HHS Public Access

Author manuscript

Adv Exp Med Biol. Author manuscript; available in PMC 2015 October 15.

Published in final edited form as:

Adv Exp Med Biol. 2013 ; 787: 383–390. doi:10.1007/978-1-4614-1590-9_42.

Wideband Monaural Envelope Correlation Perception

Joseph W. Hall III, **Emily Buss**, and **John H. Grose**

Department of Otolaryngology/Head and Neck Surgery, University of North Carolina School of Medicine, 170 Manning Dr., Chapel Hill, NC 27599, USA

Abstract

This study investigated monaural envelope correlation perception [Richards, V. M. 1987. J. Acoust. Soc. Am. 82, 1621–1630] for noise bandwidths ranging from 25 to 1600 Hz. The high-frequency side of the low band was fixed at 3000 Hz and the low-frequency side of the high band was fixed at 3500 Hz. When comodulated, the magnitude spectra of the pair of noise bands were either identical or reflected around the midpoint. Six listeners with normal hearing participated. Listeners showed similar performance for identical and reflected spectrum conditions, with best performance usually occurring for bandwidths between 200 and 800 Hz. Results were considered in terms of envelope comparisons of waveforms at the outputs of multiple peripheral filters or envelope comparisons of waveforms at the outputs of central filters set to the bandwidths of the noise stimuli. Some aspects of the results were incompatible with the account based on multiple peripheral filters. However, the results of a supplementary condition involving the gating of band sub-regions indicated that this incompatibility could be accounted for by non-optimal weighting of peripheral filter outputs.

Keywords

monaural envelope correlation perception; temporal processing; across-channel processing; frequency selectivity

1. INTRODUCTION

In a paradigm known as monaural envelope correlation perception (MECP), Richards (1987) demonstrated that listeners are able to discriminate between a stimulus consisting of two narrow bands of noise having random temporal envelopes and a stimulus in which the two bands are comodulated. That study, and most other investigations of MECP, used noise bandwidths of 100 Hz or less (e.g., Hall et al., 1993; Moore et al., 1990; Richards, 1987; Richards, 1988a; Richards, 1988b; Richards, 1989). A rationale for exploring these relatively narrow bandwidths is that they are narrower than the equivalent rectangular bandwidths (ERB) of auditory filters over a relatively large frequency region (Glasberg et al., 1990; Patterson, 1976). Therefore many MECP results can be conceptualized in terms of comparisons of temporal envelopes at the outputs of frequency-separated auditory filters.

Corresponding author: Joseph W. Hall III. jwh@med.unc.edu.

We recently explored MECP for noise bandwidths considerably wider than 100 Hz, examining a range of values up to 1600 Hz (Buss et al., 2012). As in most MECP studies, there were two noise bands of equal spectral width, a "low" band and a "high" band. The upper edge of the low band was fixed at 2000 Hz, and the lower edge of the high band was fixed at 2500 Hz. In one set of conditions, the comodulated bands were generated by assigning identical amplitude/phase values to corresponding frequency bins for the low and high bands. Therefore the low and high bands had identical temporal envelopes and spectral profiles for all bandwidths. These are referred to as "identical-spectrum" conditions. In a separate set of conditions, bands were generated that were comodulated but had *frequency-reversed* spectral profiles. These are referred to as "reflected-spectrum" conditions. The reflected-spectrum conditions were generated using a method developed by Richards (1988b) where the low and high bands have magnitude spectra that are mirror reflections of each other and phase spectra that are reversed and multiplied by negative one.

