particularly those below 30–50 Hz, listeners perceive changes in loudness over time during a single stimulus presentation. However, at higher rates (above 50–90 Hz), listeners perceive a change in pitch.

Temporal envelope fluctuations are suspected to contribute to sound grouping and segregation in several ways (for review, see Grose et al., 2005). Repetitive fluctuations in the envelope over time result in the formation of an auditory stream for listeners with NH (Dollezal et al., 2012; Grimault et al., 2002; Nie and Nelson, 2015) and CIs (Chatterjee et al., 2006; Hong and Turner, 2009). Also, comodulation of temporal envelopes across frequency leads to better speech reception of sine wave speech (Carrell and Opie, 1992) and could be used to group sounds (Bregman et al., 1985). Conversely, signal disparities can be a cue for segregation. For example, listeners with NH are sensitive to phase disparities of sinusoidally amplitude-modulated tones (Strickland et al., 1989). Further, binaural envelope decorrelation, which is perceived as a diffuseness in the head or "roughness" (Goupell and Litovsky, 2015; Whitmer et al., 2014) may be related to detection of sounds with differing

phase in binaural masking level difference experiments. It has been demonstrated that NH listeners are sensitive to both monaural (Richards, 1987) and binaural (Bernstein and Trahiotis, 1992) envelope decorrelation. Listeners with CIs are also sensitive to binaural envelope decorrelation (Goupell, 2015; Goupell and Litovsky, 2015), but their sensitivity to monaural envelope decorrelation has not been studied.

C. Purpose of current study

Representation of temporal information important for sound grouping and segregation likely varies depending upon place-of-stimulation for listeners with hearing impairment. Hence, it is likely that these spectro-temporal comparisons in listeners with hearing-impairment must be completed with asymmetric representations of temporal information. Most research concerning sound source segregation in NH listeners has been conducted with individuals that have relatively symmetric representations of temporal information within and across ears and has focused on primarily symmetric representations of temporal information in the stimulus. Sound source segregation research in individuals with hearing impairment tends to focus on the type of prosthesis being used in either ear: bilateral hearing aids (e.g., Festen and Plomp, 1990; Reiss et al., 2017), bimodal hearing (e.g., Reiss et al., 2014), single-sided deafness and electrical-acoustic integration (Bernstein et al., 2015; Gifford et al., 2013), or bilateral CIs (Bernstein et al., 2016; Loizou et al., 2009). Beyond this, most studies do not address asymmetries in temporal encoding for all places-of-stimulation despite evidence that such asymmetries exist. Thus, the contributions of asymmetric temporal encoding on the ability to extract and compare information from co-occurring envelope fluctuations have not been explored systematically. It is possible that, given the heterogeneity in fidelity of temporal encoding across places-of-stimulation in the hearing-impaired auditory system, the limit to the central auditory system is imposed by places-of-stimulation providing poorest temporal encoding.

One way to understand the impact of these differences is to provide poorer temporal encoding of one sound when compared against a second sound in NH listeners. We attempted to model varying states of temporal envelope encoding (better and poorer) by reducing the depth of AM when presenting the stimulus to NH listeners. Data from the auditory nerve of cats indicated that as modulation depth decreased, the synchronization of firing decreased for all AM rates between 50 and 250 Hz (Joris and Yin, 1992). Further, smaller AM fluctuations are homologous to the reduced dynamic range observed with poor status of the interface between CI electrodes and the auditory nerve (Bierer, 2010; Bierer and Nye, 2014).

In this work, we used a new experimental paradigm similar to decorrelation detection to probe the auditory system. The goal of this paradigm was to explore listeners' ability to compare the rate of AM across ears and place-of-stimulation when temporal encoding at one cochlear place-of-stimulation was degraded. We hypothesized that the auditory system is limited in its ability to compare temporal envelopes by the fidelity of temporal encoding in the worse place-of-stimulation. Therefore, if amplitude modulation depth is lower in one cochlear place-of-stimulation in one ear, thereby reducing the fidelity of temporal encoding, then the sensitivity to differences in AM rate across ears and/or cochlear places-of-stimulation should decrease.

II. EXPERIMENT 1: A NORMAL-HEARING MODEL OF POOR TEMPORAL ENCODING

A. Motivation

The purpose of this experiment was to determine if reduced AM depth resulted in poorer ability to discriminate between AM rates presented to a single place-of-stimulation in NH listeners. Reducing AM depth increased rate discrimination thresholds for square wave amplitude-modulated noise in NH listeners and listeners with hearing impairment (Grant et al., 1998). Additionally, using noise carriers, it is possible that high frequency hearing loss could account for better performance with some listeners compared to others (for review, see Kohlrausch et al., 2002). The only experiment that has tested the effects of AM depth on rate discrimination using tone carriers varied AM depth relative to detection threshold (Füllgrabe and Lorenzi, 2003). Thus, it was necessary to determine whether the same effects hold for the stimuli employed in Experiment 2 of this paper.

B. Methods

1. Listeners

Twelve NH listeners (age 19–25 years; mean 22.4 years) participated in this experiment. All listeners had absolute detection thresholds less than or equal to 20 decibels hearing level [dB hearing level (HL)] for octave-spaced frequencies between 250 and 8000 Hz in the ear being tested. Before completing the experiment, all listeners provided informed consent. All procedures were approved by the Human Listeners Institutional Review Board of the University of Wisconsin-Madison.

2. Stimuli and procedures

Stimuli were sinusoidally amplitude-modulated (SAM) pure tones with carrier frequencies of 4000 or 7260 Hz. These carrier frequencies were chosen because they are disparate, with carrier frequency and sidebands falling outside equivalent rectangular bandwidths for each center frequency based on the equations provided by Moore and Glasberg (1990). Therefore, they were expected to result in a very small degree of overlap in stimulation on the basilar membrane when the two SAM tones were played simultaneously in the same ear for Experiment 2. Additionally, using carrier frequencies above 1500–3000 Hz should result in phase locking only to the envelope and not carrier tone. For each listener, one cochlear place-of-stimulation was chosen by using a carrier frequency of 4000 or 7260 Hz in the left or right ear. Six listeners were tested with carrier frequencies of 4000 Hz and the other six were tested with carrier

frequencies of 7260 Hz. Seven listeners were tested in the left and five in the right ear. Each SAM tone was presented at 65 dB sound pressure level A-weighted [dB(A)] for 600 ms. Stimuli were generated in MATLAB and presented using a Tucker-Davis Technologies System3 with RP2.1, HB7, and PA5 units (digital processor, amplifier, and attenuator, respectively) through Sennheiser HD600 open-back headphones. All testing took place in a double-walled, sound-attenuating booth (Industrial Acoustics Company, Inc.).

Two AM depths were tested: 50% and 20%, corresponding to -6 and -14 dB with respect to $20 \log_{10}(m)$, where m is the modulation depth in proportion. The AM depth of 20% was chosen because pilot experiments (data not shown) suggested this was the AM depth for which listeners began to show difficulty with discriminating AM rate but remained able to detect AM. The AM depth of 50% was chosen as a comparison to provide a control where listeners easily discriminated AM rates without making AM depths extremely different from one another (e.g., 20% vs 100% AM depth). Trials were blocked by AM depth condition and the order of AM depths was counterbalanced across listeners. One practice block was given with 50% AM depth prior to testing. Minimal training was desired so that results could be more easily generalized to listeners without extensive training in other experiments. Testing took approximately 0.5 to 1 h for each listener.

The experiment employed an "oddball" three-interval, two-alternative forced-choice task where listeners were instructed to choose the highest AM rate from either the second or third interval. The first interval always consisted of a standard AM rate of 10, 30, or 90 Hz. The second and third intervals presented the same standard rate as the first interval and a higher (oddball) AM rate, respectively, or vice versa with equal probability. The inter-stimulus interval was 300 ms within each trial. Visual feedback of the correct answer was given after each trial.

Three interleaved two-down, one-up adaptive tracking staircases corresponding to standard AM rates of 10, 30, and 90 Hz were used. Three blocks of three staircases were collected per listener: first, one practice block with 50% AM depth, next, two blocks with 50% and 20% AM depth, the order of which was chosen randomly with equal probability. Other than the first (practice) block, no data were excluded from analysis for Experiment 1. On each trial, one of the staircases that was not previously completed was selected to be presented to the subject with equal probability. For each staircase, 12 turnarounds were collected to estimate AM rate discrimination threshold. To ease the task of listeners and prevent confusion about the task throughout testing, step sizes were determined following the parameter estimation by sequential testing (PEST) rules outlined in Litovsky (1997). The maximum step size for each standard was 16/3, 16, and 48 Hz, and the minimum was 1/3, 1, and 3 Hz (i.e., the same relative change with the standard AM rates of 10, 30, and 90). The maximum oddball AM rate possible was five times the standard and the minimum was 31/30 of the standard. Discrimination thresholds for AM rate were computed using a weighted maximum likelihood procedure by fitting a two-

parameter logistic function to all data within each staircase and weighting each point by the number of observations. Curve fitting was completed using version 2.5.6 of the psignifit toolbox in MATLAB via the method described by Wichmann and Hill (2001). AM thresholds are presented as the base 10 logarithm of change in AM rate divided by standard AM rate ($\Delta f/f$) at threshold. This was done to normalize thresholds, allowing for fairer comparisons across standard AM rates.

C. Results

The goal of Experiment 1 was to determine the effects of AM depth on AM rate discrimination thresholds. For reference, thresholds for each subject in each condition are shown in Supplementary Fig. 1[1] and demonstrate the considerable variability across listeners. Summarized results are plotted in Fig. 1. The x axis of Fig. 1 corresponds to the standard AM rate tested. The y axis corresponds to the difference in threshold between the 50% and 20% AM depth conditions; positive values indicate poorer performance with 20% AM depth. Error bars shown in the plot correspond to 95% confidence intervals for the difference in threshold between the 50% and 20% AM depth conditions. Change in threshold was expressed in units of $\log(\Delta f/f)$ as in Eq. (1)

$$x = \log_{10}\left(\frac{\Delta f_m}{f_m}\right), \tag{1}$$

to normalize between the different standard AM rates. A value of 0 would indicate that threshold (Δf) was equal to the standard AM rate, and a value of 1 would indicate that threshold was ten times the standard AM rate. A mixed-effects analysis of variance (ANOVA) was completed with fixed effects of AM depth and standard AM rate within-listeners, as well as presentation ear and carrier frequency across-listeners, and $\log(\Delta f/f)$ threshold as the dependent variable. A random effect of listener was included to account
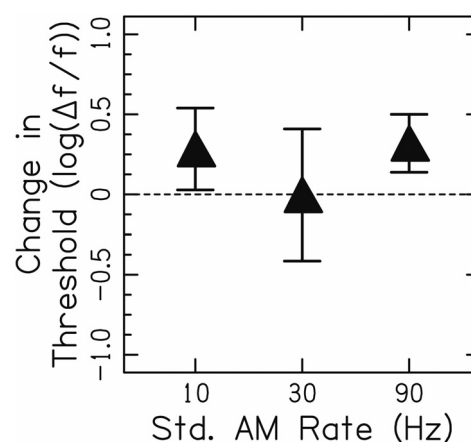

FIG. 1. 95% confidence interval for change in threshold between the 20% and 50% AM depth. The x axis corresponds to the three possible standard AM rate (i.e., the AM rate repeated twice in the three-interval, two-alternative forced-choice task). The y axis corresponds to the difference in threshold between the 20% and 50% AM depth conditions. Confidence intervals were computed assuming a t distribution. Confidence intervals for standard AM rates of 10 and 90 Hz that fall outside of zero imply that there was a significant increase in threshold as AM depth was reduced from 50% to 20%.

for variability across listeners. Thresholds for AM rate increased as AM depth decreased from 50% to 20% [$F(1,55)$ = 4.912, $p < 0.05$]. Anecdotal reports from most listeners suggested that the 20% AM depth condition was noticeably more difficult and required closer attention to complete the task.

Standard AM rate also significantly affected threshold [$F(2,55) = 4.908$, $p < 0.05$], with the 10 Hz standard AM rate having a significantly higher threshold than 90 Hz [$t(55)$ = 3.123, $p < 0.01$], but no significant differences between 10 and 30 Hz [$t(55) = 1.776$, $p = 0.188$] or 30 and 90 Hz [$t(55) = 1.348$, $p = 0.375$]. All *post hoc* tests were completed by comparing estimated marginal means with Tukey adjustment for multiple comparisons. There was no statistically significant interaction between the effects of standard AM rate and AM depth [$F(2,55) = 1.278$, $p = 0.287$]. The large decrease in change in threshold plotted in Fig. 1, suggestive of an interaction between AM rate and AM depth, is due primarily to subject TQB. Subject TQB had a substantially lower threshold for 20% AM depth compared to 50% (see Supplementary Fig. 1[1]), and may have been due to an ill fit of the maximum likelihood function to listener TQB's adaptive tracking data.

To rule out effects of differences between carrier frequencies, the choice of carrier frequency was counterbalanced across listeners. The effects of ear [$F(1,9) = 0.555$, $p = 0.478$] and carrier frequency [$F(1,9) = 0.090$, $p = 0.772$] were not statistically significant across individuals.

D. Discussion

Experiment 1 investigated whether the ability to discriminate changes in AM rate presented over sequential intervals becomes worse as the depth of AM decreases. Thresholds for AM rate discrimination relative to three standard rates (10, 30, and 90 Hz) were estimated for two different AM depths: 20% and 50%. Results indicated that when AM depth decreased, there was an increase in threshold for discriminating changes in AM rate, consistent with the prediction that poorer encoding of the temporal envelope results in poorer ability to discriminate changes in rate of fluctuations. This result is consistent with previous reports in NH and listeners with hearing impairment using square wave amplitude-modulated noise (Grant *et al.*, 1998), and listeners with NH using sinusoidally amplitude-modulated tones (Füllgrabe and Lorenzi, 2003).

There are several reasons why a reduction in AM depth might result in more difficulty in discriminating AM rate for SAM tones (for review, see Carlyon and Deeks, 2002). As mentioned previously, smaller AM depth resulted in poorer phase locking to AM in extracellular recordings of the auditory nerve from the cat (Joris and Yin, 1992) and computational models (Zilany *et al.*, 2009). Thus, one interpretation of our result suggests that poorer phase locking of the auditory nerve would worsen psychophysical performance. This conclusion is supported by studies with listeners that have CIs, showing a significantly higher percentage of correct discriminations for square wave and sawtooth waveforms, which have sharp onsets (Landsberger, 2008). Similar

improvements in performance due to sharper envelopes have been shown in tasks involving interaural timing difference discrimination in listeners with NH and CIs (Bernstein and Trahiotis, 2002; Bernstein and Trahiotis, 2009; Laback *et al.*, 2011). These improvements in performance were presumably driven by an improvement in phase locking by the auditory nerve.

