Masked detection and discrimination of tone sequences under conditions of monaural and binaural masking release

Joseph W. Hall III,^{a)} Emily Buss, and John H. Grose

Department of Otolaryngology/Head and Neck Surgery, University of North Carolina School of Medicine, 170 Manning Drive, Chapel Hill, North Carolina 27599-7070

(Received 29 April 2010; revised 7 January 2011; accepted 15 January 2011)

Experiment 1 examined detection and discrimination of monaural four-tone sequences composed of 400-, 500-, and 625-Hz sinusoids. In the baseline conditions, the masker was monaural composed of 25-Hz-wide bands of random noise centered on 320, 400, 500, 625, and 781 Hz. In the binaural masking release conditions, the noise was presented diotically. In the monaural masking release conditions, the noise was presented to the same ear as the signal, but it was comodulated. Tones had half-amplitude durations of 30, 60, or 150 ms. There was no delay between successive tones, so the rate of frequency change depended on tone duration. Listeners discriminated between sequences composed of 500–400–625–500 Hz and 500–625–400–500 Hz. Discrimination results were poor for rapid sequences in both monaural and binaural masking release conditions relative to baseline conditions. Results from experiment 2 indicated that poor discrimination for rapid sequences could also occur in the baseline conditions, provided that the frequency separation among tonal components was small. Sluggish processing in the present paradigm was not restricted to conditions relying on binaural cues. It is argued that sluggishness may reflect a long temporal window in monaural and binaural masking release conditions or an interaction between poor cue quality and task difficulty. © 2011 Acoustical Society of America. [DOI: 10.1121/1.3552885]

PACS number(s): 43.66.Dc, 43.66.Ba, 43.66.pn [MAA]

Pages: 1482–1489

I. INTRODUCTION

Several experimental findings are consistent with the idea that binaural temporal resolution is sometimes relatively poor when contrasted with monaural temporal resolution (Grantham and Wightman, 1978; Kollmeier and Gilkey, 1990; Culling and Summerfield, 1998; Stellmack *et al.*, 2005; Kolarik and Culling, 2009; Krumbholz *et al.*, 2009). This is often referred to as “binaural sluggishness.” A typical experimental example of binaural sluggishness is relatively poor sensitivity to dynamic changes in interaural difference cues. For example, sinusoidal modulation of interaural time differences occurring at rates of more than a few hertz are not heard as the movement of a lateralized image, but instead as a spatially diffuse auditory image (Blauert, 1972; Grantham and Wightman, 1978). Grantham and Wightman (1979) also demonstrated binaural sluggishness for the detection of a brief $S\pi$ tone burst presented in a noise masker with interaural phase that varied sinusoidally to give a correlation that oscillated between 1.0 and -1.0 . Detectability of the brief $S\pi$ signal was relatively good when the modulation rate was very slow and the signal was presented at the moment when the masker interaural correlation was near 1.0, but this binaural detection advantage diminished rapidly with increases in modulation rate, and was essentially absent for rates of 4 Hz or higher. Krumbholz *et al.* (2009) reported that binaurally unmasked sounds were not capable of conveying pitch cues related to temporal envelope fluctuations and suggested

that this occurred because binaural sluggishness limited the coding of envelope to frequencies that were too low to give rise to the perception of pitch. In apparent contrast to these psychoacoustical studies, physiological experiments investigating brainstem sites associated with binaural hearing have shown little evidence of binaural sluggishness (Joris *et al.*, 2006; Siveke *et al.*, 2008; Fitzpatrick *et al.*, 2009; Shackleton and Palmer, 2010). This may not indicate a conflict between behavioral and physiological data because the psychophysical findings consistent with binaural sluggishness may reflect higher order brain processes that interpret information from a range of other neural populations (Joris *et al.*, 2006; Shackleton and Palmer, 2010).

Culling and Colburn (2000) reported results consistent with an interpretation that binaural sluggishness also pertains to the processing of rapid frequency changes for low sensation level (SL) $S\pi$ tone bursts that are partially masked by an No masker. The SL was based on the masked detection thresholds of individual listeners for the same stimuli. These authors suggested that the sluggishness of the binaural auditory system, modeled in terms of a long binaural temporal window (Culling and Summerfield, 1998), blurred the temporal distinctness of the dynamic tone bursts, therefore reducing the ability of the auditory system to follow the rapid frequency changes of the temporal pattern. The specific task of the listeners in their study was to discriminate the direction of frequency change (upward from downward arpeggio) in repeated, three-tone sequences composed of 400-, 500-, and 625-Hz tone bursts. The tones had no temporal gaps between them, and the parameter was the tone-burst duration (which translated to the rate of arpeggio). The

^{a)}Author to whom correspondence should be addressed. Electronic mail: jwh@med.unc.edu

arpeggios began on a randomly selected tone and were repeated throughout a 1.6-s noise burst. Threshold was defined as the signal level required to perform this discrimination with 71% accuracy. Whereas listeners could perform the task at low SLs for both So and $S\pi$ signals when the tone-burst duration was relatively long (e.g., 100 ms) and the arpeggio rate was therefore relatively low; much higher SLs were required in the $S\pi$ condition than in the So condition when the tone-burst duration was relatively short (e.g., 32–50 ms) and the arpeggio rate was therefore relatively high. The interaction between the SL required for criterion performance and the rate of tone-burst presentation was consistent with an interpretation in terms of binaural sluggishness.

