The Effect of Dichotic Processing on the Perception of Binaural Cues

Akiko Amano-Kusumoto1, Justin M. Aronoff1,2, Motokuni Itoh1,3, Sigfrid D. Soli1

1Human Communication Science and Devices/House Research Institute, Los Angeles, CA, USA
2Department of Neurobiology/University of Southern California, Los Angeles, CA, USA
3Hearing Instruments Business Unit/Panasonic Healthcare Co., Ltd, Yokohama, Japan
{aamano, jaronoff, qito, ssoli}@hei.org

Abstract
Hearing impaired individuals often have difficulty hearing in noise because of reduced spectral resolution. Previous research suggests that dichotic processing, where information from neighboring frequency regions is sent to opposite ears, may benefit those individuals. However, dichotic processing can degrade binaural cues, reducing spatial release from masking and localization accuracy. In this study, an eight-channel filter bank was used to create diotic and dichotic stimuli as well as partial dichotic stimuli that used a combination of diotic and dichotic filters. To test the effect of dichotic processing on binaural cues, speech intelligibility in noise and sound localization were evaluated in normal hearing subjects. Results showed that dichotic processing degrades speech intelligibility in spatially separated noise and sound localization, but that degradation can be minimized by using partial dichotic filtering.

Index Terms: dichotic listening, upward spread of masking, release from masking, sound localization

1. Introduction
Individuals with a sensorineural hearing impairment often have difficulty hearing in noise because of reduced spectral resolutions. Dichotic processing algorithms have been developed to improve spectral resolution by reducing upward spread of masking. This is done by presenting adjacent filter bands to opposite ears (i.e., presenting odd bands to one ear and even bands to the other).

Previous studies have proposed using dichotic processing algorithms for digital hearing aids [1, 2, 3], and for cochlear implants [4]. Lunner et al. [1] showed preliminary results for speech recognition in noise with an eight-channel dichotic algorithm. An improvement on the order of 2 dB signal-to-noise ratio (SNR) was observed for two out of three hearing-impaired subjects as compared with diotic listening (i.e., presenting all bands to both ears).

Despite the potential benefits of dichotic processing, a major challenge is that such processing may also degrade binaural cues that are important for both speech perception in noise and sound localization. These binaural cues consist of interaural time differences (ITDs) at low frequencies and interaural level differences (ILDs) at high frequencies. Degradation of binaural cues is mainly because presenting different frequency bands to each ear potentially distorts ITDs and ILDs. In agreement with this concern, Tyler et al. [4] found that bilateral cochlear implant users were worse at localizing sounds when using dichotic processing. Similarly, Takagi et al. [3] found that hearing impaired listeners lateralization performance was significantly worse with dichotic processing. However, not all studies have found a decrease in sound localization performance. Kulkarni et al. [5] did not find a difference between localization with diotic and dichotic filters for normal hearing listeners using a hearing loss simulation.

To minimize the degradation of binaural cues, Takagi et al. [3] proposed an algorithm, referred to here as partial dichotic, where a dichotic filter configuration was used in the mid-frequency range and a diotic filter configuration was used for low and high frequency regions with four-channel filter bank (cut-off frequencies=100, 1250, and 4000 Hz). Their results showed significantly better spatial release from masking for the partial dichotic configuration, as compared to with the diotic configuration. In our study, which used methods similar to Takagi et al. [3], we evaluated a partial dichotic algorithm that combined diotic and dichotic filter configurations. Unlike the algorithm from Takagi et al. [3], we used a greater number of bands (8 bands) which could provide better spectral resolution. Additionally, we used equally spaced bandwidths on a logarithmic scale, which have been found to yield better dichotic benefits [5].

The first objective of the current study was to examine the effect of dichotic processing on the use of binaural cues for speech perception in noise and sound localization. Because we expect that dichotic processing would be detrimental for binaural cues, the second objective was to examine whether a partial dichotic configuration that used a combination of diotic and dichotic filters would minimize the disruption of binaural cues.

2. Experiment 1: Hearing in Noise Test (HINT)
The purpose of this experiment was to examine the effects of dichotic processing on speech reception thresholds (SRT), when speech and noise were presented with and without spatial separation.

2.1. Subjects
Eight normal hearing subjects (3 males and 5 females) with an average age of 38.1 years were tested. All had pure-tone thresholds ≤ 25 dB HL in both ears at audiometric frequencies from 250 Hz to 8000 Hz. All subjects were native speakers of English with no history of hearing disorders.

2.2. Stimuli
Stimuli were processed following the procedure shown schematically in Figure 1.
2.2.1. Test materials

The testing materials consisted of HINT sentences, each with 3–7 words [6], along with speech-shaped noise.