Another reason reduced AM depth might worsen the ability to discriminate AM rates presented sequentially relates to sidebands in the magnitude spectrum. Lower AM depths result in lower magnitude of sidebands in the magnitude spectrum, and higher AM rates increase the distance between the center frequency band and sidebands. However, data from Kohlrausch *et al.* (2002) suggest that when modulation depth is presented near detection threshold, sidebands would have been useful in discrimination only for the highest AM rates presented in this study (roughly >300 Hz for a carrier of 4000 Hz). In their experiments, sidebands were audible at substantially lower AM depths than those used here. From Supplementary Fig. 1[1] it appears that only participants TJI at 20% and 50% AM depth and TSF at 20% AM depth had thresholds with the standard AM rate of 90 Hz in the frequency region where sidebands would be a useful cue.

Additionally, it is thought that non-linear frequency processing in the cochlea introduces combination tones that vary depending upon the relationship between carrier frequency and higher AM rates (Lee, 1994; Ruggero *et al.*, 1992; Strickland and Dhar, 2000; Strickland and Viemeister, 1997). Lee (1994) attempted to mask combination tones by introducing low-pass noise in one of the conditions in her experiment. Adding masking noise increased threshold for AM rate discrimination at higher AM rates (160 Hz and 320 Hz when the carrier frequency was chosen randomly, 160 Hz at 500 Hz carrier frequency). However, the magnitude of combination tones is not computable, so the addition of masking noise may have masked difference tones at lower frequencies or interfered with the listener's ability to assess pitch. Low-frequency noise was also added in experiments concerning AM detection with the goal of masking combination tones (Strickland and Dhar, 2000; Strickland and Viemeister, 1997). Results from the Appendix of Eddins (1999) suggest that low-frequency distortion products (i.e., combination tones) would not contribute to AM detection for the rates used in our study (though the stimuli in the former study were noise and quasifrequency-modulated).

Results from Experiment 1 also indicated that thresholds were significantly higher for the 10 Hz compared to the 90 Hz standard AM rate. Increasing thresholds as the AM decreases below 100 Hz are consistent with previous literature on rate discrimination in general (see data reviewed by Carlyon and Deeks, 2002; Krumbholz *et al.*, 2000; Lemańaska *et al.*, 2002). Lee (1994) reported that thresholds were relatively similar across AM rate when expressed in terms of $\Delta f/f$ using SAM tones. However, the data in experiment one of their manuscript show that each of the three subjects have a small but consistent increase in threshold between the 10 and 20 Hz AM rate compared against the 80 or 160 Hz AM rate for the 4000 Hz carrier. These standard AM rates from Lee (1994) were the most similar to the

present experiment and indicate a consistent pattern with our data. Similarly, Lemańaska et al. (2002) describe that their data demonstrate a near-miss to Weber's law (or a change in threshold depending upon standard AM rate).

While the results were qualitatively consistent with previous reports (Füllgrabe and Lorenzi, 2003; Grant et al., 1998; Lee, 1994; Lemańaska et al., 2002), AM rate discrimination thresholds were much higher on average. In the present study, thresholds at 50% AM depth were between two and seven times greater than those of Lemańaska et al. (2002) and Füllgrabe and Lorenzi (2003), and roughly ten times higher than those from Lee (1994). It is important to note that several studies using SAM tones at 100% AM depth (Lemańaska et al., 2002), with varying AM depth (Füllgrabe and Lorenzi, 2003), SAM noise with 100% AM depth (Lemańaska et al., 2002), and square wave AM with varying AM depth (Grant et al., 1998) also showed higher thresholds than those reported by Lee (1994). One important difference between the present study and that of Lee (1994) and Lemańaska et al. (2002) is that prior studies used 100% modulation depth, whereas here the two modulation depths were 50% and 20%. Decreasing AM depth by 30% for SAM noise resulted in an approximate 2–3 fold increase in thresholds in previous studies (Füllgrabe and Lorenzi, 2003; Grant et al., 1998), suggesting that the difference between our results and previous results could have been driven at least in part by AM depth differences. Results from our experiment led to an effect of AM depth with similar or smaller magnitude. One other possibility for this difference from previous reports is that listeners used different listening strategies for each standard rate and interleaving standard AM rates in this study made this task more difficult than previous tasks. In agreement with this idea, the study by Lee (1994) found that when carrier frequencies were randomized within the same block of trials, thresholds for discrimination increased, although thresholds were still lower than in the present study. It is also possible that loudness across AM rate played a role in the results of this experiment. The potential impact of loudness cues is explored in the Discussion of Experiment 2. Finally, it should be noted that the listeners in the present experiment did not receive extensive training on the task.

III. EXPERIMENT 2: COMPARING TEMPORAL ENVELOPE ACROSS PLACES-OF-STIMULATION WITH ASYMMETRIC TEMPORAL ENCODING

A. Motivation

Experiment 1 established that decreased modulation depth of SAM tones led to poorer performance on an AM rate discrimination task for AM rates presented sequentially (over three intervals) in NH listeners. In this experiment, we applied these findings to examine whether the ability to make simultaneous comparisons across two places-of-stimulation might be affected by temporal envelope representations in one cochlear place-of-stimulation.

It was predicted that the auditory system relies on similar quality of representations of temporal information between places-of-stimulation. Thus, if the temporal encoding is worse in one place-of-stimulation, then the ability to make these

comparisons should be limited by the state of temporal encoding in the worse place-of-stimulation. Therefore, it was hypothesized that if AM depth was reduced in one of two places-of-stimulation, then sensitivity to differences in AM rate presented between places-of-stimulation would decrease. Further, it was predicted that if listeners compared AM rates at the same place-of-stimulation across the ears, listeners would have greater sensitivity to differences in AM rate overall due to the presence of an additional binaural beat cue in the envelope (Mcfadden and Pasanen, 1975). Binaural beats refer to the perception of a moving or diffuse sound source when the AM rate differs by a small number of cycles per second, introducing an interaural timing difference that varies over time. The addition of a contralateral cue in a previous rate discrimination experiment has also improved ability to discriminate AM rate when presented in sequence, as in Experiment 1 (Carlyon and Deeks, 2002), supporting the notion that binaural AM rate discrimination is improved relative to tasks like that used in Experiment 1.

B. Methods

1. Listeners

Eleven listeners (age 18–26; mean 21.2 years) met the same hearing screening criteria as in Experiment 1, but with the addition of having 10 dB or less asymmetry in absolute pure-tone detection thresholds across the ears.

2. Stimuli and procedures

Stimuli were presented using the same equipment as Experiment 1. SAM tones of 600 ms duration with carrier frequencies of 4000 or 7260 Hz were played to the left and/or right ear (depending on the condition being tested). The starting phase of the envelope for each SAM tone was randomly selected from a uniform distribution between 0 and 2π radians for each presentation. The level was set to 65 dB(A) overall for any stimuli presented monaurally and attenuated by 6 dB in both ears when presented binaurally to result in similar loudness for monaural and binaural stimuli.

This experiment used a one-interval, two-alternative forced-choice task, where the listener was presented with SAM tones in two different cochlear places-of-stimulation and they responded by indicating whether the two AM rates were the same or different. Listeners were given visual feedback after each response. The AM rate in one place-of-stimulation was a fixed standard of either 10 or 90 Hz and the other (variable) place-of-stimulation received an AM rate equal to or greater than the standard rate. A psychometric function was measured over varying values of $\log(\Delta f/f)$. There was a 0.5 probability of both places-of-stimulation having an equal AM rate on each trial.

There were three possible place-of-stimulation pairing configurations: Same place-of-stimulation across-ears, different place-of-stimulation across-ears, and different place-of-stimulation within-ears. Both the task and AM pairing configurations are outlined in Fig. 2 for a "different" trial (where the AM rate differed between places-of-stimulation).
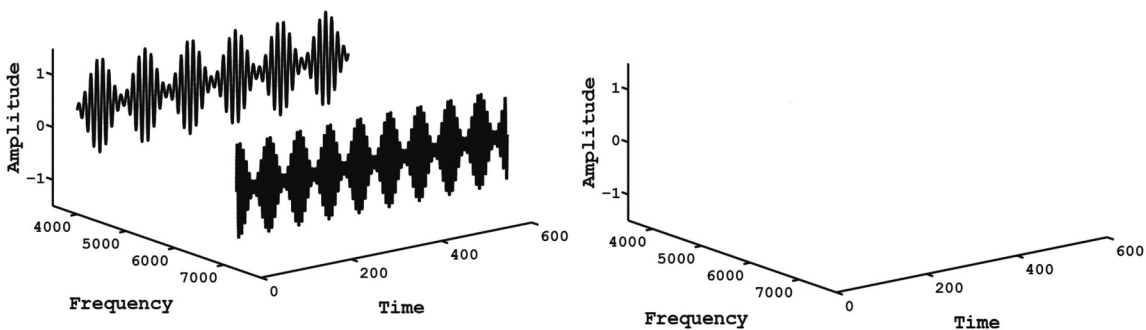
Same Place, Across Ears
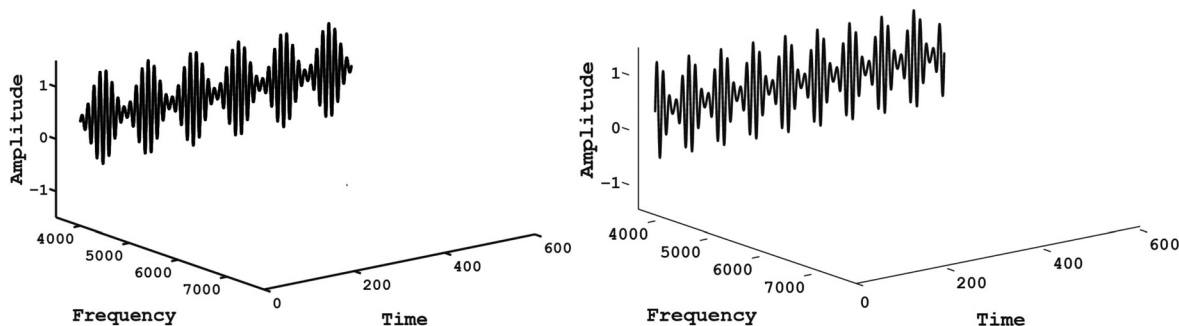
Different Place, Across Ears
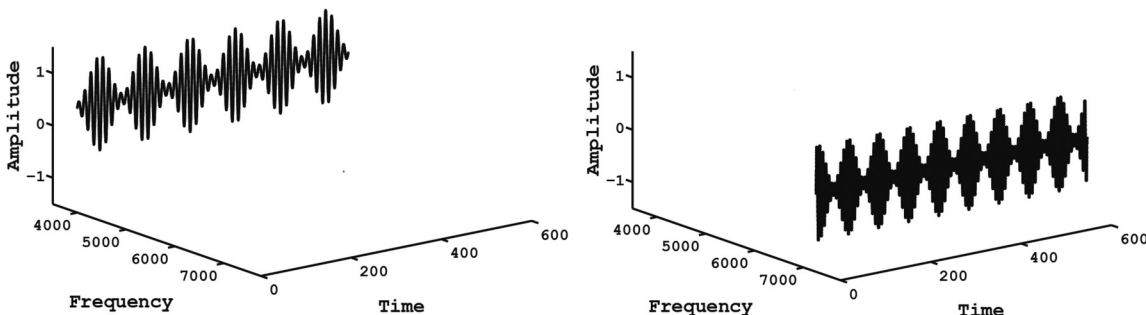
Different Place, Within Ears

Left Ear   Right Ear

FIG. 2. Illustration of AM pairing configurations. The *x* axis represents time (each stimulus had a duration of 600 ms). The *y* axis represents carrier frequency (of either 4000 or 7260 Hz). The *z*-axis represents relative amplitude of the stimulus. Each row represents a different AM rate pairing configuration. The left and right column correspond to the left and right ear, respectively. Same Place, Across Ears: Carrier frequencies were equal in both ears (either 4000 or 7260 Hz). This configuration may have resulted in the perception of binaural beats and represents temporal envelope comparisons that occur for matched place-of-stimulation across the ears. Different Place, Across Ears: Different carrier frequencies were used in each ear and this configuration represents temporal envelope comparisons completed across spectrum and the ears. Different Place, Within Ears: Different carrier frequencies were used within the same ear and this configuration represents comparisons completed across spectrum but within the same ear.

A split-plot experimental design was used with AM configuration as the whole-plot variable, and the AM depth and the ear receiving the variable AM rate as the split-plot variables (Table I). Within each listener, the place-of-stimulation containing the variable AM rate was counterbalanced. For example, in the same place, across ears condition depicted in Fig. 2, half of the time the variable AM rate was presented on the left side, and the other half on the right side. By extension, in the first and second halves of the trials, the standard AM rate was presented to the right and left, respectively.

This was done to account for any differences in performance depending upon where the standard AM rate was delivered. Location of the standard AM rate was blocked to prevent confusion for listeners (for example, see Table I).

All AM pairing configurations were repeated, once with 50% AM depth in both places-of-stimulation, and once with 20% AM depth in one place-of-stimulation and 50% AM depth in the other. The carrier frequency and ear of the place-of-stimulation with 20% AM depth, and order of AM pairing configurations were counterbalanced across listeners.

TABLE I. Example experimental conditions. Each row corresponds to a block of 420 trials. On each trial, the variable AM rate place-of-stimulation would present the standard AM rate with a probability of 0.5 or one of the variable rates shown in Table II.

Block	Place-of-Stimulation Pairing	Variable AM Rate			Std. AM Rate		
		Ear	Frequency (Hz)	AM Depth (%)	Ear	Frequency (Hz)	AM Depth (%)
1	Same Place, Across Ears	L	4000	50	R	4000	50
2	Same Place, Across Ears	L	4000	20	R	4000	50
3	Same Place, Across Ears	R	4000	50	L	4000	50
4	Same Place, Across Ears	R	4000	50	L	4000	20
5	Different Place, Across Ears	L	4000	50	R	7260	50
6	Different Place, Across Ears	L	4000	20	R	7260	50
7	Different Place, Across Ears	R	7260	50	L	4000	50
8	Different Place, Across Ears	R	7260	50	L	4000	20
9	Different Place, Within Ears	L	4000	50	L	7260	50
10	Different Place, Within Ears	L	4000	20	L	7260	50
11	Different Place, Within Ears	L	7260	50	L	4000	50
12	Different Place, Within Ears	L	7260	50	L	4000	20

Only one place-of-stimulation was chosen at the outset for each listener to receive the AM depth manipulation (either left or right ear with 4000 or 7260 Hz). The corresponding places-of-stimulation in all other configurations were then chosen accordingly. An example of the order and conditions tested is shown in Table I. For this example, the left ear with a 4000 Hz carrier was chosen for the AM depth manipulation; then in half of all conditions the AM depth was 50% and the rest were 20% for that ear and carrier frequency. In the same place, across ears configuration, the other place-of-stimulation was 4000 Hz in the right ear. For the different place, across ears configuration, the other place-of-stimulation was 7260 Hz in the right ear. For the different place, within ears configuration, the other place-of-stimulation was 7260 Hz in the right ear. In all of these cases, the "other place-of-stimulation" was presented with 50% AM depth.