The present study investigated whether the pattern of results found by Culling and Colburn (2000) is specific to binaural masking release, or whether a similar pattern also occurs under conditions of monaural masking release via comodulation of the masker bands (Hall *et al.*, 1984). As will be shown, relatively poor performance, consistent with sluggishness, was found for brief, rapid sequences in both binaural and monaural masking release conditions. It is argued below that these findings could be accounted for either in terms of relatively long time constants for both monaural and binaural masking release conditions or in terms of an interaction between cue quality and task difficulty.

II. EXPERIMENT 1. SEQUENCE DETECTION AND DISCRIMINATION

A. Methods

1. Listeners

There were seven listeners with normal hearing, aged between 20 and 56 yr (mean = 33 yr). All listeners had previous experience in psychoacoustical testing.

2. Stimuli and conditions

The stimuli consisted of pure tone sequences and narrow-band noise maskers. The task was either to detect a sequence of tones or to discriminate between a pair of four-tone sequences. The tones were sinusoids of 400, 500, and 625 Hz. The tones were gated on and off with 30-ms raised-cosine ramps and had half-rise durations of 30, 60, or 150 ms.

The masker was a set of five bands of noise, each 25-Hz wide, centered on 320, 400, 500, 625, and 781 Hz. These bands were either random noise presented monaurally (termed “Rm”), random noise presented diotically (“Ro”) or comodulated noise presented monaurally (“Cm”).¹ The stimuli were generated based on a 6103-Hz sampling rate. Each masker was generated in the frequency domain using random Gaussian draws to define the real and imaginary values of components within the passband. The comodulated noise (Cm) was generated using a single set of random draws for all five noise bands. The array was then transformed into the time domain using an inverse fast Fourier transform. The resulting waveform was composed of 2^{17} points, and, when played continuously, repeated seamlessly every 21.5 s. In all cases the maskers played continuously throughout a thresh-

old track at a pressure spectrum level of 48 dB (62 dB SPL per band).

3. Procedures

The discrimination tests used a cued, two-alternative forced choice paradigm. All trials were composed of three intervals, each containing a four-tone sequence. The standard sequence was 500–625–400–500 Hz, and the target sequence was 500–400–625–500 Hz. Each trial began with the standard sequence and was followed by two additional intervals, termed as “interval 1” and “interval 2,” one containing a replicate of the standard and the other containing the target, with equal likelihood. Each tone in a sequence was of equal amplitude and equal duration, and the four tones in each interval were presented in succession without silent gaps. The tones were always presented monaurally (Sm) to the left ear. Listening intervals were 840 ms in duration, separated by 120-ms inter-stimulus intervals and marked by lights on a handheld response box. The first tone was delayed by 60 ms relative to the onset of the light, marking the listening interval. An example trial is shown in Fig. 1, with 60-ms tone duration and the target shown as occurring in interval 1 (note the fixed 840-ms observation interval). Because the first tone was always presented 60 ms after the beginning of an 840-ms observation interval, the temporal gap between the last tone of a four-tone sequence and the first tone of the following four-tone sequence depended on tone duration.

Note that this method of stimulus presentation differs from that used by Culling and Colburn (2000), where the arpeggios began on a randomly selected tone and cycled continuously over a 1.6-s interval. Culling and Colburn noted that their task was very difficult, particularly at high rates, and could not be reliably performed by several of their prospective listeners. Most listeners who did achieve reliable performance at all rates required 5–10 h of practice. In our own pilot data using their method we also found poor or unreliable performance in several listeners, particularly at higher rates. We adopted the present method because additional pilot data indicated that it was easier for listeners while providing similar data patterns as the method used by Culling and Colburn (2000).

The detection tests used a two-alternative forced choice procedure wherein a four-tone sequence was presented in one of the two intervals selected at random, and that sequence was equally likely to be the 500–625–400–500 Hz

FIG. 1. Schematic illustration of a discrimination trial from experiment 1. The standard is a sequence of four tones; each 60 ms in duration, with frequencies ordered 500–625–400–500 Hz. In the target interval the order of the middle two tones in this sequence is reversed. In this illustration the target is in interval No. 1.

sequence as the 500–400–625–500 Hz sequence. Feedback was provided visually after each listener response.

For both discrimination and detection thresholds; the signal level was adjusted adaptively using a 3-down, 1-up tracking rule estimating 79% correct. At the beginning of a track, the signal level was adjusted in steps of 4 dB. This step size was reduced to 2 dB after the second track reversal, and the track continued until a total of eight reversals had been obtained. The final threshold estimate was calculated as the mean signal level at the last six track reversals. Three such estimates were obtained in each condition, with a fourth estimate in cases where the first three varied by 3 dB or more.

Discrimination thresholds were completed before detection thresholds. Within task, the nine conditions (3 durations $\times$ 3 maskers) were completed in a random order for all listeners except for S6 and S7, who completed the RmSm and RoSm conditions before completing the CmSm conditions.