We found that MECP performance was best for noise bandwidths between 100 and 400 Hz, but percent correct remained above chance even at the 1600-Hz bandwidth. A finding of interest was that there was no significant difference between the results of the identical- and reflected-spectrum stimuli. Two different ideas were considered to account for sensitivity to comodulation when the noise bandwidths were relatively wide. One was that the auditory system somehow combines the outputs of the peripheral auditory filters to derive wider filter bands matched in width to the noise bandwidths. The temporal envelopes at the outputs of the derived filters could then be extracted and compared. This would be consistent with the finding of comparable results for the identical- and reflected-spectrum stimuli. The other idea was that MECP could involve a matrix of comparisons between the outputs of multiple peripheral auditory filters, with good performance depending on the combination of information from the best comparisons. For the identical-spectrum noise bands, the best comparisons would presumably involve corresponding regions of the low and high bands (i.e., the lower regions of the low band with the lower regions of the high band, and the higher regions of the low band with the higher regions of the high band). For the reflected-spectrum noise bands, the best comparisons would presumably involve spectrally rotated regions of the low and high bands (i.e., the lower regions of the low band with the higher regions of the high band, and the higher regions of the low band with the lower regions of the high band). Note that a factor that would limit performance is the difference in auditory filter widths between "matched" comparison regions, due to the increase in auditory filter width with increasing stimulus frequency (e.g., Fletcher, 1940; Glasberg et al., 1990).

As noted above, MECP performance remained above chance but showed a downturn at the widest bandwidths. It is possible that this finding is related to an effect noted by Richards (1987) where MECP performance decreased at lower stimulus frequencies. In our band-widening paradigm, the low band contained progressively lower frequencies as the noise bandwidth was increased, extending as low as 400 Hz in the widest bandwidth tested. The present study used a higher spectral region to test the possibility that no downturn in performance would occur, even at the widest bandwidths, as long as the noise bands did not include relatively low frequency components.

2. Methods

2.1 Listeners and Stimuli

Six listeners with normal hearing and previous listening experience in MECP tasks were recruited. In this study, the high-frequency cutoff of the low band was fixed at 3000 Hz, and the low-frequency cutoff of the high band was fixed at 3500 Hz. Both identical-spectrum and reflected-spectrum noises were tested for bandwidths from 25 to 1600 Hz. The low-frequency bands were generated in the frequency domain by assigning random values to the magnitude and phase components within the noise pass-band. In the comodulated identical-spectrum conditions, the high band was constructed from the same frequency-ordered draws as the low band. In the comodulated reflected-spectrum conditions, the high band was constructed from the same draws as the low band, but the Fourier components were assigned to sequential frequency bins in reverse order, and component phases were multiplied by −1. Independent random draws were used for the high-frequency noise bands in the random envelope conditions. Bands were presented for 400 ms, including 30-ms raised cosine ramps. Each band was presented at a level of 65 dB SPL. For some bandwidths of the identical-spectrum noise, some listeners reported hearing a tonal pitch corresponding to the spectral separation between matched stimulus components in the two bands (e.g., a pitch associated with 900 Hz for the 400-Hz bandwidth condition (2600–3000 Hz and 3500–3900 Hz)). A continuous, 500-Hz-wide band of masking noise centered on this difference frequency was introduced to mask this unintended cue, except for the 1600-Hz bandwidth condition where the masking band would have overlapped with part of the low band. The masking noise had a level of 47 dB SPL.

2.2 Procedure

Performance was quantified as percent correct on fixed-blocks of 25 trials. At least five blocks were completed, with additional blocks obtained if there was evidence of learning. A 3AFC procedure was used, with intervals separated by 300-ms. In two intervals, the noise bands had random envelopes, and in one, chosen at random, the bands were comodulated. Listeners responded by pressing one of three buttons on a response box. Feedback was provided after each response. Conditions were completed in random order. At the beginning of each block, listeners were given the opportunity to listen to examples of random versus comodulated bands. Here, comodulated bands were always presented in a known interval (the second of three listening intervals). The listener was told to listen to as many reminder trials as desired by continuing to press buttons 1 or 3 in response to each trial and to press button 2 when ready to start an experimental block.