A seven-point psychometric function (see Table II for AM rates) was collected, with each point representing a change in AM rate relative to 10 or 90 Hz standard AM rates. Thirty repetitions were collected per variable AM rate in each condition, resulting in 840 trials per condition, with three possible AM pairing configurations and two possible AM depth conditions. This resulted in a total of 5040

stimulus presentations throughout the experiment. Testing took approximately 6 h, with 1–2 h of training. Testing was divided over two to three sessions on different days.

The goal in every trial was to correctly identify whether the AM rate was the same in both places-of-stimulation. If AM rates were different, the variable place-of-stimulation had a higher AM rate. Trials for both of the standard AM rates were interleaved within one block. The AM rate on each trial was roved according to Table II. Interleaving standard AM rate and rate roving were both completed to discourage listeners from using a single-place-of-stimulation strategy to complete the task (e.g., responding "different" for high AM rates and "same" for low AM rates). Rate roving was not completed for one listener (THK) because they were tested before rate roving was implemented. We attempted to ensure that participants were completing the task by comparing the two SAM tones to one another. If participants exhibited non-monotonic functions during the beginning of a new block, they were given a break, a practice block where only the two highest values of $\log(\Delta f/f)$ were included, and data for that block were collected *de novo*. Unfortunately, some raw data still exhibited non-monotonic trends.

TABLE II. AM rates corresponding to each $\log(\Delta f/f)$ for Experiment 2. The columns show the extent of AM roving for 10 and 90 Hz standard AM rates. The N/A row refers to values taken on by the standard AM rate [technically $\log(\Delta f/f) = -\infty$].

$\log(\Delta f/f)$	Std. Rate: 10 Hz			Std. Rate: 90 Hz		
	Min. (Hz)	Mean (Hz)	Max. (Hz)	Min. (Hz)	Mean (Hz)	Max. (Hz)
N/A (Std.)	7.50	10.00	13.33	67.50	90.00	120.00
-0.71 ± 0.28	8.95	11.94	15.92	80.59	107.45	143.27
-0.42 ± 0.28	10.37	13.82	18.43	93.31	124.41	165.88
-0.12 ± 0.28	13.16	17.54	23.39	118.40	157.86	210.48
0.17 ± 0.28	18.65	24.87	33.16	167.87	223.83	298.44
0.47 ± 0.28	29.49	39.33	52.43	265.44	353.93	471.90
0.76 ± 0.28	50.87	67.83	90.44	457.85	610.47	813.96
1.06 ± 0.28	93.03	124.05	165.39	837.31	1116.41	1488.55

C. Results

In this experiment, pairs of SAM tones were presented across the ears at the same or different places-of-stimulation, or within the same ear at different places along the cochlea (see Fig. 2). Examples of raw data from two listeners in the same place, across ears condition are presented in Fig. 3 showing that performance could be quite variable. Some listeners performed extremely well (e.g., listener TIG), with a small amount of bias and large number of correct responses. Others performed very poorly (e.g., listener THK), with substantial bias toward choosing "different" and small number of correct responses. As a reminder, THK was tested before AM rate roving was implemented. Individuals varied considerably, and some psychometric functions were non-monotonic. Raw data from all listeners are provided in Supplementary Fig. 2.[1] Notably, non-monotonic psychometric functions were more likely to occur in the 20%:50% AM depth condition.

Figure 4 shows sensitivity by AM pairing configuration, AM depth in each ear, and $\log(\Delta f/f)$. Proportion of correct responses were converted to d' (Green and Swets, 1966). The data in Fig. 4 suggest that sensitivity was greater for standard AM rates of 90 compared with 10 Hz, demonstrated in all three AM pairing configurations by a higher asymptotic sensitivity for the 90 Hz standard AM rate conditions. Further, average sensitivity was greater in the 50%:50% compared against the 20%:50% AM depth conditions for all AM pairing configurations, demonstrated by a rightward shift for the open symbols and dashed lines.

To determine the effects of $\log(\Delta f/f)$ on d' and interaction with blocking variables (standard AM rate, AM depth, and AM pairing condition), $\log(\Delta f/f)$ vs d' were fit with a four-parameter logistic mixed effects model with each combination of blocking variables, where the mid-point was a random effect by listener (i.e., data from Fig. 4).

Unfortunately, the covariance structure did not converge, and it would not be appropriate to report the results of the model. As a second approach, mixed effects ANOVA was completed with d' as the dependent variable and $\log(\Delta f/f)$, AM pairing configuration, AM rate standard, and reduction in AM depth as fixed effects, with listener as a random-effect. The effect of carrier frequency and ear for the place-of-stimulation across listeners were also explored as fixed effects. Mixed effects ANOVAs fit fewer parameters compared to the four-parameter logistic model, a substantially less complex covariance structure, and converged without issue in this case. Results indicated that $\log(\Delta f/f)$ had a significant effect on d' $[F(6,842) = 402.711, p < 0.0001]$ on average across all blocking variable combinations. This was unsurprising because sensitivity to differences in AM rate should increase as the degree of difference between standard and variable AM rates increases.

To confirm that using different carrier frequencies representing the place-of-stimulation with poorer phase locking did not affect performance, a fixed effect of carrier frequency was included in the ANOVA. Effects of carrier frequency were tested because previous experiments investigating the ability to discriminate interaural time differences in the envelope showed differences in sensitivity depending upon carrier frequency (e.g., Bernstein and Trahiotis, 2009). Results indicated that carrier frequency with the AM depth manipulation did not have a significant effect across listeners $[F(1,9) = 0.319, p = 0.586]$. This result, along with the lack of a significant effect of carrier frequency in Experiment 1, suggests that listeners were able to compare AM rates regardless of carrier frequencies employed in this experiment. It is important to note that only 11 people participated in Experiment 2 and carrier frequency with the AM depth manipulation was an across-listener variable, so variability across listeners could have masked an effect.
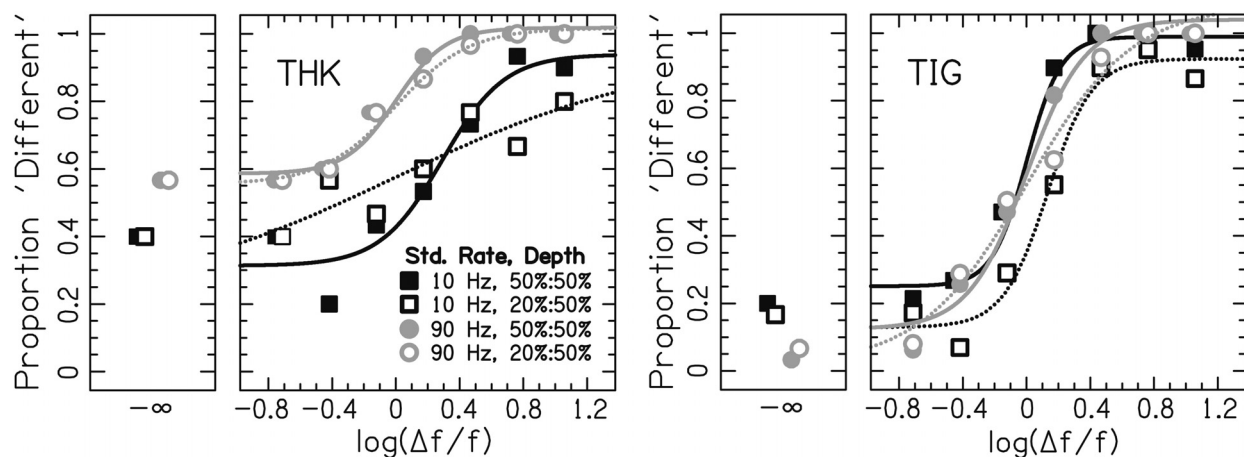

FIG. 3. Example raw data from two listeners in the same place, across ears condition. Listener codes for each individual are given in the top left corner. Open shapes and dotted lines correspond to the 20%:50% AM depth condition, and closed shapes and solid lines correspond to the 50%:50% AM depth condition. Performance for 10 Hz standard rate is shown in black and 90 Hz standard rate is shown in grey. The y axis corresponds to the proportion of "different" responses across all trials. The x axis corresponds to the difference in AM rate between the standard and variable AM rate (see Table I for values of AM rates in Hz). The small panel on the left [a $\log(\Delta f/f)$ of $-\infty$] represents the proportion of "different" responses when AM rates were equal. Ideal performance occurs when the proportion of "different" responses is 0 for the small left panel, and 1 for the larger, right panel. Sensitivity (d') can be calculated directly from the raw data. For raw data for all listeners, please see Supplementary Fig. 2 (Footnote 1).
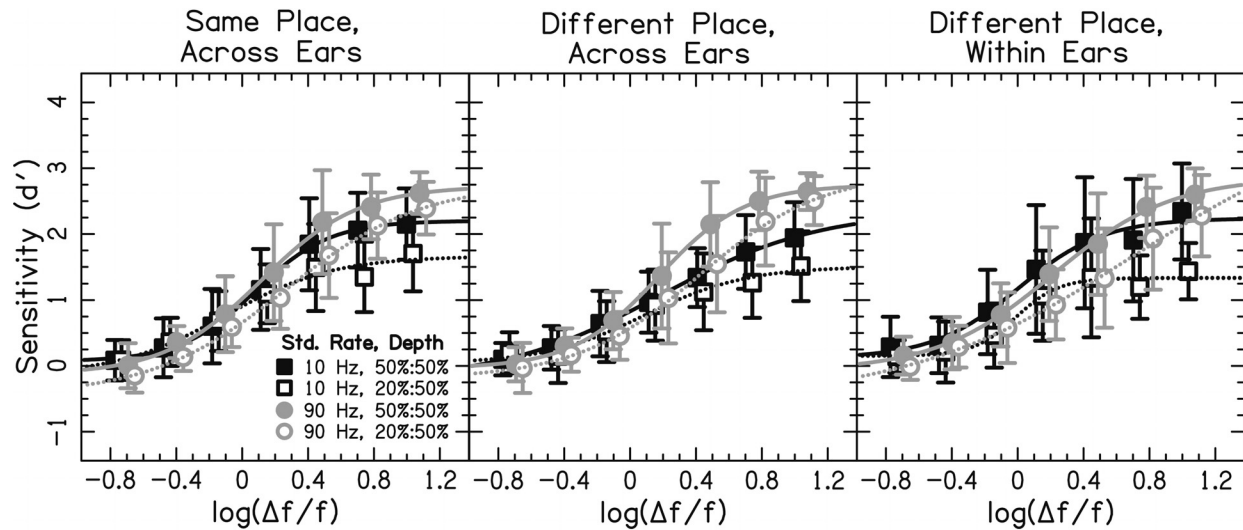
J. Acoust. Soc. Am. **146** (2), August 2019

Anderson *et al.*   1197

FIG. 4. Mean ± one standard deviation sensitivity across listeners for all three conditions. Open shapes and dotted lines correspond to the 20%:50% AM depth condition, and closed shapes and solid lines correspond to the 50%:50% AM depth condition. Performance for 10 Hz standard rate is shown in black and 90 Hz standard rate are shown in grey. Each panel corresponds to a different AM pairing condition (see Fig. 2). The *y* axis corresponds to sensitivity in *d'* (Green and Swets, 1966). The *x* axis corresponds to the difference in AM rate between the standard and variable AM rate (see Table I for values of AM rates in Hz).

1. Effects of asymmetric temporal encoding

The primary purpose of this experiment was to explore if sensitivity to differences in AM rate between places-of-stimulation decreases when AM depth is reduced for one place-of-stimulation. Reduced AM depth significantly decreased *d'* [$F(1,842) = 81.920$, $p < 0.0001$], suggesting that sensitivity to differences in AM rate between places-of-stimulation decreased on average when depth was reduced in only one of the two places-of-stimulation. There was also a significant interaction between $\log(\Delta f/f)$ and reduced AM depth [$F(1,842) = 4.569$, $p < 0.001$], suggesting that the slope of the psychometric function changed when AM depth was reduced. Due to the limitations of using an ANOVA compared with a four-parameter logistic model, a comparison of the slope of psychometric functions fitted within the same model between AM depth conditions was not possible.

Mean thresholds for both standard AM rates and AM pairing configurations are reported in Fig. 5. Thresholds (0.707 proportion correct) were estimated by fitting a four-parameter logistic function to the proportion of "different" responses across variable AM rates within each subject (data and fitted curves shown in Supplementary Fig. 2[1]). Five thresholds with standard AM rates of 10 Hz (THQ different place, across ears; THQ different place, within ears; THS different place, within ears; THU different place, across ears; TJF same place, across ears) could not be estimated because the logistic curve did not cross 0.707 for $\log(\Delta f/f)$ values tested in the experiment (Supplementary Fig. 2[1]). All thresholds in the same AM pairing configuration (for 20%:50% or 50%:50% AM depth) where thresholds could not be estimated were excluded from analyses for that subject. Thresholds are generally consistent with the results of the ANOVA, although the considerable variability across subjects is clear from Fig. 5.

2. Effects of standard AM rate

In general, for AM rates above 50–90 Hz the listener experiences a pitch cue. Thus, the listening strategy in this experiment was expected to change depending upon the standard AM rate. Normalizing differences in AM rate using
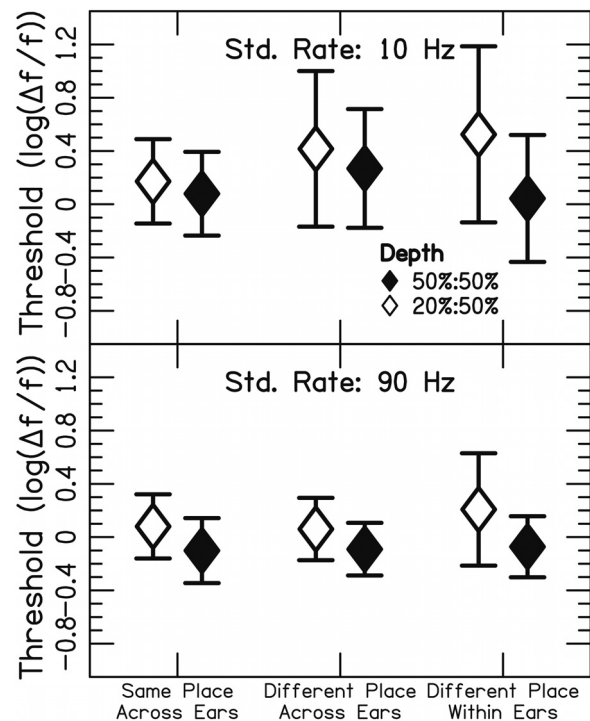

FIG. 5. Mean ± one standard deviation thresholds for each standard AM rate. Top and bottom rows correspond to standard AM rates of 10 and 90 Hz, respectively. Open and closed shapes correspond to the 20%:50% and 50%:50% AM depth conditions. The *y* axis corresponds to threshold (defined at 0.707 proportion correct). The *x* axis corresponds to the AM pairing condition (see Fig. 2). Thresholds were highly variable across listeners (for raw data for each listener, see Supplementary Fig. 2, Footnote 1).