4. Training

Prior to data collection, listeners practiced discrimination of the two four-tone sequences in the absence of masking noise, with tones fixed in level at 65-dB SPL. As in the primary discrimination conditions described above; the first interval was the standard sequence, and the target sequence was presented in either the second or third interval selected at random. Feedback was given following each listener's response, and each block consisted of 30 trials. Listeners started with the longest (150 ms) duration stimulus and practiced with that stimulus until they achieved better than 85% correct on three blocks in a row. At that point the listener began testing with the intermediate duration (60 ms), and practice continued until the 85% criterion had been met three times in a row. Once listeners reached criterion with the shortest (30 ms) duration they began adaptive testing in the presence of a background masker. One listener who was not able to achieve criterion performance with the shortest duration signal within an hour of practice was excused from the study.

B. Results and discussion

Figure 2 summarizes the detection and discrimination results for the seven listeners and also shows the mean data across all listeners. We will first consider the detection results shown with dashed lines. There were several general trends evident in the data of all listeners. The detection thresholds improved similarly with increasing signal duration across the three masker conditions with a mean effect of 5.9 dB. This effect is reflected in Fig. 2 as the relatively parallel nature of the dashed lines that are associated with the detection conditions. Another general trend was that detection thresholds were better (lower) in the masking release conditions (RoSm, CmSm) than that in the baseline masking condition (RmSm). For the three signal durations, comodulation was associated with an average masking release of 10.7–11.9 dB, and diotic masker presentation was associated with an average binaural masking level difference (MLD) of 12.5–13.2 dB. A repeated measures analysis of variance (ANOVA) was performed on the detection data. There were two within-subjects

factors: masking condition (RmSm, CmSm, and RoSm) and signal duration (30, 60, and 150 ms). There was a significant effect of masking condition ($F_{2, 12} = 314$; $p < 0.001$), a significant effect of signal duration ($F_{2, 12} = 244$; $p < 0.001$), but no interaction between masking condition and duration ($F_{4, 24} = 1.8$; $p = 0.16$). Thus, although the binaural (RoSm) and comodulation (CmSm) conditions were associated with lower thresholds than the baseline monaural (RmSm) condition, the effects of signal duration on detection were similar among the masker conditions.

The trends apparent in the discrimination results are similar in some respects to the results for detection. For example, performance tended to improve with increasing stimulus duration and was generally better in the masking release conditions (CmSm and RoSm) than in the monaural baseline condition (RmSm). An important difference between detection and discrimination, however, is that whereas the improvement with increasing signal duration was similar in magnitude across masker conditions for signal detection, this improvement appears to be greater in the masking release conditions than in the baseline condition for the discrimination results. This is reflected in Fig. 2 by the fact that the solid lines are generally steeper for the masking release conditions than that for the baseline (RmSm) condition. For example, whereas the discrimination threshold improved by 9.5 dB across signal durations from 30 to 150 ms in the baseline condition, the improvement was 14.7 and 17.6 dB, respectively, for the binaural and comodulation conditions. As a consequence, masking release for discrimination grew with increasing signal duration. Comparing data for the 30- and 150-ms signal durations, the average masking release rose from 5.1 to 10.4 dB for binaural masking release and from 1.5 to 9.6 dB for monaural masking release. An ANOVA supported these impressions. This analysis showed a significant effect of masking condition ($F_{2, 12} = 78.6$; $p < 0.001$), a significant effect of duration ($F_{2, 12} = 60.2$; $p < 0.001$), and a significant interaction between masking condition and duration ($F_{4, 24} = 10.1$; $p < 0.001$). This interaction reflects the fact that, in general, the discrimination thresholds for the three maskers were more similar at the 30 ms duration than at the longer durations, where the thresholds for the masking release conditions were considerably better than that for the baseline (RmSm) condition. Note that although many of the data are consistent with this interaction, there are exceptions: for example, S4 had relatively parallel functions with increasing signal duration for the baseline and binaural conditions.

In summary, the present results indicated some close similarities between the monaural and binaural masking release findings. For detection, there was no interaction between masking condition and duration. Both the monaural and binaural masking release remained stable and relatively large across the duration conditions. For discrimination, there was an interaction between condition and duration. This reflects the fact that thresholds improved with increasing duration more steeply for both masking release conditions than for the baseline condition. Therefore, both the monaural and the binaural masking release decreased with decreasing duration.

FIG. 2. Stimulus detection and discrimination thresholds from experiment 1 plotted as a function of signal duration. Symbol shape and line style reflect task and masker condition, as defined in the legend. The thresholds of individual listeners appear in separate panels, with mean thresholds in the right-most column and middle row.

One similarity between the present result and the results of Culling and Colburn (2000) is the relatively poor discrimination performance for the short-duration (and, therefore, rapidly presented) tone sequences in the binaural conditions (NoS π in Culling and Colburn, and RoSm in the present study). An additional finding of the present study was that the pattern of discrimination results was similar for the comodulated and binaural conditions, with relatively poor performance in both masking release conditions for the brief signal condition.