2.2.2. Head-related transfer functions (HRTFs)

Head-related transfer functions (HRTFs; [7, 8]) were used to simulate sound field testing based on a loudspeaker placed at ear level at a distance of 1 m. The right ear HRTF for azimuth $A^R$ was identical to the left ear HRTF for azimuth $360^\circ - A^R$, as anatomical symmetry was assumed. Previous research studies [7, 8] demonstrated comparable performance for sound-field as anatomical symmetry was assumed. The complete description of the process for deriving the HRTFs is given in [7]. Appropriate HRTFs were applied to both speech and noise, resulting in the speech source coming from $0^\circ$ and the noise source either from $0^\circ$ (noise front: NF), $+90^\circ$ (noise right: NR), or $-90^\circ$ (noise left: NL) azimuths.

2.2.3. Filter bank

An eight-channel filter bank was used to increase frequency resolution for the dichotic processing algorithm [1] as compared to the study by Takagi et al. [3]. The lowest band (Band 1) and the highest band (Band 8) were designed with low-pass (cut-off frequency=300 Hz) and highpass filters (cut-off frequency=4500 Hz), respectively. Cutoff frequencies were determined 3 dB below the passband value. The rest of the 6 filters (Bands 2–7) were designed with bandpass filters equally spaced on a logarithmic scale from 300 Hz to 4500 Hz (Center frequencies for Bands 2–7 were 386 Hz, 605 Hz, 951 Hz, 1594 Hz, 2345 Hz, and 3683 Hz). Figure 2 shows the magnitude response of the eight-channel filter bank. Each band was implemented using 1024th order linear-phase finite impulse response (FIR) filters designed with a Butterworth filter-design method using a sampling frequency of 24000 Hz.

2.2.4. Filter configuration

Three filter configurations: diotic, dichotic and partial dichotic, were used to examine the impact of dichotic processing on binaural cues. Each configuration is shown in Figure 3.

1. **Diotic configuration.** Outputs from all filter bands were mixed together for presentation to both ears. This configuration was used to examine the effects of the filter processing alone.

2. **Dichotic configuration.** Outputs of odd filter bands (Bands 1, 3, 5, and 7) were presented to the odd ear (R/L), and even bands (Bands 2, 4, 6, and 8) were presented to the even ear (L/R). This configuration was used to examine the effects of dichotic processing.

3. **Partial dichotic configuration.** Outputs of the lowest and highest two bands (Bands 1, 2, 7, and 8) were presented diotically and the 4 bands in the middle (Bands 3–6) were presented dichotically (adjacent bands were presented to different ears). This configuration was used to examine whether ITDs and ILDs were preserved.

The odd/even ear was counterbalanced for the dichotic and partial dichotic configurations.

2.3. Procedure

The experiment was conducted under headphones (Sennheiser HDA 220) in a single-walled sound room. Speech was always presented from $0^\circ$ azimuth, while noise was presented either from $0^\circ$ (NF), $+90^\circ$ (NR), or $-90^\circ$ (NL) azimuth. The speech and noise levels were scaled to meet the desired signal-to-noise ratios (SNRs) prior to the eight-channel filter bank being applied. Small level differences ($\pm 2.5$ dB) introduced to both ears as results of filtering configurations did not affect performance, and were uncompensated.

HINT measures the SRT where the listener can correctly identify sentences 50% of the time [6]. Throughout the test, the level of the noise was held at 65 dB(A) and the level of the speech was varied to determine a final threshold using a one-up/one-down procedure. The subject’s task was to repeat the sentence aloud immediately after the sentence was presented.
A response was marked correct by the tester when all words in a sentence were repeated correctly.

Each subject completed two HINT lists (one for the NF condition and one for the NR or NL condition) for each of the three filter configurations. The selection of NR/NL and the odd/even ear were counterbalanced across subjects. The order of the filter configuration was randomized across subjects.

2.4. Results and discussion

Robust statistical techniques and measures were adopted to minimize the potential effect of any outliers or non-normality in the data (see the Appendix in [8] for details and justifications for these techniques). A percentile-t bootstrap repeated measures analysis of variance (ANOVA) with 20 % trimmed means was conducted to determine the effect of the filter configuration.

For the noise side (NR and NL) test condition, there was a significant main effect of filter configuration ($F_{crit} = 8.33, F_t = 31.76$, where $F_t > F_{crit}$ indicates significant results for $\alpha = 0.05$). The 20 % trimmed mean SRTs were $-10.58, -8.47$, and $-10.48$ dB for the diotic, dichotic and partial dichotic filter configurations, respectively. Post-hoc pairwise comparisons with percentile bootstraps and 20 % trimmed mean SRTs revealed that performance with the diotic configuration was significantly better than that with the dichotic configuration (adjusted $p < 0.01$). In contrast, there was no significant difference between the diotic and the partial dichotic filter configurations ($p = 0.28$). The partial dichotic filter did not significantly affect SRTs relative to the diotic configuration.

With normal hearing subjects, we did not expect performance to be better with the dichotic configuration than with the diotic configuration. In conditions with spatial separation, ITD and ILD cues are degraded with the dichotic configuration, leading to a diotic benefit, defined as the difference between the diotic SRT and the dichotic or partial dichotic SRT in the comparable test condition. Therefore, a pairwise comparison with percentile bootstraps and 20 % trimmed means was performed on diotic benefit between dichotic and partial dichotic configurations. The result showed that the diotic benefit was significantly larger for the dichotic configuration ($p < 0.01$), suggesting that the diotic configuration was more detrimental than the partial dichotic configuration.