$\log(\Delta f/f)$ made it possible to directly compare performance between standard AM rates, and indicated a significant effect of standard AM rate [$F_{(1,842)} = 52.926$, $p < 0.0001$] and interaction between $\log(\Delta f/f)$ and standard AM rate [$F_{(1,842)} = 15.966$, $p < 0.0001$], where the 90 Hz standard AM rate had a higher average d'. The addition of a pitch cue may account for enhanced performance at 90 Hz. The interaction between $\log(\Delta f/f)$ and standard AM rate suggests that the slope of the psychometric function differed between 10 and 90 Hz. However, the interaction between standard AM rate and AM depth was not significant [$F_{(1,842)} = 0.006$, $p = 0.940$], implying that the ability for listeners to compare AM rate between places-of-stimulation may be impaired when AM depth is reduced in one place-of-stimulation regardless of standard AM rate.

3. Effects of AM pairing configuration

Pairs of SAM tones were presented in three different configurations, representing the three primary ways that envelope fluctuations could be compared in real-world listening (outlined in Fig. 2). It was predicted that the same place, across ears condition might result in better performance overall because of the addition of a binaural beat cue (Mcfadden and Pasanen, 1975). There was not sufficient statistical evidence to conclude that AM pairing configuration affected d'

$[F_{(2,842)} = 2.076$, $p = 0.127]$. However, there was a significant interaction between standard AM rate and AM pairing configuration [$F_{(2,842)} = 3.901$, $p < 0.05$]. This suggests that the ability to compare envelope fluctuations between places-of-stimulation may differ depending on which places-of-stimulation are compared for lower frequency envelope cues where pitch cues are not available. For 10 Hz, d' was significantly higher for: different place, within ears compared to the different place, across ears [$t(842) = 2.863$, $p < 0.05$], same place, across ears compared to the different place, across ears [$t(842) = 2.871$, $p < 0.05$], but not the different place, within ears compared to same place, across ears [$t(842) = 0.007$, $p = 1.000$] configurations. For 90 Hz, d' no significant difference was observed between any AM pairing configurations: within ears, different place compared to the different place, across ears [$t(842) = -0.894$, $p = 0.948$], same place, across ears compared to the different place, across ears [$t(842) = -0.064$, $p = 1.000$], and the different place, within ears compared to same place, across ears [$t(842) = -0.830$, $p = 0.962$] configurations. *Post hoc* tests were completed using estimated marginal means with Tukey adjustment for multiple comparisons.

The 95% confidence intervals for change in threshold between the 20%:50% and 50%:50% AM depth conditions were created for each AM pair configuration and standard AM rate. Confidence intervals are displayed in Fig. 6 and
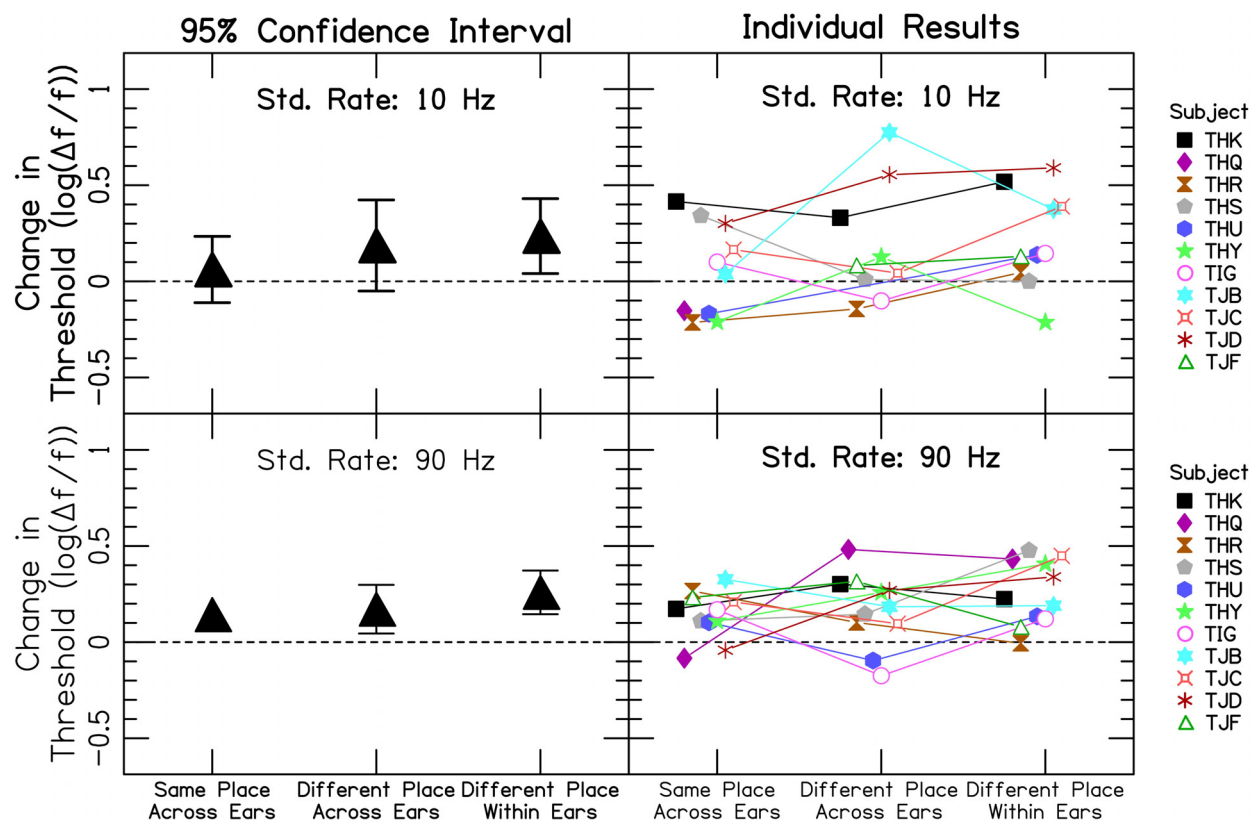

FIG. 6. (Color online) 95% confidence intervals and individual results for change in threshold (0.707 proportion correct) between the 20%:50% and 50%:50% AM depth conditions. The top and bottom rows correspond to standard AM rates of 10 and 90 Hz, respectively. The left and right columns correspond to 95% confidence intervals and individual results, respectively. The x axis corresponds to the AM pairing configuration (see Fig. 2). The y axis corresponds to the change in threshold between the 20%:50% and 50%:50% AM depth conditions. Values above zero indicated that listeners worsened on the task when AM depth was reduced from 50% to 20% in one place-of-stimulation. Confidence intervals were computed assuming a t distribution. Values above zero imply a significant increase in threshold from the 50%:50% to 20%:50% AM depth conditions. Thresholds for listeners were excluded if they fell above or below the values for $\log(\Delta f/f)$ tested in the experiment. Listener codes and corresponding symbols are given on the far right.

J. Acoust. Soc. Am. **146** (2), August 2019

Anderson *et al.*   1199

suggest a change in the effect of AM depth with AM rate and AM pairing configuration. Results from confidence intervals differ slightly from the results of the ANOVA. Though there was not a statistically significant interaction between reduced modulation depth and AM pairing configuration [$F(2,842) = 2.620$, $p = 0.073$], it can be seen from Fig. 6 that a change in threshold between the 20%:50% and 50%:50% AM depth conditions was only observed for the different place, within ears condition for the standard AM rate of 10 Hz. With the standard AM rate of 90 Hz, there was a positive difference between the 20%:50% AM depth conditions in all AM pair configurations. Note that the ANOVA was completed using the entire psychometric function for each subject, while confidence intervals only show the change in threshold between 20%:50% and 50%:50% AM depth.

D. Discussion

Results from this experiment demonstrated that when comparing the temporal envelope between two places-of-stimulation, if the AM depth was reduced in one place-of-stimulation, then sensitivity to differences in rate of AM decreased. This finding held for both standard AM rates of 10 and 90 Hz. Moreover, there was an interaction between AM pairing configuration and standard AM rate. For the 10 Hz standard AM rate, sensitivity was greatest when comparing temporal envelope fluctuations at a different place, across ears. For the 90 Hz standard AM rate, sensitivity was similar across all AM pairing configurations.

Most listeners reported that Experiment 2 was exceptionally difficult. Thus, some preventative stimulus manipulations were foregone to prevent distraction or increased difficulty. It is possible that listeners could have relied on changes in loudness to discriminate between different AM rates. Zhang and Zeng (1997) evaluated the effects of AM rate on loudness perception of SAM noise in NH listeners. Their results suggest that loudness may have changed for the AM rates employed in the present study. The best way to account for this would have been to rove the level in each cochlear place-of-stimulation independently. Instead, level was fixed within each block. However, recall that in Experiment 2 the AM rate was roved between trials and standard rates of 10 and 90 Hz were interleaved for all listeners except THK. If listeners experienced changes in loudness associated with AM rate between places-of-stimulation, the average loudness and difference in loudness should have changed on each trial. Further, the root-mean square level between different AM depths and AM rates was fixed. Finally, when asked to describe how they completed the task, most listeners indicated that they formed categories by AM rate, with the lowest rates corresponding to noticeable changes in loudness over time that could be compared between cochlear places-of-stimulation, and the highest AM rates corresponding to changes in pitch or roughness.

To address whether loudness could have confounded the results from Experiment 2, the loudness for individual SAM tones was estimated across AM rate using the loudness model from Moore *et al.* (2016). Results from the model are
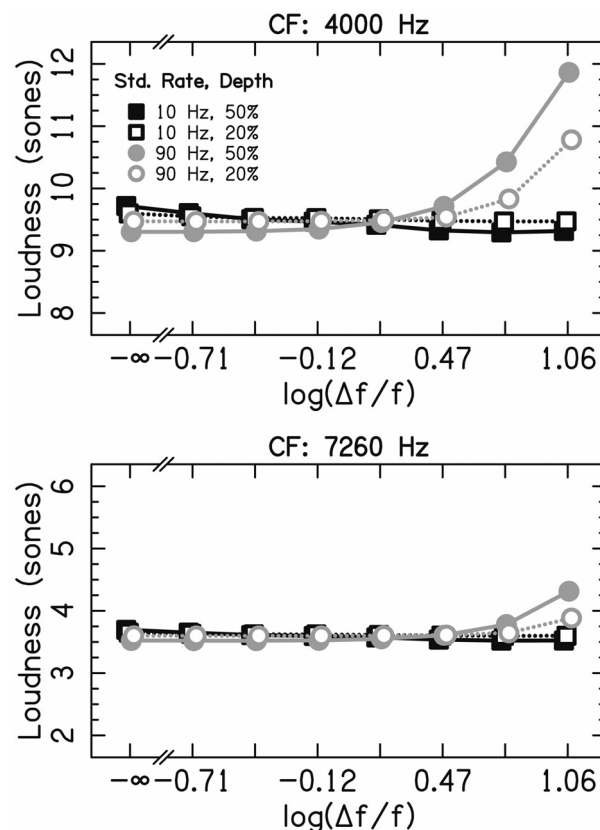

FIG. 7. Estimated loudness by AM rate. Loudness was estimated using the model from Moore *et al.* (2016). Results are plotted as in Fig. 4 for comparison, but each point represents a single AM rate (for specific AM rates, see Table II). The *y* axis corresponds to loudness in sones. The top and bottom panels correspond to SAM tones with carrier frequencies of 4000 and 7260 Hz, respectively. Please note the difference in scale for each panel (7260 Hz resulted in much less loudness overall). Different scales were used to make the change in loudness across AM rate visually apparent.

plotted in Fig. 7. Stimuli with a 4000 Hz carrier resulted in substantially greater estimated loudness than 7260 Hz. Estimated loudness was relatively consistent for most AM rates at 20% or 50% AM depth for both center frequencies. Estimated loudness decreased slightly as AM rates were increased from 10 to 100 Hz and increased substantially for the highest AM rates (mean of 610.47 and 1116.41 Hz) as in Zhang and Zeng (1997), especially for 50% AM depths. Critically, in the results of Zhang and Zeng (1997) AM rate had different effects on loudness at different modulation depths for SAM noise. Figure 7 indicates that estimated loudness was relatively consistent across rate for 20% and 50% AM depth, suggesting that loudness remained relatively consistent across AM rate for 50% and 20% AM depth for the five smallest values of $\log(\Delta f/f)$.

To account for results that may have been driven by loudness differences, the results from three additional ANOVAs are included in Table III. For one ANOVA, the analysis was conducted with the two largest values of $\log(\Delta f/f)$ excluded to account for cases where loudness might have provided a useful cue. Since loudness differences did not occur for the largest values of $\log(\Delta f/f)$ for the standard AM rate of 10 Hz, two additional ANOVAs were computed for standard AM rates of 10 Hz and 90 Hz separately to evaluate the effects of $\log(\Delta f/f)$, AM depth, and AM pairing configuration. For the ANOVA

TABLE III. ANOVAs summarized by main effects and interactions. The second and fourth columns (labeled 10 and 90 Hz, and 90 Hz only) were computed excluding data for the two highest values of log($\Delta f/f$).

Effect	Original ANOVA		10 and 90 Hz		10 Hz only		90 Hz only	
	F-Statistic	p	F-Statistic	p	F-Statistic	p	F-Statistic	p
log($\Delta f/f$)	$F(6,842) = 402.711$	<0.0001	$F(4,596) = 241.261$	<0.0001	$F(6,422) = 152.927$	<0.0001	$F(4,298) = 199.959$	<0.0001
AM Depth	$F(1,842) = 81.920$	<0.0001	$F(1,598) = 32.900$	<0.0001	$F(1,422) = 45.961$	<0.0001	$F(1,298) = 37.688$	<0.0001
Std. Rate	$F(1,842) = 52.926$	<0.0001	$F(1,598) = 1.766$	0.184	–	–	–	–
Pairing Configuration	$F(2,842) = 2.076$	0.127	$F(2,598) = 2.224$	0.109	$F(2,422) = 6.252$	<0.01	$F(2,298) = 0.194$	0.824
Center Frequency	$F(1,9) = 0.319$	0.586	$F(1,9) = 0.403$	0.541	$F(1,9) = 0.064$	0.806	$F(1,9) = 0.088$	0.773
log($\Delta f/f$) × AM Depth	$F(1,842) = 4.569$	<0.001	$F(4,598) = 3.653$	<0.01	$F(6,422) = 5.429$	<0.0001	$F(4,298) = 3.154$	<0.05
log($\Delta f/f$) × Std. Rate	$F(1,842) = 15.966$	<0.0001	$F(4,598) = 3.653$	<0.01	–	–	–	–
log($\Delta f/f$) × Pairing Configuration	$F(12,842) = 0.829$	0.621	$F(8,598) = 0.963$	0.464	$F(12,422) = 1.009$	0.439	$F(8,298) = 1.319$	0.233
AM Depth × Std. Rate	$F(1,842) = 0.006$	0.940	$F(1,598) = 3.223$	0.073	–	–	–	–
AM Depth × Pairing Configuration	$F(2,842) = 2.620$	0.073	$F(2,598) = 1.145$	0.319	$F(2,422) = 4.324$	<0.05	$F(2,298) = 0.017$	0.983
Std. Rate × Pairing Configuration	$F(2,842) = 3.901$	<0.05	$F(2,598) = 1.954$	0.143	–	–	–	–

corresponding to 90 Hz standard AM rate alone, the two highest values of log($\Delta f/f$) were excluded from the analysis. Together, the new analyses attempting to account for estimated loudness changes due to AM rate suggest that loudness differences in the largest values of log($\Delta f/f$) for standard AM rate of 90 Hz may have driven greater sensitivity compared to the standard AM rate of 10 Hz. Additionally, the second set of ANOVAs confirm that for the standard AM rate of 90 Hz, differences in sensitivity between AM pairing configurations were driven in part by the largest two values of log($\Delta f/f$).