III. EXPERIMENT 2. DISCRIMINATION FOR SEQUENCES HAVING CLOSE SPECTRAL SPACING

One way to view the brief-tone discrimination findings of experiment 1 is in terms of temporal sluggishness limiting performance for brief, rapid sequences in both the binaural and monaural masking release. This could be interpreted in terms of relatively long temporal windows in the masking release conditions. However, it is important to note that relatively poor frequency discrimination has been found in binaural and monaural masking release paradigms, even for long-duration signals (Gebhardt *et al.*, 1971; Henning and Wartini, 1990; Henning, 1991; Hall *et al.*, 1997). Gebhardt *et al.* (1971) measured tone detection and masked frequency discrimination for 275 ms tones, and found that frequency difference limens were approximately two times larger for NoS π than NoSo stimuli when tones were presented at matched SLs.² Henning and Wartini (1990) found that frequency discrimination for 7-dB SL masked signals was

improved by a factor of about 1.8 as the signal duration increased from 18 to 82 ms for both NoSo and NoS π stimuli. This would suggest that the effect of signal duration on binaural frequency discrimination cannot be explained in terms of different temporal window durations.

An alternative interpretation of the brief-tone discrimination data of experiment 1 relates to the quality of frequency information available in baseline and masking release conditions. Hall *et al.* (1997) measured tone detection and pitch ranking for 400 ms pure tones. In pitch ranking, one of three different frequencies was presented with the middle frequency fixed at 500 Hz. The three tones varied in their spectral separation across conditions, and on a given trial the task of the listener was to report whether the tone was low, mid, or high in pitch. For wide spectral separation (e.g., 357–500–700 Hz), MLDs were large for both detection and pitch ranking. However, for narrow spectral separation (e.g., 488–500–512 Hz), MLDs for pitch ranking were small due to relatively high NoS π thresholds. A similar trend was found for monaural performance in modulated noise, where comodulation masking release (CMR) for pitch ranking was relatively small for narrow spectral separations. Hall *et al.* (1997) noted that the finding of poor pitch ranking of pure tones under conditions of binaural and monaural masking release was consistent with a poorer quality of cues for frequency discrimination in the masking release conditions than in the baseline conditions when measured at comparable, relatively low SLs.

The idea that cue quality is reduced in masking release conditions raises a possible interpretation of the results of

experiment 1 that is unrelated to temporal resolution. Assume for the moment that the quality of frequency coding at similar, low SLs was better in the baseline conditions than in the binaural and monaural masking release conditions of the present experiment 1 and in the binaural conditions of Culling and Colburn (2000). The discrimination of dynamic frequency changes might nevertheless have been relatively good in the masking release conditions at the longer durations tested, because the frequency separation in the task was relatively large (20% of the center frequency). However, the increased task difficulty associated with brief, rapid sequences of signals could interact with the reduced quality of frequency information in the masking release conditions, resulting in substantial performance degradation in the masking release conditions. In this case, the difficulty might not be related to a long time constant, but instead to more central auditory processes associated with the analysis of rapidly presented sequences. This point of view receives some support from the fact that listeners often require significant practice to perform order judgments for brief, rapid sequences (Divenyi and Hirsh, 1974; Culling and Colburn, 2000).

By the above reasoning, the kind of duration effect observed in the masking release conditions might also be found in the baseline monaural condition with appropriate stimuli. Relatively good performance in masking release conditions may occur in the long-duration masking release conditions, even with the assumption of relatively poor frequency coding, as long as the frequency separation is relatively large; it is only for brief stimuli that the poor frequency coding results in sharply poorer performance. A similar effect of duration might be found for the baseline condition for a smaller frequency separation. Here, discrimination performance might be relatively good for long duration tones, but the increased task difficulty associated with processing short tones in a rapid sequence could result in substantial performance degradation similar to that which occurs in the masking release conditions for more widely spaced tones. This possibility was tested in the present experiment by reducing the frequency spacing between the tones.

A. Listeners, stimuli, and procedure

Listeners S1, S4, S6, and S7 participated. The conditions were identical to the monaural (RmSm) conditions of experiment 1, except that the frequency spacing of the tones was reduced and the masker level was reduced by 10 dB. The masker level was reduced in order to avoid uncomfortably loud signals, which pilot data indicated could occur for some of the stimuli used here. Pilot data also indicated that it was problematic to identify a single frequency spacing that was similarly challenging for all listeners and yet did not result in the threshold track hitting the ceiling value (90 dB SPL). The low and high signal frequencies in experiment 1 were separated by a factor of 1.25 from the center frequency of 500 Hz. In the present experiment, listeners ran the training trials in quiet with progressively narrower tone spacing until a spacing factor was identified that supported approximately 90% correct for the 30-ms duration. This spacing was selected for further testing. For two listeners (S6 and S7) the

spacing originally identified had proved to be too large, as reflected in improved performance over time. Data were therefore replaced for these two listeners at a smaller (more difficult) spacing factor. The final values used were a factor of 1.025 (S1), 1.05 (S6 and S7), or 1.075 (S4). These monaural conditions were run in random order, blocked by condition. The $\times 1.025$ and $\times 1.05$ separation conditions were too narrow to allow discrete, 25-Hz wide masking bands. For the $\times 1.025$ separation, the masker was a contiguous band from 463 to 538 Hz; for the $\times 1.05$ separation, the masker was a contiguous band from 441 to 564 Hz. These maskers were presented at 69 dB SPL, equal to the level of five non-overlapping bands in the $\times 1.25$ separation conditions of experiment 1. The procedure was the same as that used in experiment 1 for tone discrimination.