3. Results and Discussion

Mean results for identical- and reflected-spectrum conditions are shown in Figure 1. All listeners achieved performance levels above chance in all conditions. As can be seen from the standard deviations shown in Figure 1, individual differences were relatively small for the middle bandwidths. At the 200 and 400-Hz bandwidths, the total range in performance among listeners was 76% to 98% correct. However, the inter-individual variation was quite large at the bandwidth extremes, ranging from 41–82% correct at the 25-Hz bandwidth, and from 45–90% correct at the 1600-Hz bandwidth. Listeners reported that the signal interval was associated with a roughness percept for bandwidths of 100 Hz or higher. As in our previous study using lower-frequency noise bands, best performance occurred at intermediate bandwidths. Also as in our previous study, the functions for the identical- and reflected-spectrum stimuli were grossly similar. A repeated-measures ANOVA was performed with two levels of spectral profile (identical- and reflected-spectrum), and seven levels of bandwidth. The analysis was performed on rationalized arcsine (RAU) transformed percent correct scores in order to stabilize the error variance associated with proportional data (Studebaker, 1985). This analysis showed a main effect of bandwidth ($F_{6,30} = 28.9$, $p < 0.001$), but no effect of spectral profile ($F_{1,5} = 0.046$, $p = 0.84$). The two-way interaction was not significant ($F_{6,30} = 2.21$, $p = 0.07$). The main effect of bandwidth was further evaluated with within-subjects contrasts. The quadratic contrast with bandwidth was significant ($F_{1,5} = 158.0$, $p < 0.001$), consistent with the trend for performance to decrease at bandwidth extremes (see Figure 1). As a coarse test of where the performance decrease began with the increase in bandwidth, t-tests were done comparing data for the 200-Hz bandwidth to data for the two widest bandwidths. Comparing the 200 and 800-Hz bandwidths showed no significant decrease in performance for either the identical-spectrum noise ($t_5=0.50$;$p=0.64$) or the reflected-spectrum noise ($t_5=1.52$;$p=0.18$). However, comparing the 200 and 1600-Hz bandwidths showed significant declines in performance for both the identical-spectrum noise ($t_5=3.83$;$p=0.012$) and the reflected-spectrum noise ($t_5=4.84$;$p=0.004$). The present results suggest that the reduction in MECP performance at the 1600-Hz bandwidth may be a robust feature of the effect rather than one related to the inclusion of lower frequencies where MECP cues are relatively weak.

A result that could be seen as being inconsistent with an interpretation in terms of envelope comparisons between the outputs of "matched" peripheral auditory filters is the reduction in performance that occurred at the widest bandwidth for the reflected-spectrum condition. In the identical-spectrum condition, increasing the bandwidth from 800 Hz to 1600 Hz resulted in larger frequency separations between all corresponding frequency regions of the low and high bands. This would result in more poorly-matched auditory filters for all corresponding frequencies in the 1600-Hz case, which might account for the drop in performance between the 800- and 1600-Hz bandwidths. In the reflected-spectrum case, band widening from 800 to 1600 Hz introduces corresponding stimulus components into increasingly mismatched auditory filters, while maintaining stimulation of the more closely matched filters associated with the 800-Hz bandwidth condition. The added frequencies in the 1600-Hz bandwidth condition would not necessarily hurt performance if the listener were able to weight information optimally. One possibility is that listeners do not weight information optimally. If this interpretation is valid, better performance in the 1600-Hz reflected spectrum condition might occur if cues were present to encourage weighting of the better-matched filters. We attempted to examine this idea in two supplementary conditions.

3.1 Supplementary Conditions

The two supplementary conditions involved reflected-spectrum, 1600-Hz bandwidth stimuli and used temporal gating manipulations intended to perceptually isolate particular 800-Hz portions of the lower and higher 1600-Hz-wide bands. In one of the conditions, the aim was to perceptually isolate the *better-matched* filters associated with the low and high bands. Here, the lower half of the low band and the higher half of the high band were presented *continuously*, and the higher half of the low band and the lower half of the high band were presented only during the three listening intervals. This gating pattern was intended to promote preferential weighting of the better-matched filters. The complementary condition was also run, where the higher half of the low band and the lower half of the high band were presented continuously, and the lower half of the low band and the higher half of the high band were presented only during the listening intervals. This gating pattern was intended to promote preferential weighting of the more poorly-matched filters.