It is also possible that loudness over time of stimulus presentation differed according to log($\Delta f/f$). This reflects one possible cue that listeners could have used to perform the experiment. At very low rates, listeners perceived fluctuating loudness over time for a single SAM tone, so when multiple, low-rate SAM tones are presented simultaneously, it seems likely that the perceived loudness will vary with the relationship between the phase of each SAM tone. Additional experiments would be required to determine the role of loudness changes over stimulus presentation in this task. Specifically, loudness changes at low AM rates due to loudness summation from two components depending upon phase could also apply to a previous experiment where SAM tones were presented to different places-of-stimulation and AM phase was manipulated but AM rate was kept constant (Strickland et al., 1989). In the present experiment, AM rate roving and randomization of phase would have made it difficult to make decisions on each trial according to the loudness summation of both SAM tones.

Removing the two highest values of log($\Delta f/f$) would have kept the maximum AM rate in Experiment 2 at 471.90 Hz (see Table II). Experiments by Kohlrausch et al. (2002) imply that these sidebands are not useful in detecting AM rate in a task like that used in Experiment 1, so it is possible that sidebands were not very useful in this experiment.

Removing the two highest values of log($\Delta f/f$) would also decrease the utility of combination tones. Since combination tones depend upon the relationship between the AM rate and carrier, the detectability of these tones may change with AM rate. Thus, one possible explanation for a difference in slope between the psychometric function for 10 Hz

compared to 90 Hz standard AM rates is an increased access to combination tones at larger values of log($\Delta f/f$) for 90 Hz. We do not feel that there is sufficient evidence concerning which combination tones could contribute most greatly affect discrimination of AM rate, but combination tones would be most audible at higher rates. Therefore, preservation of effects after excluding the highest rates tested in Experiment 2 provides some evidence to suggest that the effects of AM depth at 90 Hz were not due to combination tones. Experiments with patients that use CIs could address this in more detail since CIs directly stimulate the auditory nerve and do not rely on the cochlear filtering that produces combination tones.

The goal for the task in Experiment 2 was to require listeners to compare pairs of AM rates presented simultaneously between cochlear places-of-stimulation. For the different place, within ears AM pairing configuration in Experiment 2, it is possible that NH listeners could have used an overlapping region on the basilar membrane to make decisions on each trial (Kreft et al., 2013). The best way to account for this overlap would have been to use a low-level masking noise. During pilot testing, several listeners reported the inclusion of masking noise as being distracting, so no masking noise was used. Moreover, the carrier frequencies used in Experiment 2 were quite disparate (4000 or 7260 Hz). Performance in the different place, within ears condition was worse than other the different place, across ear conditions for 10 Hz (Fig. 5). Thus, the monaural overlap in excitation on the basilar membrane did not appear to provide any substantial advantage over the different place, across ears AM pairing configuration. In the same place, across ears condition, it is possible that listeners could use sidebands which change with AM rate to complete the task. A listener could detect a non-zero interaural level difference using the sidebands in either ear. Similarly, a masking noise would have been needed to prevent the use of sidebands but was not included.

Results were considerably variable across listeners (for example, see Fig. 3; for raw data, see Supplementary Fig. 2[1]). Variability in naive NH listeners has been documented on a variety of psychophysical tasks (Johnson et al., 1986; Kidd

et al., 2007; Lutfi and Liu, 2007), though not with this specific psychophysical paradigm or that used in Experiment 1 (Füllgrabe and Lorenzi, 2003; Grant *et al.*, 1998; Lee, 1994; Lemańaska *et al.*, 2002). It is possible that the task difficulty contributed to variability across listeners. Fortunately, the design and statistical analyses implemented in this study account for differences across listeners. This result suggests that a within-subject design will be important to consider if this experiment is implemented in listeners with hearing loss. Further, listeners with hearing loss would be expected to demonstrate greater variability and could be compared against NH listeners.

It was predicted that if one AM pairing configuration resulted in the best performance, it would be the same place, across ears condition because of the addition of a binaural beat cue at low values of $\log(\Delta f/f)$. Binaural beat cues from the envelope result when small differences in AM rate exist between each ear and the same center frequency is used (Mcfadden and Pasanen, 1975). Thus, binaural beat cues could provide an additional cue in the same place, across ears AM pairing configuration beyond those available for the other AM pairing configurations. Surprisingly, the different place, across ears AM pairing configuration led to better performance than the same place, across ears AM pairing configuration for the 10 Hz standard AM rate. Note that sensitivity to interaural timing differences is worse for 32 Hz (lowest rate tested) compared to 128 Hz with SAM tones (Bernstein and Trahiotis, 2002). It may be that a binaural beat cue is not useful for very low standard AM rates, or that it simply was not sufficient to result in better performance. Note that the magnitude of the effect of AM depth was smaller for the same place, across ears compared to the different place, within ears AM pairing configurations (Fig. 6 and Table III). The overall better performance of 90 Hz was most likely due to the addition of a pitch cue or sharper slope of the envelope for 90 Hz compared to 10 Hz (Bernstein and Trahiotis, 2002; Bernstein and Trahiotis, 2009; Dietz *et al.*, 2016; Laback *et al.*, 2011). This result also supports the notion that listeners did not use sidebands to complete the task, since sidebands would have been most useful in the same place, across ears AM pairing configuration.

The paradigm in Experiment 2 was similar to decorrelation detection experiments (Bernstein and Trahiotis, 1992; Goupell and Litovsky, 2015; Richards, 1987), except that the differences in envelope across places-of-stimulation were deterministic, resulting in changes in the quality of sounds with increasing rate. It is possible that the task could have been completed using the amount of decorrelation between each place of stimulation, which could have changed with AM rate. To assess this possibility, the cross correlation was computed as in previous studies (Goupell and Litovsky, 2015) and is reported in Supplementary Fig. 3.[1] The only changes in correlation occurred at the standard AM rate and changed systematically with envelope phase. The cross correlation is relatively constant across all other values of $\log(\Delta f/f)$. Since the proportion of "different" AM rates in this experiment was similar when the standard AM rate was presented to each place-of-stimulation and when small values of $\log(\Delta f/f)$ were used (Fig. 3 and Supplementary Fig. 2[1]), it seems that

decorrelation did not provide a useful cue for listeners to complete the task. It should be noted, however, that the magnitude of change in cross-correlation with envelope phase decreased when the AM depth was 20%:50%. Thus, differences in phase across places-of-stimulation could be less useful if the AM depth is reduced in one place-of-stimulation (e.g., interaural timing differences in the envelope).

Previous research evaluating thresholds for discriminating between interaural time differences in the envelope changed with frequency (Bernstein and Trahiotis, 1994). Thus, it was important to determine whether the center frequency of the ear that received the AM depth manipulation affected results. With respect to Experiment 2, the same place, across ears AM pairing configuration might then be expected to see a change in sensitivity according to carrier frequency. In this case, no effect of carrier frequency was found, suggesting that results could be interpreted similarly regardless of carrier frequency. The difference in our results could be related to the fact that listeners in Bernstein and Trahiotis (1994) were older than the undergraduate students that participated in this experiment.

IV. GENERAL DISCUSSION

In everyday auditory environments, listeners with hearing impairment struggle in separating speech from noise. It is not obvious which factors contribute most to poor performance listening in noise, and it may be that factors vary from one individual to another. One particular factor that is thought to vary within and across individuals is the fidelity of temporal envelope encoding in the auditory nerve in either ear. Even within the same listener, there may be places-of-stimulation that have poor to excellent temporal envelope encoding. Specifically, periodic fluctuations in the temporal envelope can be used to form auditory streams when presented in a sequence (Grimault *et al.*, 2002) and differences between the temporal envelopes are detectable within-(Richards, 1987) and across-ears (Goupell and Litovsky, 2015; Whitmer *et al.*, 2014) when presented simultaneously. While many factors can contribute to poorer temporal envelope encoding, at least some of these factors affect virtually all listeners with hearing-impairment.

This study investigated a NH simulation of poor temporal encoding at specific places-of-stimulation. The depth of AM was reduced for stimuli to represent poorer temporal envelope encoding. Experiment 1 demonstrated that performance discriminating between intervals with different AM rates worsened when AM depth was reduced from 50% to 20%. In Experiment 2, a new psychophysical task where listeners compared rates of AM across ears and places-of-stimulation was used (see Fig. 2). Results indicated that reducing AM depth from 50% to 20% worsened sensitivity (Fig. 6). Together, these results suggest that the ability to compare information in the envelope across places-of-stimulation might be impaired when the temporal encoding is poor for one of the places-of-stimulation.

The critical part of this conclusion is that when only one of the places-of-stimulation encodes temporal information poorly, the auditory system cannot make accurate

comparisons of temporal envelope fluctuations. From the review by Grose *et al.* (2005), it is clear that there are many examples where the temporal envelope can be used to discriminate between sound sources. This experiment focused specifically on the case of short duration, periodic sounds, with the unique advantage of comparing performance within and across the ears. This is the first study to systematically compare performance discriminating between AM rate within and across the ears and cochlear place. Importantly, while no effect of AM pairing configuration was shown here on average, sensitivity changed depending upon the AM pair configuration for low standard AM rates (10 Hz).

In the present study, listeners compared one place-of-stimulation against a 10 or 90 Hz standard AM rate, reflecting two speech-relevant perceptual processes: slower fluctuations that result in, e.g., word segmentation cues and faster fluctuations that result in pitch cues (Rosen, 1992). Results from both experiments showed an effect of standard AM rate on the ability to judge differences in AM rate when presented sequentially (Experiment 1) and simultaneously presented across places-of-stimulation (Experiment 2), with listeners being more sensitive to changes at 90 Hz relative to 10 Hz.

A. Relation to listeners with hearing impairment

Listeners with hearing impairment can have highly-varying temporal representations in the auditory nerve. For example, for listeners that receive a CI and have either NH or use a hearing aid in the other ear, electric and acoustic information must be integrated across the ears to distinguish between different sound sources. Similarly, for listeners that receive a hybrid CI, highly phase-locked, electrically encoded information must be compared against acoustically encoded information within the same ear. For CI users, greater distance between the electrode array and auditory nerve has been related to increases in threshold, and likely results in poorer spectro-temporal representations (Bierer, 2010). Similarly, long durations of deafness are associated with poorer phase locking and loss of auditory nerve fibers, as well as deterioration of dendrites (Leake and Hradek, 1988; Nadol, 1997; Shepherd *et al.*, 2004). Thus, it is apparent that temporal representations can vary highly within the same individual; yet little research has systematically focused on the implications of differing temporal representations.

The present study suggests that the ability to make spectro-temporal comparisons across pairs of cochlear locations worsens when temporal encoding is poor at one place-of-stimulation. This result implies that comparisons of temporal information are limited by the worst temporal representation. Results from Experiment 2 suggest that the ability to compare temporal information within the same ear may be more heavily impacted by poor temporal encoding than comparisons across ears at low rates of modulation (see Fig. 6). This may be due to the addition of a binaural beat cue (Mcfadden and Pasanen, 1975) in across-ear conditions, and requires further investigation in listeners with hearing impairment. It is important to note that simply because the change in threshold between AM depth conditions was smaller in the across-ear compared to the within-ear condition, this does not suggest that the functional implications are less severe. That is, the binaural system relies on precise timing information to detect the location of sounds.

Recent work in listeners with bilateral CIs suggests that sensitivity to binaural cues may be predicted by the ear with worse sensitivity to temporal information (Ihlefeld *et al.*, 2015). These results are similar in spirit to studies in listeners with NH, CIs, or single-sided deafness showing that as the envelope attack slope increases, which should result in highly-synchronous firing of the auditory nerve, listeners become increasingly sensitive to interaural timing differences in the envelope (Bernstein and Trahiotis, 2002, 2009; Dietz *et al.*, 2016; Laback *et al.*, 2011). None of the studies in NH have investigated what occurs when temporal envelope encoding is asymmetric across the ears, however. Thus, it may be important to more thoroughly investigate changes in binaural sensitivity under asymmetric temporal envelope encoding.

Some previous studies have demonstrated improved performance of individuals with hearing loss compared to NH in tasks involving AM detection using low AM rates (e.g., Schlittenlacher and Moore, 2016). Additionally, some experiments have suggested that hearing loss might improve temporal resolution for stimuli with AM (e.g., Henry *et al.*, 2014). It has been suggested that differences in performance are due to loudness recruitment and a lack of compression associated with hearing loss (Jennings *et al.*, 2018; Schlittenlacher and Moore, 2016). Some studies show an improvement in AM detection for listeners with hearing-impairment compared to NH when the same sound pressure levels (not sensation levels) are used (Jennings *et al.*, 2018), while others show no difference (Schlittenlacher and Moore, 2016). Regardless, these results suggest that the representation of AM rate may be asymmetric between different places-of-stimulation for individuals with hearing impairment.

Before this task or similar tasks are implemented in listeners with hearing loss in the future, investigators should consider several issues. Listeners with NH in this study exhibited considerable variability across individuals (see Fig. 3 for one example and Supplementary Figs. 1 and 2[1] for raw data). This suggests that an across-subjects experimental design using similar tasks to compare performance between groups would need a very large sample size to attain the necessary statistical power to detect effects. A more efficient approach may be to use a within-subjects experimental design and compare temporal envelope processing abilities across a variable within the same person as has been completed in several studies in individuals with hearing loss (e.g., Garadat *et al.*, 2013; Ihlefeld *et al.*, 2015; Landsberger, 2008; Zhou and Pfingst, 2012). In this study, investigators spent considerable time training listeners and discussing listeners' perception of changes in AM rate. This seemed to be an especially effective approach as it helped listeners understand the task and identify the perceptual changes they experienced as AM rate changed (e.g., rhythm, roughness, timbre, pitch). Experimenters were careful not to tell listeners what they "should" hear as AM rate changed. Finally, experimenters documented listeners' descriptions of changes in stimuli for specific ranges of AM rates. This was helpful

if listeners became frustrated or began a new session on a different day.