B. Results and discussion

Figure 3 summarizes the sequence discrimination results for the four listeners. Discrimination thresholds for the wide spacing of $\times 1.25$ are replotted from the data of experiment 1 for reference. The spacing used in the reduced spacing condition for each listener is indicated in the upper right portion of the associated panel. As found previously in conditions involving the 1.25 frequency separation, thresholds in the monaural conditions with reduced tone spacing improved with increasing signal duration. This function was approximately parallel to the thresholds in the RoSm condition with wider signal separation, with one exception. For S4, thresholds in the RoSm and RmSm conditions with narrow spacing diverged for the 30 ms duration; for this listener, the effect of tone-burst duration was more pronounced in the RmSm

FIG. 3. Sequence discrimination thresholds from experiment 2 plotted as a function of signal duration. Symbol shape reflects signal and masker conditions, as defined in the legend. The thresholds of individual listeners appear in separate panels. Thresholds for the RmSm and RoSm conditions of experiment 1, both using a $\times 1.25$ frequency spacing, are replotted from Fig. 2.

condition with reduced spacing than in the other two conditions. Comparing results in the 30- and 150-ms conditions, thresholds were improved by an average of 10.2 dB (8.9–12.4 dB) in the RmSm ($\times 1.25$) condition, 15.1 dB (8.5–25.8 dB) in the RoSm ($\times 1.25$) condition, and 18.2 dB (12.7–24.1 dB) in the RmSm reduced spacing condition. The effect of signal duration was smallest for S7. For this listener, increasing the signal duration improved thresholds by 10.0 dB, 11.7 dB, and 12.7 dB in the RmSm, RoSm, and reduced spacing RmSm conditions, respectively. An ANOVA was performed with within-subject factors of duration (30, 60, and 150 ms) and spectral separation (RmSm: $\times 1.25$ and RmSm: reduced). This analysis showed a significant effect of spectral separation ($F_{1,3} = 21.0$; $p = 0.02$), a significant effect of duration ($F_{2,6} = 47.2$; $p < 0.001$), and a significant interaction between spectral separation and signal duration ($F_{2,6} = 7.9$; $p = 0.02$). The interaction confirms that there was a significantly steeper improvement in performance with increasing signal duration for the reduced frequency separation in comparison with the $\times 1.25$ separation.

These results support the notion that discrimination of frequency sequences for rapidly presented, brief tones can require relatively high signal-to-noise ratios even in the absence of masking release, as long as the frequency separation among the signals is relatively small. Note that although listener S4 appeared to show relatively little indication of binaural sluggishness in the RoSm condition of experiment 1, this listener showed poor monaural performance for discrimination of rapid sequences both in the RmSm reduced spacing condition (exp. 3) and the CmSm condition (exp. 1). That is, there is more evidence of monaural than binaural sluggishness for this listener.

IV. GENERAL DISCUSSION

A. Sluggishness for frequency discrimination and pitch perception

The present results support the previous findings of Culling and Colburn (2000) showing that under conditions of binaural masking release, listeners generally require relatively high sensation levels to attain criterion performance in discriminating sequences of rapidly presented brief tones. In experiment 1, listeners showed this type of effect under conditions that produce MLD or CMR. One interpretation of this outcome is that relatively long temporal windows underlie performance under conditions of both monaural and binaural masking release. This interpretation is consistent with the fact that the poor performance in the masking release conditions occurred when the stimuli were brief and when the task was based upon discrimination rather than detection. The idea that some aspects of monaural processing might indicate sluggishness is not new. For example, Moore and Sek (1995) suggested that the mechanism for decoding phase-locking information was sluggish in monaural hearing.

A factor complicating interpretation of the present results for brief stimuli in terms of long temporal windows is that cues for frequency discrimination may be of reduced quality for low-SL signals in both binaural and monaural masking release paradigms, even for signals of relatively

long duration (Gebhardt *et al.*, 1971; Henning, 1991; Hall *et al.*, 1997). This raises the possibility that poor performance for sequence discrimination based on frequency might not be best understood in terms of a long temporal window. Instead, such poor performance may arise due to relatively poor cues for frequency discrimination interacting with increased task difficulty when the stimuli consist of rapid sequences of brief signals. This alternative interpretation received support in experiment 2, where it was found that a similar effect (markedly poor performance for a rapid sequence) was obtained in a non-masking release condition where the spectral distance among signals was relatively small.