An important feature of the continuous portions of the 1600-Hz-wide bands is that they were normally random but were comodulated in the signal interval. This was accomplished by gating off the random bands while simultaneously gating on the comodulated bands. Because the total level of the bands was unchanged over time, this transition was "seamless". The random bands were also gated off in the *non-signal* listening intervals, with separate random bands simultaneously gated on, in order to insure that the difference across intervals was related to comodulation rather than to an uncontrolled feature related to gating. As in the main experiment, the listening intervals were 400-ms in duration and were separated by 300 ms. All six listeners from the main experiment completed data collection for these conditions. Assuming that the gating manipulations resulted in the intended effects, we had the following expectations for the supplementary conditions when compared to the 800-Hz and 1600-Hz reflected-spectrum conditions of the main experiment: (1) performance for the 1600-Hz bandwidth supplementary condition with gating intended to isolate the better-matched filters should be similar to that for the 800-Hz bandwidth condition from the main experiment; (2) performance for the 1600-Hz bandwidth supplementary condition with gating intended to isolate the better-matched filters should be better than that for the 1600-Hz bandwidth condition from the main experiment; (3) performance for the 1600-Hz bandwidth condition from the main experiment should be better than that for the 1600-Hz bandwidth supplementary condition with gating intended to isolate the more poorly-matched filters.

Figure 2 shows the results of these four conditions, plotted in order of expected performance. As can be seen, some aspects of the findings were consistent with expectations. A repeated measures ANOVA was performed to compare the four reflected-spectrum conditions shown in Figure 2. In the analysis, the conditions were ordered by expected performance. This analysis showed a significant effect of condition ($F_{1,5}$=7.47; p=0.003). Planned contrasts between adjacent conditions indicated:

- **a.** The 800-Hz bandwidth condition did not differ significantly from the 1600-Hz bandwidth condition with gating intended to isolate the better-matched filters ($F_{1,5}$=0.007; p=0.938).
- **b.** The 1600-Hz bandwidth condition with gating intended to isolate the better-matched filters was associated with better performance than the 1600-Hz condition of the main experiment ($F_{1,5}$=9.355; p=0.028).

c. The 1600-Hz condition from the main experiment did not differ significantly from the condition with gating intended to isolate the more poorly-matched filters ($F_{1,5}$=0.557; p=0.489).

These results could be taken as allaying concerns about interpretation in terms of comparisons between "matched" peripheral filters, suggesting that the performance decrease in the main experiment when the reflected-spectrum bandwidth increased from 800 to 1600 Hz may be due to non-optimal weighting, where better-matched and more poorly-matched auditory filters are given similar weighting. When gating cues were provided that may discourage inclusion of more poorly-matched filters, a performance benefit was found. This outcome is also compatible with interpretation in terms of a relatively wideband central filter. Here, the width of this arbitrarily wide filter could be driven by the bandwidths of the gated stimulus components (800 Hz). With an assumption of reduced sensitivity to modulation at relatively high modulation frequencies (Bernstein et al., 2002; Dau et al., 1999; Kohlrausch et al., 2000; Viemeister, 1979), performance would be expected to be better for an 800-Hz-wide band than a 1600-Hz-wide band.

It is not clear how to interpret the finding that the 1600-Hz condition from the main experiment did not differ significantly from the condition with gating intended to isolate the more poorly-matched filters. From the standpoint of comparisons involving peripheral filters, the expectation was that performance should have been worse for the supplementary condition where noise components associated with the poorly matched filters were present only during the listening intervals.