B. Implications of asymmetric temporal representations

One example of how poorer signal encoding can interfere with speech in noise understanding is contralateral interference, where information from the poorer ear interferes with accessing information in the better ear (Bernstein *et al.*, 2017; Gallun *et al.*, 2007; Goupell *et al.*, 2016; Goupell *et al.*, 2018). This has been observed in NH listeners (Gallun *et al.*, 2007), but the implications for listeners with hearing impairment are not immediately clear. Within the bilateral CI population, it appears that longer durations of deafness might be related to poorer ability to ignore information in the worse ear (Goupell *et al.*, 2016; Goupell *et al.*, 2018). Interference has also been demonstrated in subsets of patients with single-sided deafness (Bernstein *et al.*, 2017).

The present experiments provide evidence to suggest that the ability to separate auditory objects will be negatively impacted by asymmetric temporal encoding. As this report demonstrates, the ability to make spectro-temporal comparisons across pairs of cochlear places-of-stimulation likely worsens when the amount of phase locking decreases for one place-of-stimulation. Poorer spectro-temporal comparisons have downstream implications for the model presented by Shinn-Cunningham (2008), making auditory objects less salient and therefore less able to compete for attention. Listeners with hearing impairment exhibit extraordinary heterogeneity with respect to temporal encoding and performance on speech reception tasks in noise. More research is required to understand the implications of asymmetries in temporal encoding on patient outcomes to potentially improve patient care.

Previous studies have demonstrated that turning off electrodes where patients with CIs are insensitive to temporal cues can improve speech in noise understanding (Garadat *et al.*, 2013; Zhou and Pfingst, 2012). This paper provides one example of a sound source segregation mechanism that might be improved when electrodes for places-of-stimulation where the patient is not very sensitive to temporal fluctuations are turned off in a patient's programming.

C. Normal-hearing simulation of asymmetric temporal encoding

The experiments presented in this manuscript altered the representation of AM rate in the auditory nerve by reducing the depth of AM. The goal of this manipulation was to reduce the dynamic range of the stimulus and degree of phase locking for auditory nerve fibers. Changes in loudness as AM rate was varied could have been useful in performance. To address this potential confound, Fig. 7 shows estimated loudness due to changes in AM rate. The AM rates resulting in the largest differences in estimated loudness were not tested in Experiment 1 and additional ANOVAs are provided in the Discussion of Experiment 2 excluding these cases. The results imply that listeners' sensitivity to differences in AM rate decreased when AM depth was reduced for both Experiment 1 and 2, even after accounting for changes in loudness.

It is difficult to determine the role of combination tones in the current experiments. While the existence of combination tones has been verified psychophysically and physiologically, there is not a widely-accepted model to account for the magnitude of each combination tone. In the experiments by Lee (1994), the author suggests that that the difference tone (equal to the rate of AM) could contribute to AM rate discrimination in a paradigm similar to Experiment 1. If the difference tone is audible, then it would have presumably affected results in Experiment 2 as well. If participants in our experiments could use the difference tone to make judgments on differences in AM rate, then this would affect the higher AM rates used in this study ($>90\,$Hz). However, it is also possible that the difference tone did not contribute much to perceptual results, or that additional combination tones could be used to make AM rate distinctions. Ultimately, the role of combination tones in AM rate discrimination requires further investigation.

D. Summary and conclusions

In real listening environments, sound sources span across frequency and are present in both ears. Segregating sound sources requires ongoing spectro-temporal comparisons within and across the ears (Bregman, 1990; Grose *et al.*, 2005; Shinn-Cunningham, 2008). Experiment 2 explored the simplest case; where listeners indicated whether pairs of places-of-stimulation were the same or different, representing the temporal envelope with good temporal fidelity in one place and poorer temporal fidelity in the other place. The results from this manuscript suggest that the accuracy of these spectro-temporal comparisons may be determined by the places-of-stimulation with poorest temporal encoding. Asymmetries in temporal encoding provide one mechanism to explain difficulties separating sound sources when listening in complex auditory environments for listeners with hearing impairment.

[1]See supplementary material at https://doi.org/10.1121/1.5121423 for Supplementary Figs. 1–3.

Bernstein, J. G. W., Goupell, M. J., Schuchman, G. I., Rivera, A. L., and Brungart, D. S. (**2016**). "Having two ears facilitates the perceptual separation of concurrent talkers for bilateral and single-sided deaf cochlear implantees," Ear Hear. **37**(3), 289–302.
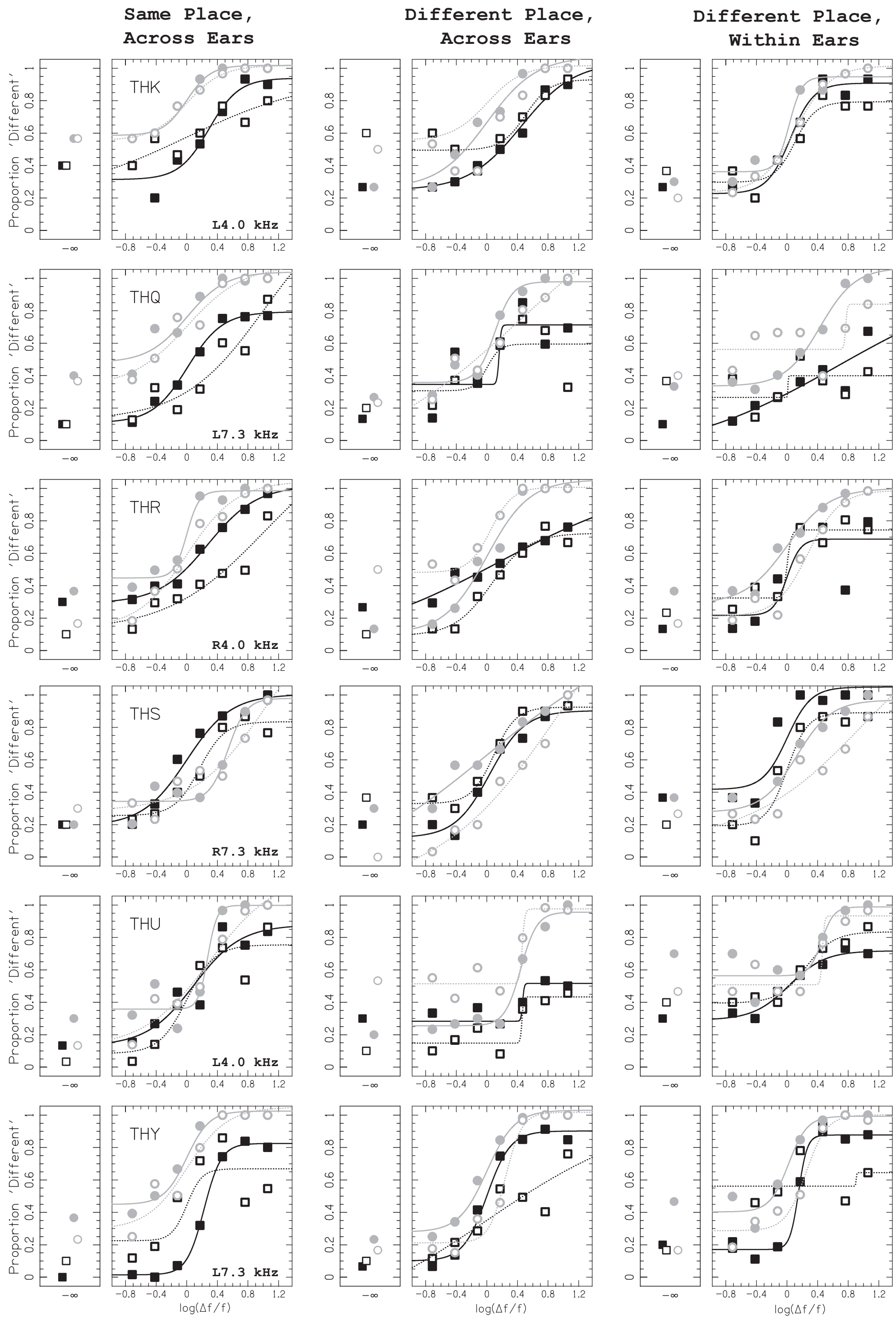
Bernstein, J. G. W., Goupell, M. J., Wess, J. M., Stakhovskaya, O. A., and Brungart, D. S. (**2017**). "Having two ears can facilitate or interfere with the perceptual separation of concurrent talkers for bilateral and single-sided deafness cochlear-implant listeners [Abstract]," J. Acoust. Soc. Am. **141**(5), 4031.

Bernstein, J. G. W., Iyer, N., and Brungart, D. S. (**2015**). "Release from informational masking in a monaural competing-speech task with vocoded copies of the maskers presented contralaterally," J. Acoust. Soc. Am. **137**(2), 702–713.

Bernstein, L. R., and Trahiotis, C. (**1992**). "Discrimination of interaural envelope correlation and its relation to binaural unmasking at high frequencies," J. Acoust. Soc. Am. **91**(1), 306–316.

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**(6), 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**(3), 1026–1036.

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**(5), 3234–3242.

Bierer, J. A. (**2010**). "Probing the electrode-neuron interface with focused cochlear implant stimulation," Trends Amplif. **14**(2), 84–95.

Bierer, J. A., and Nye, A. D. (**2014**). "Comparisons between detection threshold and loudness perception for individual cochlear implant channels," Ear Hear. **35**(6), 641–651.

Bregman, A. S. (**1990**). *Auditory Scene Analysis: The Perceptual Organization of Sound* (Bradford Books, MIT Press, Cambridge, MA).

Bregman, A. S., Abramson, J., Doehring, P., and Darwin, C. J. (**1985**). "Spectral integration based on common amplitude modulation," Percept. Psyhophys. **37**(5), 483–493.

Brungart, D. S. (**2001**). "Informational and energetic masking effects in the perception of two simultaneous talkers," J. Acoust. Soc. Am. **109**(3), 1101–1109.

Carlyon, R. P., and Deeks, J. M. (**2002**). "Limitations on rate discrimination," J. Acoust. Soc. Am. **112**(3), 1009–1025.

Carrell, T. D., and Opie, J. M. (**1992**). "The effect of amplitude comodulation on auditory object formation in sentence perception," Percept. Psyhophys. **52**(4), 437–445.

Chatterjee, M., Sarampalis, A., and Oba, S. I. (**2006**). "Auditory stream segregation with cochlear implants: A preliminary report," Hear. Res. **222**(1–2), 100–107.

Dietz, M., Wang, L., Greenberg, D., and McAlpine, D. (**2016**). "Sensitivity to interaural time differences conveyed in the stimulus envelope: Estimating inputs of binaural neurons through the temporal analysis of spike trains," J. Assoc. Res. Otolaryngol. **17**(4), 313–330.

Dollezal, L.-V., Beutelmann, R., and Klump, G. M. (**2012**). "Stream segregation in the perception of sinusoidally amplitude-modulated tones," PLoS ONE **7**(9), e43615.

Drullman, R., Festen, J. M., and Plomp, R. (**1994**). "Effect of temporal envelope smearing on speech reception," J. Acoust. Soc. Am. **95**(2), 1053–1064.

Eddins, D. A. (**1999**). "Amplitude-modulation detection at low- and high-audio frequencies," J. Acoust. Soc. Am. **105**(2), 829–837.

Festen, J. M., and Plomp, R. (**1990**). "Effects of fluctuating noise and interfering speech reception threshold for impaired and normal hearing," J. Acoust. Soc. Am. **88**(4), 1725–1736.

Füllgrabe, C., and Lorenzi, C. (**2003**). "The role of envelope beat cues in the detection and discrimination of second-order amplitude modulation (L)," J. Acoust. Soc. Am. **113**(1), 49–52.

Gallun, F. J., Mason, C. R., and Kidd, G. J. (**2007**). "The ability to listen with independent ears," J. Acoust. Soc. Am. **122**(5), 2814–2825.

Garadat, S. N., Zwolan, T. A., and Pfingst, B. E. (**2013**). "Using temporal modulation sensitivity to select stimulation sites for processor MAPs in cochlear implant listeners," Audiol. Neurootol. **18**(4), 247.

Gifford, R. H., Dorman, M. F., Skarzynski, H., Lorens, A., Polak, M., Driscoll, C. L. W., Roland, P., and Buchman, C. A. (**2013**). "Cochlear implantation with hearing preservation yields significant benefit for speech recognition in complex listening environments," Ear Hear. **34**(4), 413–425.

Goldwyn, J. H., Bierer, S. M., and Bierer, J. A. (**2010**). "Modeling the electrode-neuron interface of cochlear implants: Effects of neural survival,

electrode placement, and the partial tripolar configuration," Hear. Res. **268**(1–2), 93–104.

Goupell, M. J. (**2015**). "Interaural envelope correlation change discrimination in bilateral cochlear implantees: Effects of mismatch, centering, and onset of deafness," J. Acoust. Soc. Am. **137**(3), 1282–1297.

Goupell, M. J., Kan, A., and Litovsky, R. Y. (**2016**). "Spatial attention in bilateral cochlear-implant users," J. Acoust. Soc. Am. **140**(3), 1652–1662.

Goupell, M. J., and Litovsky, R. Y. (**2015**). "Sensitivity to interaural envelope correlation changes in bilateral cochlear-implant users," J. Acoust. Soc. Am. **137**(1), 335–349.

Goupell, M. J., Stakhovskaya, O. A., and Bernstein, J. G. W. (**2018**). "Contralateral interference caused by binaurally presented competing speech in adult bilateral cochlear-implant users," Ear Hear. **39**(1), 110–123.

Grant, K. W., Summers, V., and Leek, M. R. (**1998**). "Modulation rate detection and discrimination by normal-hearing and hearing-impaired listeners," J. Acoust. Soc. Am. **104**(2), 1051–1060.

Green, D. M., and Swets, J. A. (**1966**). *Signal Detection Theory and Psychophysics*, 1st ed. (Peninsula Publishing, Los Altos Hills, CA).

Grimault, N., Bacon, S. P., and Micheyl, C. (**2002**). "Auditory stream segregation on the basis of amplitude- modulation rate," J. Acoust. Soc. Am. **111**(3), 1340–1348.

Grose, J. H., Hall, J. W. I., and Buss, E. (**2005**). "Across-channel spectral processing," Int. Rev. Neurobiol. **70**, 87–119.

Henry, K. S., Kale, S., and Heinz, M. G. (**2014**). "Noise-induced hearing loss increases the temporal precision of complex envelope coding by auditory-nerve fibers," Front. Neuro. **8**(20), 1–10.

Hong, R. S., and Turner, C. W. (**2009**). "Sequential stream segregation using temporal periodicity cues in cochlear implant recipients," J. Acoust. Soc. Am. **126**(1), 291–299.

Ihlefeld, A., Carlyon, R. P., Kan, A., Churchill, T. H., and Litovsky, R. Y. (**2015**). "Limitations on monaural and binaural temporal processing in bilateral cochlear implant listeners," J. Assoc. Res. Otolaryngol. **16**(5), 641–652.

Jennings, S. G., Chen, J., Fultz, S. E., Ahlstrom, J. B., and Dubno, J. R. (**2018**). "Amplitude modulation detection with a short-duration carrier: Effects of a precursor and hearing loss," J. Acoust. Soc. Am. **143**(4), 2232–2243.