Culling and Colburn (2000) noted that although the frequency discrimination for rapid tones was poorer for NoS π than NoSo stimulation, their results were inconsistent with the equivalent rectangular duration (ERD) of 110 ms measured previously by Culling and Summerfield (1998): NoS π performance in the most rapid presentation conditions was better than that predicted by the 110-ms time constant. It is not clear how to view this inconsistency. One possibility is that the inconsistency occurred because the poor NoS π performance for rapid tones in the Culling and Colburn experiment was not related to binaural temporal resolution but, instead, to an interaction between cue quality and task difficulty. However, it is also possible that the poor NoS π performance was related to poor binaural temporal resolution, although not as poor as was estimated previously by Culling and Summerfield (1998).

In another paradigm examining binaural sluggishness for supra-threshold signals, Krumbholz *et al.* (2009) reported that binaurally unmasked sounds were ineffective in conveying pitch cues related to temporal envelope fluctuations. In one experiment, a set of So or S π harmonic tones band-pass filtered between 0.1 and 1.0 kHz was masked by a spectrally broader band of No masking noise. Masked thresholds were obtained for both detection and musical interval recognition. For recognition, the listener had to identify whether a harmonic complex corresponded to one of three musical intervals: a whole note, a musical third, or a musical fifth. Performance was determined for five nominal fundamental frequencies, including 48, 56, 64, 96, and 128 Hz. A key element of the approach was that pitch could be based on spectral cues for the higher fundamental frequencies, where the harmonic components were spectrally resolved, but could be based only on temporal envelope cues at the lower fundamental frequencies, where the harmonic components were too close together to be resolved. The main finding was that musical interval recognition thresholds for the S π harmonic tones worsened relatively steeply at the lower fundamental frequencies, consistent with poor use of temporal envelope cues. This finding was consistent with binaural sluggishness precluding the use of temporal envelope cues associated with pitch perception. The results of the present study suggest a possible alternative interpretation, namely that the steep worsening of S π musical interval recognition with decreasing fundamental frequency may have resulted from an interaction between reduced cue quality and task difficulty. In this case, the task difficulty may have increased with decreasing fundamental frequency because of the relatively weaker pitch associated

with the envelope cues arising from unresolved harmonics (e.g., [Shackleton and Carlyon, 1994](#)). A second possible factor related to task difficulty concerns the relationship between signal frequency and NoS π thresholds. The harmonic tones were band-pass filtered between 0.1 and 1.0 kHz, and NoS π thresholds increase with increasing signal frequency over this wide frequency range ([Hirsh, 1948](#)). It is therefore likely that the higher harmonics of the S π signal were somewhat reduced in terms of the magnitudes of their internal representations. Because the “best” frequency region for cues related to temporal envelope was likely to be at the high end of the signal pass-band (where auditory filters are widest and temporal envelope cues would therefore be best), the envelope-based task may have been more difficult for the S π than for the So signal.

B. Relation to other paradigms associated with binaural sluggishness

Several investigators have previously suggested that different binaural tasks may involve different aspects of binaural processing and therefore manifest sluggishness to different degrees ([Kollmeier and Gilkey, 1990](#); [Akeroyd and Bernstein, 2001](#); [Bernstein et al., 2001](#)). A remaining challenge is to better understand the stimulus features that are associated with sluggishness. One general suggestion is that binaural sluggishness effects are likely to pertain in masked detection tasks when the binaural difference cues in the noise background vary over time ([Yost, 1985](#); [Hall et al., 1998](#); [Goupell and Hartmann, 2007](#)). This notion is based on the fact that several studies using binaurally dynamic maskers clearly indicate binaural sluggishness in detection tasks ([Grantham and Wightman, 1979](#); [Kollmeier and Gilkey, 1990](#); [Culling and Summerfield, 1998](#); [Holube et al., 1998](#)), whereas other studies using binaurally static maskers (e.g., a diotic noise) have indicated relatively acute temporal processing of signals producing binaural difference cues ([Robinson and Trahiotis, 1972](#); [Akeroyd and Bernstein, 2001](#); [Bernstein et al., 2001](#); [Buss and Hall, 2011](#)). This idea, that binaural temporal resolution is acute in maskers having stable binaural cues, is compatible with the findings that the MLD does not decrease in magnitude when signals are very brief ([Green, 1966](#); [Robinson and Trahiotis, 1972](#); [Buss and Hall, 2011](#)), and that the threshold for a brief S π signal masked by a similarly brief No masker *improves* if the masker is given a forward fringe (e.g., [Robinson and Trahiotis, 1972](#)). The maskers were diotic in the present study and in the studies of [Culling and Colburn \(2000\)](#) and [Krumbholz et al. \(2009\)](#). If the view is valid that binaural temporal resolution is relatively good in masking noise having stable binaural cues, an interpretation of these studies in terms of an interaction between reduced cue quality and task difficulty would be favored over one involving a relatively long binaural temporal window.