It is possible that further insight into the mechanisms likely to underlie MECP performance may be gleaned from examination of individual differences. In contrast to the results of most of our listeners, one listener (S2) showed very reliable, comparable performance for both supplemental conditions (84.8% for the condition where the intent was to isolate the better-matched filters, and 86.4% for the condition where the intent was to isolate the more poorly-matched filters. The results of this listener are more consistent with a wideband analysis interpretation than a peripheral filter interpretation. It is possible that multiple mechanisms may underlie performance and that there are individual differences in the balance of the mechanisms. Ongoing research is investigating the validity of the gating cue for achieving perceptual isolation of noise sub-regions, and the nature of the stimulus cues that give rise to good performance.

Acknowledgments

This research was supported by NIH NIDCD grant R01-000418.

References

Bernstein LR, Trahiotis C. Enhancing sensitivity to interaural delays at high frequencies by using "transposed stimuli". J Acoust Soc Am. 2002; 112:1026–1036. [PubMed: 12243151]

Buss E, Hall JW, Grose JH. Monaural envelope correlation perception for bands narrower or wider than a critical band. J Acoust Soc Am. 2012 under review.

Dau T, Verhey J, Kohlrausch A. Intrinsic envelope fluctuations and modulation-detection thresholds for narrow-band noise carriers. J Acoust Soc Am. 1999; 106:2752–2760. [PubMed: 10573891]

Fletcher H. Auditory Patterns. Rev Mod Phys. 1940; 12:47–65.

Glasberg BR, Moore BCJ. Derivation of auditory filter shapes from notched-noise data. Hear Res. 1990; 47:103–138. [PubMed: 2228789]

Hall JW, Grose JH. Monaural envelope correlation perception in listeners with normal hearing and cochlear impairment. J Speech Hear Res. 1993; 36:1306–1314. [PubMed: 8114497]

Kohlrausch A, Fassel R. The influence of carrier level and frequency on modulation and beat-detection thresholds for sinusoidal carriers. J Acoust Soc Am. 2000; 108:723–734. [PubMed: 10955639]

Moore BCJ, Emmerich DS. Monaural envelope correlation perception, revisited: Effects of bandwidth, frequency separation, duration, and relative level of the noise bands. J Acoust Soc Am. 1990; 87:2628–2633. [PubMed: 2373797]

Patterson RD. Auditory filter shapes derived with noise stimuli. J Acoust Soc Am. 1976; 59:640–654. [PubMed: 1254791]

Richards VM. Monaural envelope correlation perception. J Acoust Soc Am. 1987; 82:1621–1630. [PubMed: 3693703]

Richards, VM. Aspects of monaural synchrony detection. In: Duifhuis, H.; Horst, J.; Wit, H., editors. Basic Issues in Hearing: Proceedings of the 8th International Symposium on Hearing. Academic Press; New York: 1988a. p. 317-322.

Richards VM. Components of monaural envelope correlation perception. Hear Res. 1988b; 35:47–58. [PubMed: 3182409]

Richards VM. Comparing monotic, diotic, and dichotic presentation modes in synchrony detection. J Acoust Soc Am Suppl. 1989; 1:s143.

Studebaker GA. A "rationalized" arcsine transform. J Speech Hear Res. 1985; 28:455–462. [PubMed: 4046587]

Viemeister NF. Temporal modulation transfer functions based upon modulation thresholds. J Acoust Soc Am. 1979; 66:1364–1380. [PubMed: 500975]