As expected, there was no significant effect of the filter configuration for the NF condition ($F_{crit} = 5.13, F_t = 0.99$). The 20 % trimmed mean SRTs were $-2.8$ dB, $-3.3$ dB, and $-2.8$ dB for the diotic, dichotic and partial dichotic filter configurations, respectively. The difference across filter configurations was always smaller than the measurement error associated with either individual SRT score ($\pm 0.5$ dB, adjusted for sample size) [8], shown with dashed horizontal lines in Figure 4. These results suggest that dichotic processing did not have a significant effect on speech perception performance when the speech and noise were co-located.

The HINT results showed that the dichotic processing did not degrade subjects’ speech perception when speech and noise were presented without spatial separation. This suggests that the normal hearing subjects were able to integrate the speech information despite it being divided across ears [9]. Also, the results suggest that dichotic processing can detrimentally affect binaural cues. In contrast, the partial dichotic configuration, in which lowest and highest frequency bands were diotic, did not significantly degrade performance even when the speech and noise were spatially separated.

3. Experiment 2: Source Azimuth Localization Test (SALT)

The purpose of this experiment was to examine the effects of dichotic processing on sound localization. The Source Azimuth Localization Test (SALT [8]) was used, where subjects were asked to indicate the location of a broadband impulsive sound (a gunshot). The same eight normal hearing subjects as in Experiment 1 participated in Experiment 2.

3.1. Stimuli

A broadband impulsive sound was processed with filters described in Section 2.2. The HRTFs, corresponding to one of 12 target locations from 97.5° to 262.5° azimuths with 15° in-
3.2. Procedure

Subjects were given a localization chart, which visually displays the stimulus locations, and were instructed to identify the numbered location from which the sound originated. Subjects were familiarized with the stimulus and locations prior to each test, by presenting the stimulus at each location, first in ascending order and then in descending order. A stimulus was also presented at a reference azimuths (90° for ascending-order familiarization and 270° for descending-order familiarization) prior to each familiarization presentation. All subjects completed one block of 24 stimuli (2 × 12 locations). Subjects who, at minimum, were able to distinguish adjacent groups of three locations with an accuracy of at least 75 % for the first block were presented with another block of 24 stimuli. The level of stimuli was set to 55 dB(A).

Each subject was tested with all three filter configurations, presented in a randomized order. As with Experiment 1, for the dichotic and partial dichotic configurations, the reference ear was counterbalanced across subjects.

3.3. Results and discussion

Localization performance was measured in terms of the root-mean squared (RMS) error in degrees. To compare the effect of filter configuration, a percentile-t bootstrap repeated measures analysis of variance (ANOVA) was conducted based on the 20 % trimmed means. There was a significant effect of filter configurations ($F_{2,14} = 5.12, F_1 = 6.70$; the 20 % trimmed mean RMS errors were 21.5°, 31.8°, and 22.3° for the diotic, dichotic, and partial dichotic configurations, respectively). Post-hoc pairwise comparisons revealed that performance was significantly better with the diotic than the dichotic configuration (adjusted $p < 0.01$; shown with the plus mark in Figure 5). In contrast, RMS error for the dichotic and partial dichotic configurations did not differ significantly ($p = 0.28$).

A pairwise comparison with percentile bootstraps and 20 % trimmed means showed that the diotic benefit, again defined as the difference in RMS error between the diotic configuration and either the dichotic or partial dichotic configurations, was significantly larger for the dichotic testing than for partial dichotic testing ($p < 0.01$; shown with the asterisk in Figure 5), suggesting that the dichotic configuration was more detrimental than the partial dichotic configuration. As with the results from Experiment 1, the results from Experiment 2 suggest that partial dichotic filters minimize the disruption of ITD and ILD cues that occurs with dichotic filtering.

4. Conclusions

In this study, a dichotic filtering algorithm was implemented with an eight-channel filter bank with equally spaced filters on a logarithmic scale. As for the first objective of the study, the results from HINT showed that when there is no spatial separation between speech and noise, presenting all information in both ears or dividing it across ears did not significantly affect performance for normal hearing subjects. In contrast, results from HINT with spatially separated noise and SALT showed that when binaural cues are important, the dichotic filter configuration was significantly detrimental. For the second objective, the partial dichotic filter configuration that combines diotic and dichotic processing preserved binaural cues.

In summary, the result of the current study suggests that the partial dichotic configuration has a potential to yield a dichotic benefit for listeners with reduced spectral resolution, and to minimize the disruption of binaural cues. However, by including diotic frequency regions, dichotic benefit may also be reduced with this configuration. This tradeoff between dichotic benefits and the availability of binaural cues needs to be examined in hearing-impaired listeners.

5. References