[Akeroyd and Summerfield \(1999\)](#) reported a set of results that could be seen to challenge the idea that binaural sluggishness does not pertain for background noise having stable binaural cues. In their study, the task was to detect that a brief epoch in the temporal center of a 500-ms, 100-Hz wide No noise was replaced by uncorrelated (Nu) noise,

with the duration of the epoch varying adaptively during a threshold run. Although this paradigm was unique, the cues are somewhat similar to those that are available when detecting a brief S π signal in an ongoing No noise. In both paradigms, the task can be viewed in terms of sensitivity to a brief interval during which interaural correlation decreases from 1.0. The threshold obtained was approximately 10 ms at a center frequency of 500 Hz. The ERD of the binaural temporal window was then modeled using a detection criterion based on the just noticeable difference (JND) for detecting a reduction in interaural correlation from 1.0 in a second experiment, where the signal was 500-ms noise burst. Using this criterion, the ERD of the binaural temporal window was estimated to be quite long, approximately 160 ms. The authors noted, however, that this estimate would be smaller if it was assumed that the listener could take advantage of multiple looks across the 500-ms stimulus in the JND experiment. If binaural sluggishness did *not* play a role in the task, there may have been an opportunity for a large number of looks over the 500-ms interval, and that temporal integration would have led to an underestimate of the criterion decorrelation for a brief (e.g., 10 ms) signal. It is possible that data on temporal integration for interaural correlation discrimination would provide an insight into this question, but we are not aware of the previous studies that have investigated this for narrow stimulus bandwidths. In the absence of such data, it is difficult to evaluate the findings of [Akeroyd and Summerfield \(1999\)](#) with respect to the idea that binaural sluggishness might not pertain for background noise having stable binaural cues.

V. SUMMARY AND CONCLUSIONS

The present binaural results are consistent with the previous findings of [Culling and Colburn \(2000\)](#) showing that listeners required relatively high sensation levels to attain criterion performance in discriminating sequences of rapidly presented brief tones. The present study found a similar effect in a CMR paradigm and in a monaural baseline condition for a reduced spectral separation between tones. One interpretation of these results is that both monaural and binaural processing may be consistent with relatively long temporal windows under some conditions. It was also argued that poor performance for rapid sequence discrimination based on frequency could arise due to relatively poor cues for frequency discrimination interacting with increased task difficulty when stimuli are presented rapidly. A remaining challenge is to determine whether particular phenomena are best explained in terms of relatively protracted temporal windows or may be more suitably accounted for by alternative interpretations, such as an interaction between cue quality and task difficulty.

ACKNOWLEDGMENTS

Helpful comments on a previous version of this work were provided by M. Akeroyd, H. Gockel, and an anonymous reviewer. This work was supported by NIH NIDCD, R01 00397.