Johnson, D. M., Watson, C. S., and Jensen, J. K. (**1986**). "Individual differences in auditory capabilities, I," J. Acoust. Soc. Am. **81**(2), 427–438.

Joris, P. X., and Yin, T. C. T. (**1992**). "Responses to amplitude-modulated tones in the auditory nerve of the cat," J. Acoust. Soc. Am. **91**(1), 215–232.

Kidd, G. R., Watson, C. S., and Gygi, B. (**2007**). "Individual differences in auditory abilities," J. Acoust. Soc. Am. **122**(1), 418–435.

Kohlrausch, A., Fassel, R., and Dau, T. (**2002**). "The influence of carrier level and frequency on modulation and beat-detection thresholds for sinusoidal carriers," J. Acoust. Soc. Am. **108**(2), 723–734.

Kreft, H. A., Nelson, D. A., and Oxenham, A. J. (**2013**). "Modulation frequency discrimination with modulated and unmodulated interference in normal hearing and in cochlear-implant users," J. Assoc. Res. Otolaryngol. **14**(4), 591–601.

Krumbholz, K., Patterson, R. D., and Pressnitzer, D. (**2000**). "The lower limit of pitch as determined by rate discrimination," J. Acoust. Soc. Am. **108**(3), 1170–1180.

Laback, B., Zimmermann, I., Majdak, P., Baumgartner, W.-D., and Pok, S.-M. (**2011**). "Effects of envelope shape on interaural envelope delay sensitivity in acoustic and electric hearing," J. Acoust. Soc. Am. **130**(3), 1515–1529.

Landsberger, D. M. (**2008**). "Effects of modulation wave shape on modulation frequency discrimination with electrical hearing," J. Acoust. Soc. Am. **124**(2), EL21–EL27.

Leake, P. A., and Hradek, G. T. (**1988**). "Cochlear pathology of long term neomycin induced deafness in cats," Hear. Res. **33**(1), 11–33.

Lee, J. (**1994**). "Amplitude modulation rate discrimination with sinusoidal carriers," J. Acoust. Soc. Am. **96**(4), 2140–2147.

Lemańska, J., Sęk, A. P., and Skrodzka, E. B. (**2002**). "Discrimination of the amplitude modulation rate," Arch. Acoust. **27**(1), 3–21.

Litovsky, R. Y. (**1997**). "Developmental changes in the precedence effect: Estimates of minimum audible angle," J. Acoust. Soc. Am. **102**(3), 1739–1745.

Litvak, L. M., Delgutte, B., and Eddington, D. K. (**2003**). "Improved temporal coding of sinusoids in electric stimulation of the auditory nerve using desynchronizing pulse trains," J. Acoust. Soc. Am. **114**(4), 2079–2098.
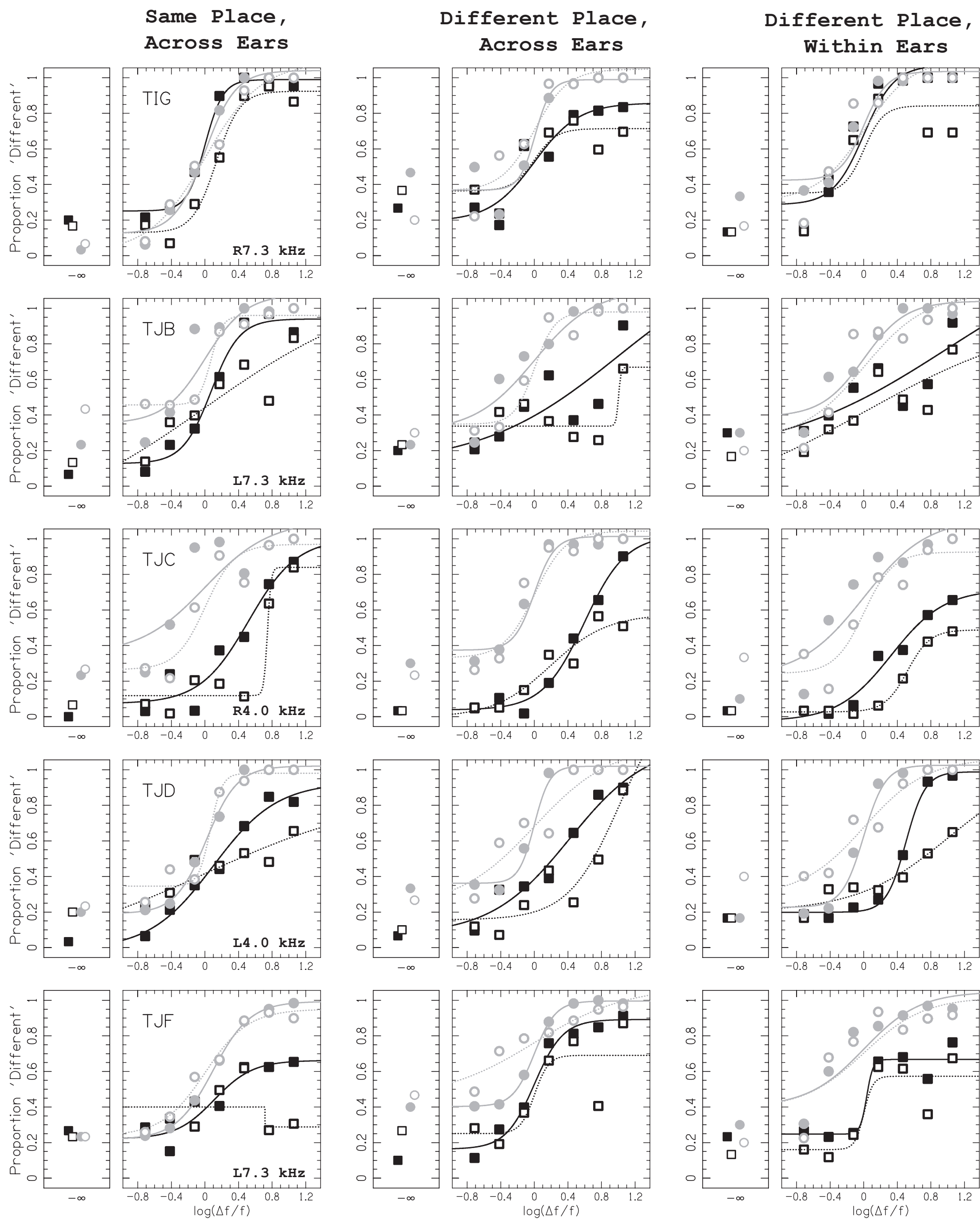
Loizou, P. C. (**2006**). "Speech processing in vocoder-centric cochlear implants," in *Cochlear and Brainstem Implants*, edited by R. Møller (Karger, Basel, Switzerland), pp. 109–143.

Loizou, P. C., Hu, Y., Litovsky, R., Yu, G., Peters, R., Lake, J., and Roland, P. (**2009**). "Speech recognition by bilateral cochlear implant users in a cocktail-party setting," J. Acoust. Soc. Am. **125**(1), 372–383.

Lutfi, R. A., and Liu, C. (**2007**). "Individual differences in source identification from synthesized impact sounds," J. Acoust. Soc. Am. **122**(2), 1017–1028.

Mcfadden, D., and Pasanen, E. G. (**1975**). "Binaural beats at high frequencies," Science **190**(4212), 394–396.

Moore, B. C. J., and Glasberg, B. R. (**1990**). "Derivation of auditory filter shapes from notched-noise data," Hear. Res. **47**, 103–138.

Moore, B. C. J., Glasberg, B. R., Varathanathan, A., and Schlittenlacher, J. (**2016**). "A loudness model for time-varying sounds incorporating binaural inhibition," Trend. Hear. **20**, 1–16.

Nadol, J. J. (**1997**). "Patterns of neural degeneration in the human cochlea and auditory nerve: Implications for cochlear implantation," Otolayngol. Head. Neck Surg. **117**(3), 220–228.

Nadol, J. J., Young, Y.-S., and Glynn, R. J. (**1989**). "Survival of spiral ganglion cells in profound sensorineural hearing loss: Implications for cochlear implantation," Ann. Otol. Rhinol. Laryngol. **98**(6), 411–416.

Nie, Y., and Nelson, P. B. (**2015**). "Auditory stream segregation using amplitude modulated bandpass noise," Front. Psychol. **6**(1151), 1–11.

Oxenham, A. J., and Bacon, S. P. (**2003**). "Cochlear compression: Perceptual measures and implications for normal and impaired hearing," Ear Hear. **24**(5), 352–366.

Reiss, L. A. J., Ito, R. A., Eggleston, J. L., and Wozny, D. R. (**2014**). "Abnormal binaural spectral integration in cochlear implant users," J. Assoc. Res. Otolaryngol. **15**, 235–248.

Reiss, L. A. J., Shayman, C. S., Walker, E. P., Bennett, K. O., Fowler, J. R., Hartling, C. L., Glickman, B., Lasarev, M. R., and Oh, Y. (**2017**). "Binaural pitch fusion: Comparison of normal-hearing and hearing-impaired listeners," J. Acoust. Soc. Am. **141**(3), 1909–1920.

Richards, V. M. (**1987**). "Monaural envelope correlation perception," J. Acoust. Soc. Am. **82**(5), 1621–1630.

Rosen, S. (**1992**). "Temporal information in speech: Acoustic, auditory and linguistic aspects," Philos. Trans. R. Soc. Lond. B **336**(1278), 367–373.

Rubinstein, J. T. Y., Wilson, B. S., Finley, C. C., and Abbas, P. J. (**1999**). "Pseudospontaneous activity: Stochastic independence of auditory nerve fibers with electrical stimulation," Hear. Res. **127**(1–2), 108–118.

Ruggero, M. A., Robles, L., Rich, N. C., and Recio, A. (**1992**). "Basilar membrane responses to two-tone and broadband stimuli," Philos. Trans. R. Soc. Lond. B **336**, 307–315.

Schlittenlacher, J., and Moore, B. C. J. (**2016**). "Discrimination of amplitude-modulation depth by subjects with normal and impaired hearing," J. Acoust. Soc. Am. **140**(5), 3487–3495.

Shannon, R. V., Zeng, F., Kamath, V., Wygonski, J., and Ekelid, M. (**1995**). "Speech recognition with primarily temporal cues," Science **270**(5234), 303–304.

Shepherd, R. K., and Hardie, N. A. (**2001**). "Deafness-induced changes in the auditory pathway: Implications for cochlear implants," Audiol. Neurootol. **6**(6), 305–318.

Shepherd, R. K., Roberts, L. A., and Paolini, A. G. (**2004**). "Long-term sensorineural hearing loss induces functional changes in the rat auditory nerve," Eur. J. Neurosci. **20**(11), 3131–3140.

Shinn-Cunningham, B. G. (**2008**). "Object-based auditory and visual attention," Trends Cogn. Sci. **12**(5), 182–186.

Spoendlin, H., and Schrott, A. (**1989**). "Analysis of the human auditory nerve," Hear. Res. **43**, 25–38.

Strickland, E. A., and Dhar, S. (**2000**). "An analysis of quasi-frequency-modulated noise and random-sideband noise as comparisons for amplitude-modulated noise," J. Acoust. Soc. Am. **108**(2), 735–742.

Strickland, E. A., and Viemeister, N. F. (**1997**). "The effects of frequency region and bandwidth on the temporal modulation transfer function," J. Acoust. Soc. Am. **102**(3), 1799–1810.

Strickland, E. A., Viemeister, N. F., Fantini, D. A., and Garrison, M. A. (**1989**). "Within- versus cross-channel mechanisms in detection of envelope phase disparity," J. Acoust. Soc. Am. **86**(6), 2160–2166.

Whitmer, W. M., Seeber, B. U., and Akeroyd, M. A. (**2014**). "The perception of apparent auditory source width in hearing-impaired adults," J. Acoust. Soc. Am. **135**(6), 3548–3559.

Wichmann, A. F., and Hill, N. J. (**2001**). "The psychometric function: I. Fitting, sampling, and goodness of fit," Percept. Psyhophys. **63**(8), 1293–1313.

Zhang, C., and Zeng, F. (**1997**). "Loudness of dynamic stimuli in acoustic and electric hearing," J. Acoust. Soc. Am. **102**(5), 2925–2934.

Zhou, R., Abbas, P. J., and Assouline, J. G. (**1995**). "Electrically evoked auditory brainstem response in peripherally myelin-deficient mice," Hear. Res. **88**(1–2), 98–106.

Zhou, N., and Pfingst, B. E. (**2012**). "Psychophysically based site selection coupled with dichotic stimulation improves speech recognition in noise with bilateral cochlear implants," J. Acoust. Soc. Am. **132**(2), 994–1008.

Zilany, M. S. A., Bruce, I. C., and Carney, L. H. (**2014**). "Updated parameters and expanded simulation options for a model of the auditory periphery," J. Acoust. Soc. Am. **135**(1), 283.

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**(5), 2390–2412.
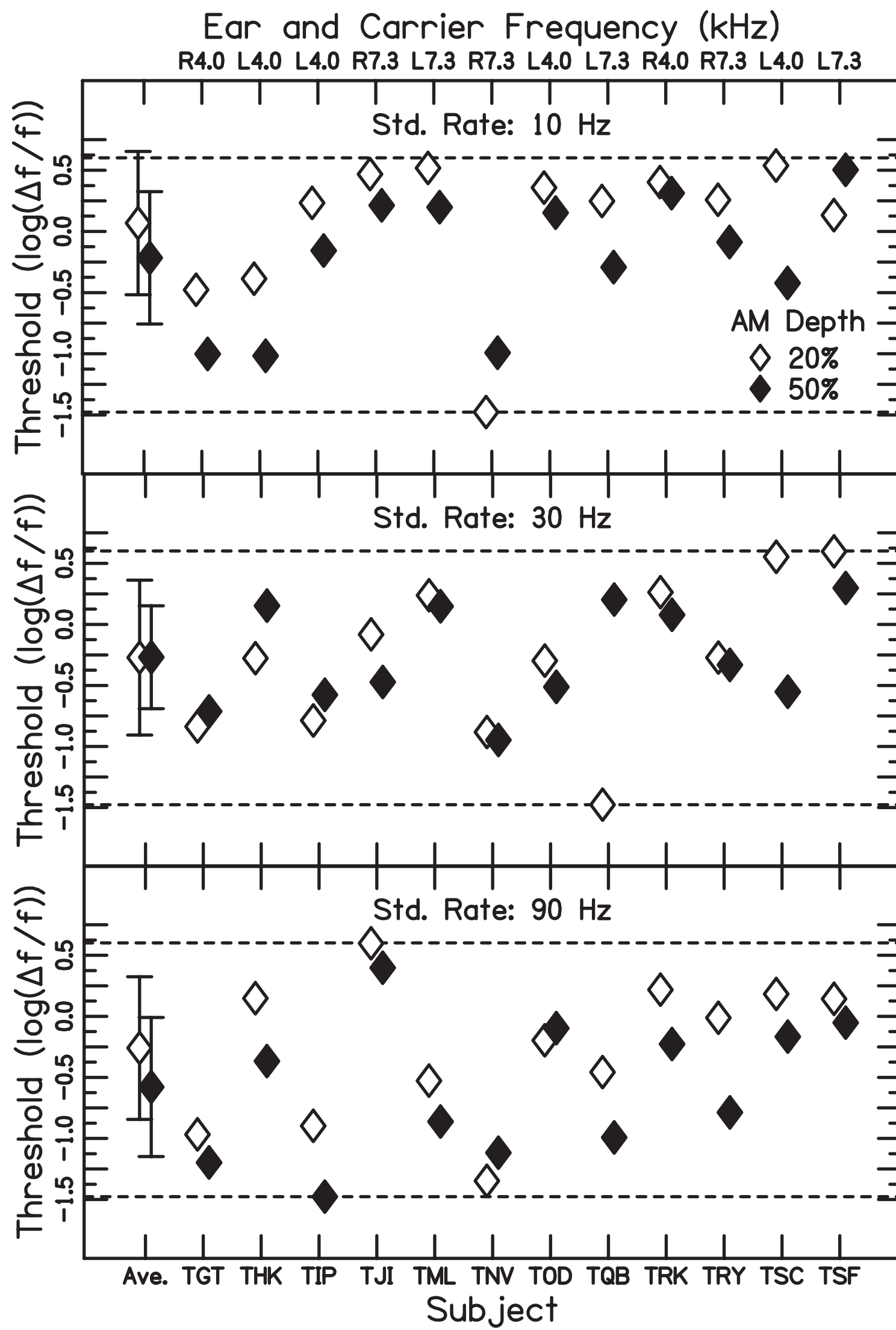
SUPPLEMENTARY FIG. 1. Raw thesholds from individual subjects in Experiment 1. Each row represents a differnt standard AM rate shown in the center of each panel.