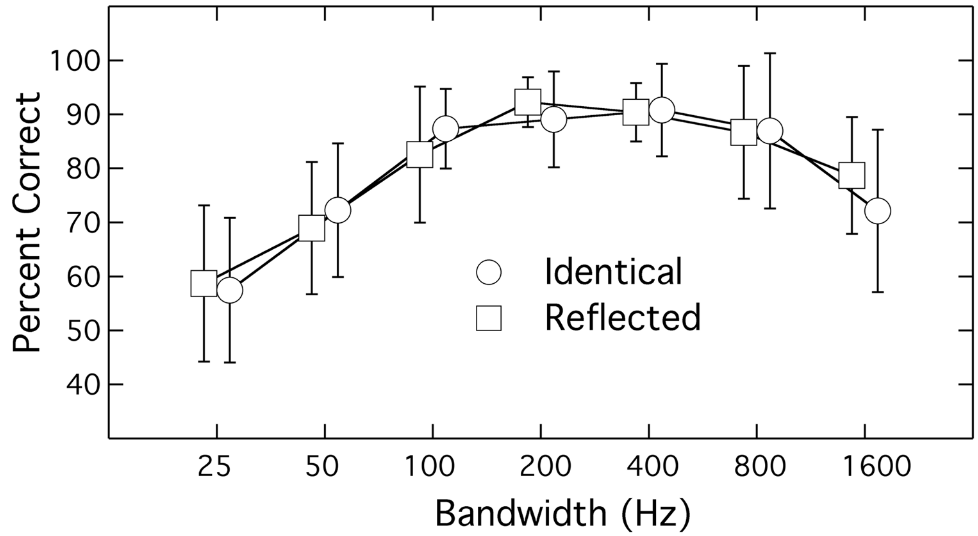

Figure 1.

Circles and squares show average percent correct for identical- and reflected-spectra, respectively. Vertical lines show plus and minus 1 standard deviation. Abscissa values are offset for clarity.

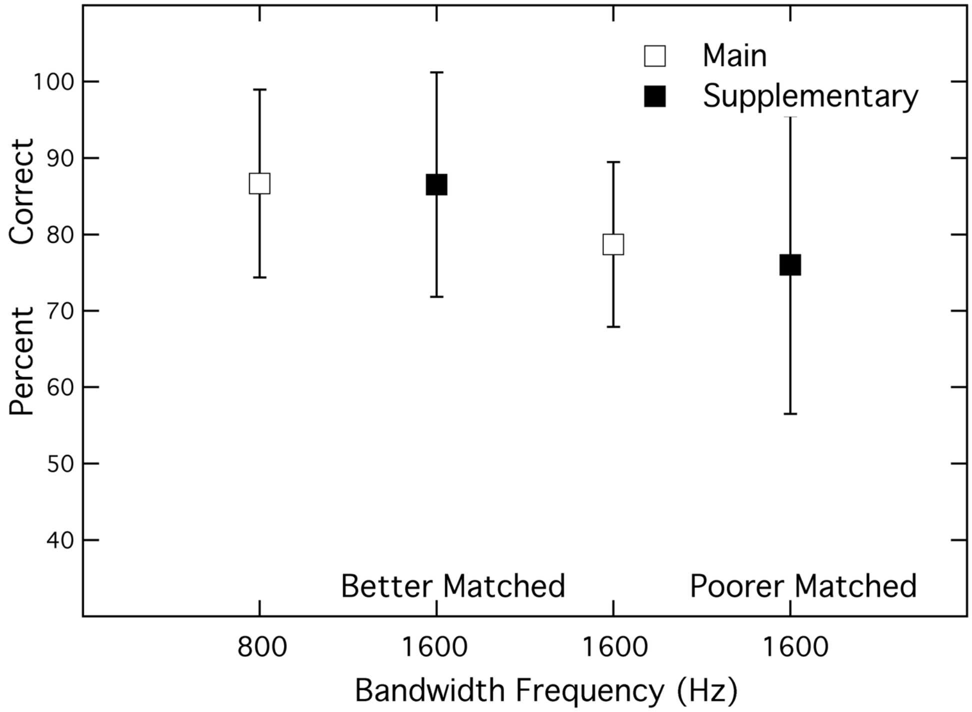

Figure 2.
Percent correct for the 800-Hz and 1600-Hz reflected-spectrum stimuli of the main experiment, and for the two supplementary 1600-Hz reflected spectrum stimuli. Gating in the supplementary conditions was intended to isolate either the better or more poorly matched auditory filters. Vertical lines show plus and minus 1 standard deviation.