- ¹The spectral separation used here was chosen based on that used by Culling and Colburn (2000). With this separation, it is likely that the CMR is based upon both within-channel and across-channel monaural cues (Schooneveldt and Moore, 1987; Moore *et al.*, 1990; Grose *et al.*, 2009).
- ²It should be noted that matched SL does not correspond to matched loudness, as the loudness of a signal grows more slowly with increasing level for NoS π than NoSo stimulation (Townsend and Goldstein, 1971).
- Akeroyd, M. A., and Bernstein, L. R. (2001). "The variation across time of sensitivity to interaural disparities: Behavioral measurements and quantitative analyses," *J. Acoust. Soc. Am.* **110**, 2516–2526.
- Akeroyd, M. A., and Summerfield, A. Q. (1999). "A binaural analog of gap detection," *J. Acoust. Soc. Am.* **105**, 2807–2820.
- Bernstein, L. R., Trahiotis, C., Akeroyd, M. A., and Hartung, K. (2001). "Sensitivity to brief changes of interaural time and interaural intensity," *J. Acoust. Soc. Am.* **109**, 1604–1615.
- Blauert, J. (1972). "On the lag of lateralization caused by interaural time and intensity differences," *Audiology* **11**, 265–270.
- Buss, E., and Hall, J. W. (2011). "Effects of non-simultaneous masking on the binaural masking level difference," *J. Acoust. Soc. Am.* **129**, 907–919.
- Culling, J. F., and Colburn, H. S. (2000). "Binaural sluggishness in the perception of tone sequences and speech in noise," *J. Acoust. Soc. Am.* **107**, 517–527.
- Culling, J. F., and Summerfield, A. Q. (1998). "Measurement of the binaural temporal window using a detection task," *J. Acoust. Soc. Am.* **103**, 3540–3553.
- Divenyi, P. L., and Hirsh, I. J. (1974). "Identification of temporal order in three-tone sequences," *J. Acoust. Soc. Am.* **56**, 144–151.
- Fitzpatrick, D. C., Roberts, J. M., Kuwada, S., Kim, D. O., and Filipovic, B. (2009). "Processing temporal modulations in binaural and monaural auditory stimuli by neurons in the inferior colliculus and auditory cortex," *J. Assoc. Res. Otolaryngol.* **10**, 579–593.
- Gebhardt, C. J., Goldstein, D. P., and Robertson, R. M. (1971). "Frequency discrimination and the MLD," *J. Acoust. Soc. Am.* **51**, 1228–1232.
- Goupell, M. J., and Hartmann, W. M. (2007). "Interaural fluctuations and the detection of interaural incoherence. III. Narrowband experiments and binaural models," *J. Acoust. Soc. Am.* **122**, 1029–1045.
- Grantham, D. W., and Wightman, F. L. (1978). "Detectability of varying interaural temporal differences," *J. Acoust. Soc. Am.* **63**, 511–523.
- Grantham, D. W., and Wightman, F. L. (1979). "Detectability of a pulsed tone in the presence of a masker with time-varying interaural correlation," *J. Acoust. Soc. Am.* **65**, 1509–1517.
- Green, D. M. (1966). "Interaural phase effects in the masking of signals of different durations," *J. Acoust. Soc. Am.* **39**, 720–724.
- Grose, J. H., Buss, E., and Hall, J. W. III. (2009). "Within- and across-channel factors in the multiband comodulation masking release paradigm," *J. Acoust. Soc. Am.* **125**, 282–293.
- Hall, J. W., Grose, J. H., and Dev, M. B. (1997). "Signal detection and pitch ranking in conditions of masking release," *J. Acoust. Soc. Am.* **102**, 1746–1754.
- Hall, J. W., Grose, J. H., and Hartmann, W. M. (1998). "The masking-level difference in low-noise noise," *J. Acoust. Soc. Am.* **103**, 2573–2577.
- Hall, J. W., Haggard, M. P., and Fernandes, M. A. (1984). "Detection in noise by spectro-temporal pattern analysis," *J. Acoust. Soc. Am.* **76**, 50–56.
- Henning, G. B. (1991). "Frequency discrimination, amplitude discrimination, and the binaural masking-level difference: Some anomalous results," in *Auditory Physiology and Perception*, edited by Y. Cazals, L. Demany, and K. Horner (Pergamon Press, Oxford), pp. 507–512.
- Henning, G. B., and Wartini, S. (1990). "The effect of signal duration on frequency discrimination at low signal-to-noise ratios in different conditions of interaural phase," *Hear. Res.* **48**, 201–208.
- Hirsh, I. J. (1948). "Influence of interaural phase on interaural summation and inhibition," *J. Acoust. Soc. Am.* **20**, 536–544.
- Holube, I., Kinkel, M., and Kollmeier, B. (1998). "Binaural and monaural auditory filter bandwidths and time constants in probe tone detection experiments," *J. Acoust. Soc. Am.* **104**, 2412–2425.
- Joris, P. X., van de Sande, B., Recio-Spinoso, A., and van der Heijden, M. (2006). "Auditory midbrain and nerve responses to sinusoidal variations in interaural correlation," *J. Neurosci.* **26**, 279–289.
- Kolarik, A. J., and Culling, J. F. (2009). "Measurement of the binaural temporal window using a lateralisation task," *Hear. Res.* **248**, 60–68.
- Kollmeier, B., and Gilkey, R. (1990). "Binaural forward and backward masking: Evidence for sluggishness in binaural detection," *J. Acoust. Soc. Am.* **87**, 1709–1719.
- Krumholz, K., Magezi, D. A., Moore, R. C., and Patterson, R. D. (2009). "Binaural sluggishness precludes temporal pitch processing based on envelope cues in conditions of binaural unmasking," *J. Acoust. Soc. Am.* **125**, 1067–1074.
- Moore, B. C., and Sek, A. (1995). "Effects of carrier frequency, modulation rate, and modulation waveform on the detection of modulation and the discrimination of modulation type (amplitude modulation versus frequency modulation)," *J. Acoust. Soc. Am.* **97**, 2468–2478.
- Moore, B. C. J., Glasberg, B. R., and Schooneveldt, G. P. (1990). "Across-channel masking and comodulation masking release," *J. Acoust. Soc. Am.* **87**, 1683–1694.
- Robinson, D. E., and Trahiotis, C. (1972). "Effects of signal duration and masker duration on detectability under diotic and dichotic listening conditions," *Percept. Psychophys.* **12**, 333–334.
- Schooneveldt, G. P., and Moore, B. C. J. (1987). "Comodulation masking release (CMR): Effects of signal frequency, flanking band frequency, masker bandwidth, flanking band level, and monotic vs dichotic presentation flanking bands," *J. Acoust. Soc. Am.* **82**, 1944–1956.
- Shackleton, T. M., and Carlyon, R. P. (1994). "The role of resolved and unresolved harmonics in pitch perception and frequency modulation discrimination," *J. Acoust. Soc. Am.* **95**, 3529–3540.
- Shackleton, T. M., and Palmer, A. R. (2010). "The time course of binaural masking in the inferior colliculus of guinea pig does not account for binaural sluggishness," *J. Neurophysiol.* **104**, 189–199.
- Siveke, I., Ewert, S. D., Grothe, B., and Wiegand, L. (2008). "Psychophysical and physiological evidence for fast binaural processing," *J. Neurosci.* **28**, 2043–2052.
- Stellmack, M. A., Viemeister, N. F., and Byrne, A. J. (2005). "Monaural and interaural temporal modulation transfer functions measured with 5-kHz carriers," *J. Acoust. Soc. Am.* **118**, 2507–2518.
- Townsend, T. H., and Goldstein, D. P. (1971). "Suprathreshold binaural unmasking," *J. Acoust. Soc. Am.* **51**, 621–624.
- Yost, W. A. (1985). "Prior stimulation and the masking-level difference," *J. Acoust. Soc. Am.* **78**, 901–907.