The y-axis corresponds to threshold (70.7% correct) calculated using a weighted maximum likelihood procedure (psignfit). The x-axis corresponds to subject identifier.

A specific carrier frequency and ear was used for each subject, indicated on the x-axis at the top of the figure. The far left, abbreviated "Ave." represents the mean ± 1

standard deviation.

Same Place, Across Ears | Different Place, Across Ears | Different Place, Within Ears

THK L4.0 kHz
THQ L7.3 kHz
THR R4.0 kHz
THS R7.3 kHz
THU L4.0 kHz
THY L7.3 kHz

Proportion 'Different'

log(Δf/f)

SUPPLEMENTARY FIG. 2. Raw data from each subject, plotted as in Fig. 3. Each row corresponds to a subject, with their identifier displayed in the left-most column. Each column corresponds to an AM pairing configuration (see Fig. 2). The carrier frequency and ear for which the AM depth manipulation was made is given in the bottom right-hand corner of the panel furthest to the left for each listener.
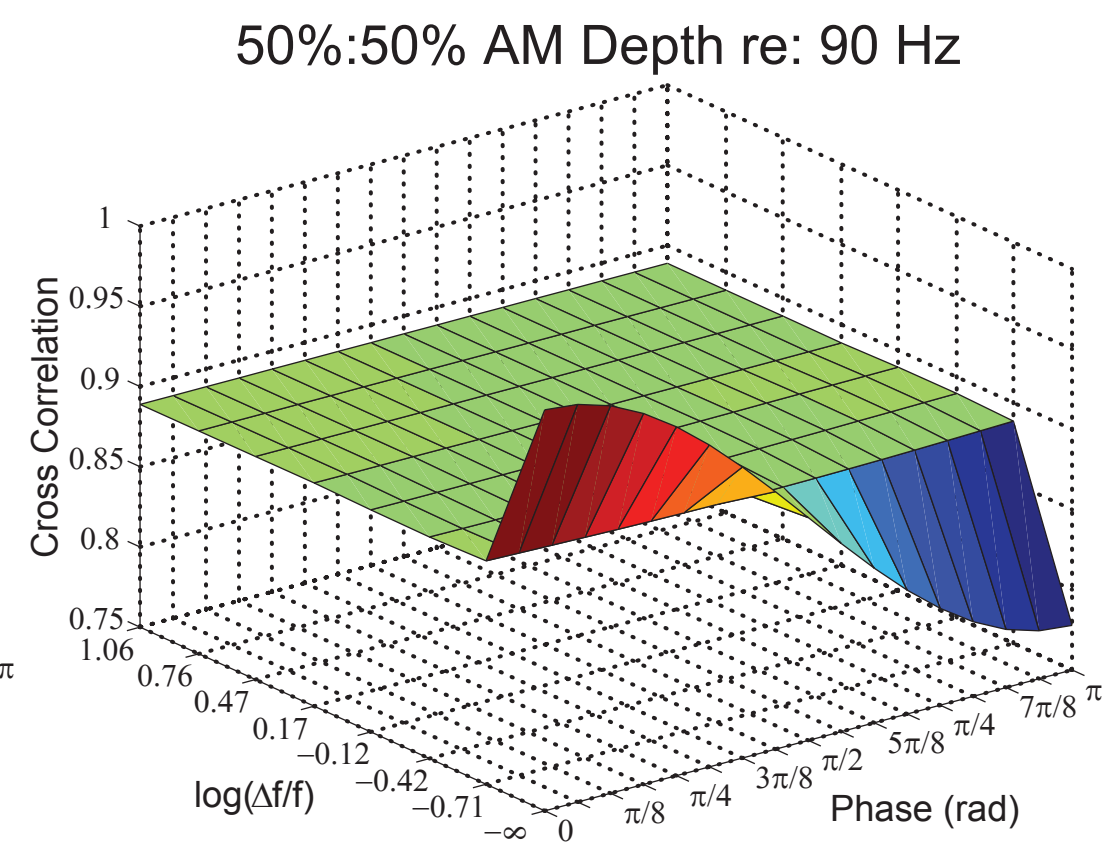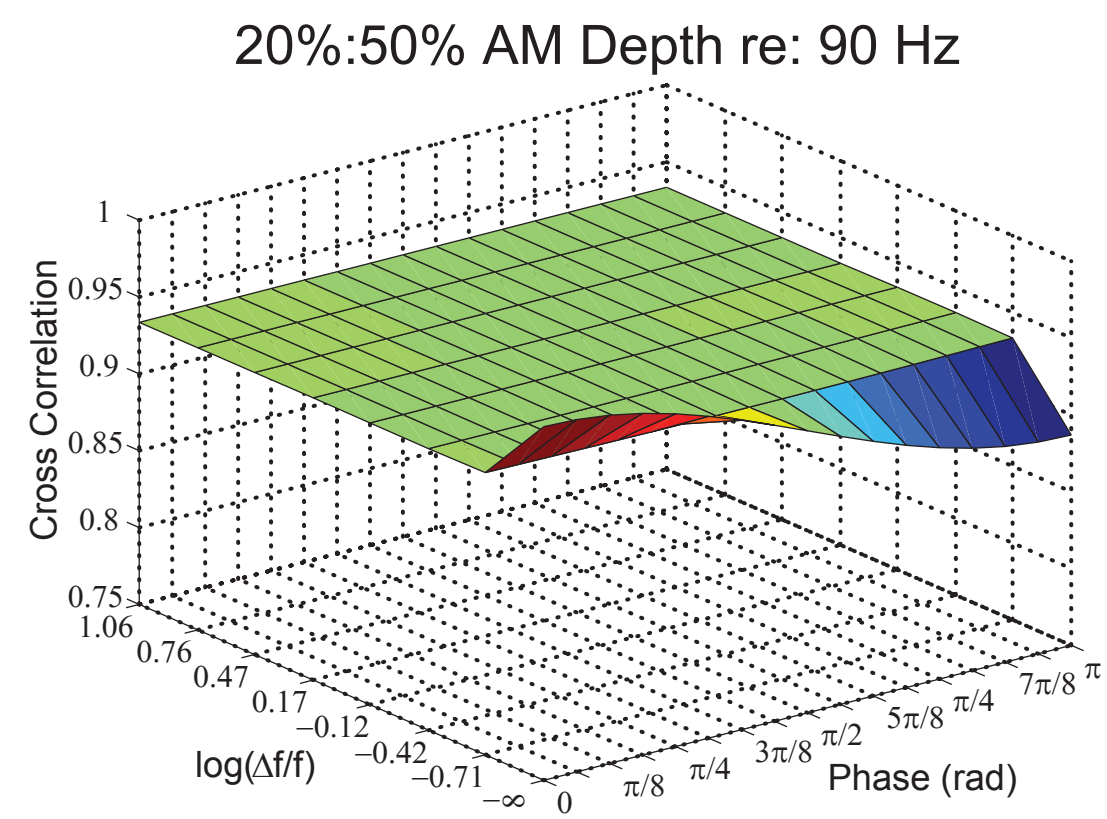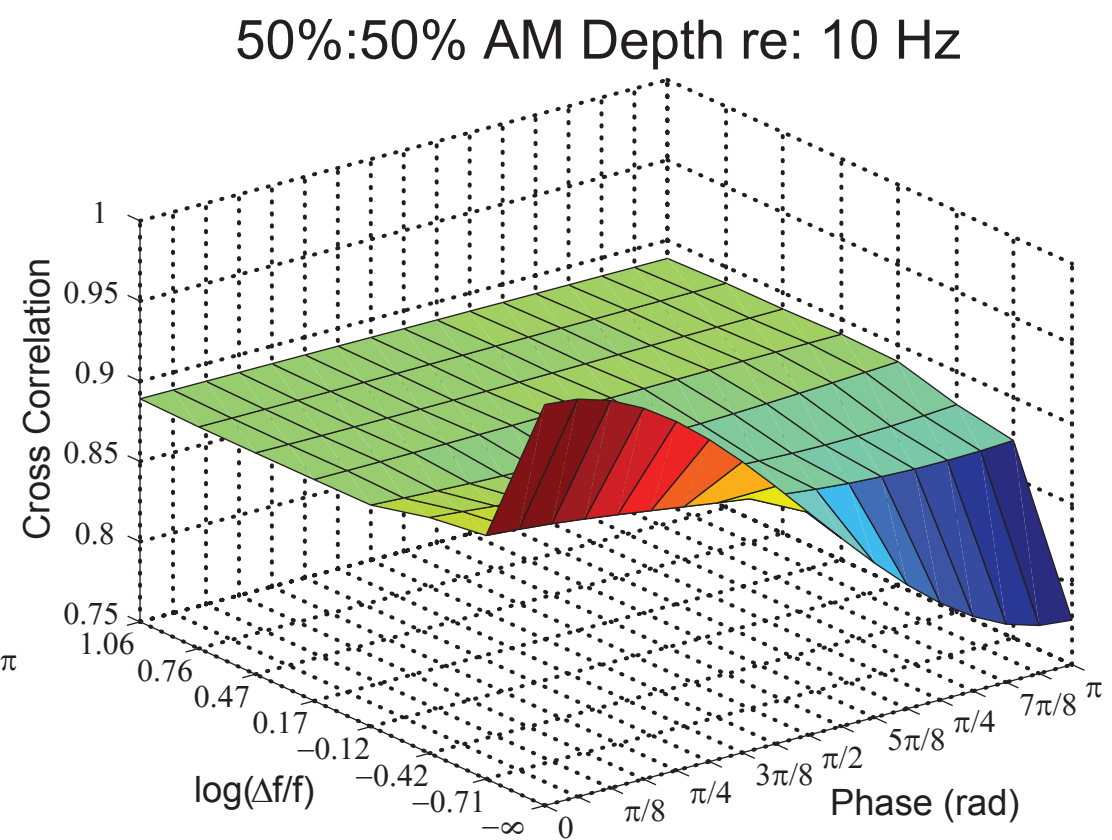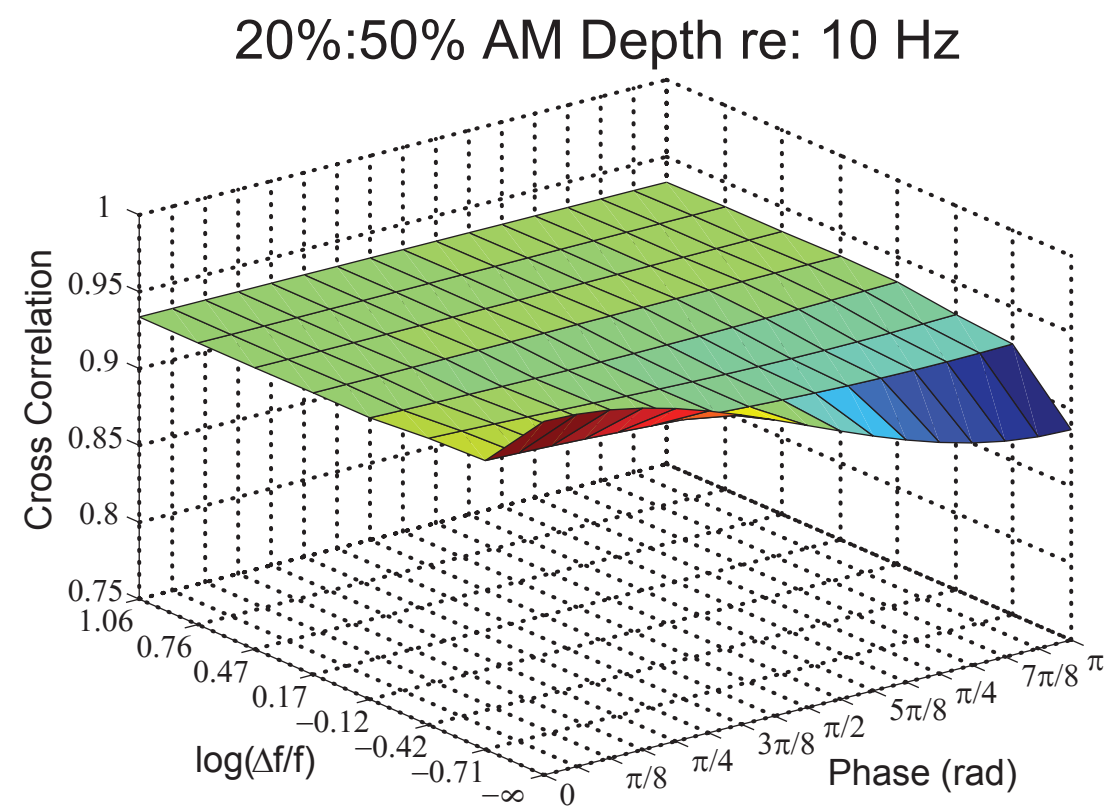
SUPPLEMENTARY FIG. 3. Cross-correlation between a signal with 50% AM depth at the standard AM rate and 0 phase (indicated in the figure title). The x-axis indicates the phase of the second signal in radians. The y-axis indicates the rate of the second signal in units of $\log(\Delta f/f)$. A value of negative infinity indicates that the rate of the second signal is equal to the standard AM rate.