1pPPb10. Extending Schroeder-phase masking: Influence of direction and shape of masker instantaneous frequency

Evelyn Hoglund*, YongHee Oh, Joseph F. Hribar, Kelsi J. Wittum, Megan L. Strang and Lawrence Feth

*Corresponding author's address: Speech and Hearing Science, Ohio State University, 110 Pressey Hall, Columbus, Ohio 43210, hoglund.1@osu.edu

Schroeder (1970) devised an algorithm to produce low peak factor signals. Schroeder signals with equal amplitude spectra, but reversed phase spectra, reveal large differences in masker effectiveness for listeners with normal hearing (Smith, et al., 1986). Results reported here extend previous work to include detection of multiple bursts of the same frequency, and multiple bursts that increase or decrease in frequency. Signal frequencies were selected to correspond with harmonics in the maskers. Results indicate that changing the frequency of the signal amplifies the difference between the Schroeder-phase maskers, but the direction of the change does not. When a frequency modulated tone is substituted for the Schroeder maskers, masked threshold depends on the shape of the instantaneous frequency (IF) function, as well as the direction of change. For linear FM (with IF similar to the Schroeder maskers), masked thresholds are comparable to the Schroeder-phase maskers. However, for logarithmic FM, IF changes at a constant ERB/s and directional differences are much smaller. A modified channel model (Oh, 2012) shows substantial differences in the basilar membrane response to these different maskers. [Research supported by a grant from the Office of Naval Research #N000140911017.]

Published by the Acoustical Society of America through the American Institute of Physics
INTRODUCTION

Schroeder (1970) devised an algorithm to produce broad bandwidth signals with low peak factors. Using complementary pairs of the Schroeder signals as maskers leads to large differences in masked thresholds that have been attributed to phase dispersion in the inner ear (Smith, et. al., 1986; Kohlrausch and Sander, 1995). The large difference in masked threshold for identical signals in complementary masker pairs reported for listeners with normal hearing was greatly reduced for listeners with mild to moderate sensory-neural hearing loss (Summers and Leek, 1998; Summers, 2000). For the previous work, the signals were primarily tones at various audiometric frequencies, or spoken sentences. The difference between masking for listeners with normal hearing and those with cochlear hearing impairments were attributed to the interaction between the Schroeder-phase maskers and the nonlinear processing found in healthy cochleae but absent in listeners with damaged outer hair cells.

The primary goal of the of the work reported here is to explore Schroeder-phase masking for signals more complex than the single frequency sinusoids used previously, but less complicated that spoken sentences. A second goal is to explore differences in masking due to differences in the instantaneous frequency (IF) of the masker. Signals were composed of multiple-bursts of brief tones either at a single frequency, or changing in frequency with each burst. In addition to the Schroeder-phase maskers, complementary pairs of frequency-modulated tones were used to mask these signals. The FM tone maskers had IF functions either rising or falling in frequency. The shape of the IF function was either linear (constant Hz/s), or logarithmic (constant ERB/s) in time.

METHODS

Subjects

Six normal hearing young adult listeners are reported, with a total of 10 ears included in the measurements. Normal hearing was defined as air conduction thresholds ≤20 dB HL at the standard audiometric frequencies (250-8000 Hz), with normal otoscopic findings. No prior experience with psychoacoustics research was required for participation. Subjects were able to complete all conditions in 5 one to two hour sessions, with a maximum of ten hours per ear.

Equipment

All experiments were conducted in three sound attenuated booths. Signals were generated digitally with Matlab (version 2012a), processed through RME Hammerfall DSP Multiface II audio interfaces, and presented monaurally over Sennheiser HD280 headphones.

Procedures

The experiments were conducted using the Single-Interval Adjustment Matrix (SIAM) procedure (Kaernbach, 1990). The signals were presented at a predetermined level (60, 70, or 80 dB SPL), and the masker was initially set to match the signal level, and then adjusted until the signal was just detectable at 75% accuracy. The step size for the masker level was initially set at 4 dB, and was reduced to 2 dB after the third reversal, then finally to 1 dB after the fifth reversal. Each run continued until there were 14 reversals, with the first three reversals discarded, and four reversals included in the calculation of the threshold estimate. Each trial included a 500 ms stimulus interval, followed by a 150 ms response interval, with visual cues for each. The next trial would start immediately after the subject responded. Thresholds were calculated using three separate runs for each condition. Occasionally, subjects ran more than three runs for a condition, if one or more run yielded an inconsistent result for the threshold estimate. These inconsistencies were generally a result of errors early in the run, resulting in thresholds that varied from the other runs by greater than 5 dB. When this occurred, the two runs with the closest threshold results were retained and the next run replaced the discarded run in the threshold calculation.
Stimuli

Target stimuli consisted of 10-ms sinusoidal tone bursts separated by 10 ms of silence. For single frequency signals, the signals were 10, 4 kHz bursts. For the spectrally varying signals, tone frequencies were 500, 700, 900, 1100, 1400, 1800, 2300, 3000, 3700, and 4600 Hz. These frequencies were selected to match harmonics contained within the Schroeder-phase maskers, and to approximate tones spaced in alternate ERBs. Targets spaced in approximately alternate ERBs were selected to maximize independence of the tones, and to minimize interaction between tones across frequency. Each 10 ms burst was gated on and off with quarter sine wave 5-ms rise and decay times, with no steady state component, and separated by 10 ms silent intervals, in order to minimize interaction between tones in time. Spectrotemporally varying signals were presented in ascending or descending order.

Schroeder-phase maskers consisted of harmonics 2-50 of a 100 Hz fundamental frequency tone, combined according to the algorithm proposed by Schroeder (1970). Both the Schroeder + and - maskers were used. Additionally, FM tone maskers were used with rising or falling instantaneous frequency functions covering the same frequency range as the Schroeder signals. The FM maskers had either a linear frequency sweep (linear+ and -) or a logarithmic frequency sweep (log+ and -). The linear FM tones exhibit equal energy for a constant bandwidth, such as 1 Hz. The logarithmic FM tones exhibit equal energy per octave and approximately equal energy per ERB.

RESULTS

Growth-of-masking functions averaged for the 10 ears are shown in Figure 1. Signal level is plotted on the abscissa, and masker level required to achieve 75% correct detections is on the ordinate. Each panel shows results for a different signal type, and the three different maskers are indicated by symbol color. The Schroeder-phase effect, SPE (Summers and Leek, 1998) or its equivalent for FM maskers is given in the table with Figure 1. There, the difference in masked threshold for the rising and falling frequency maskers is averaged across the three signal presentation levels. At 4 kHz, Summers and Leek reported a difference ranging from 10 to 15 dB for listeners with normal hearing. The signals used by Summers and Leek were single bursts of 260 ms duration. Those used in the present study covered a total duration of 190 ms with a 50% duty cycle. Thus, the multiple-burst signals contained less than half the energy of the signals at the same nominal presentation levels used by Summers and Leek.

The SPE for the multiple-burst tones at a fixed 4 kHz frequency is smaller than those reported by Summers and Leek for their long duration, single burst at 4 kHz. However, SPE results for the spectrotemporal signals either in the rising (STU) or falling (STD) signal conditions appear to be equal to those for the long duration signal. The equivalent rising-falling masker differences for log and linear FM maskers are quite similar to one another regardless of signal type.
DISCUSSION

In general, the Schroeder-phase effect reported by Summers and Leek (1998) is evident in the results reported here. Changing the signal from a single long-duration tone to a series of repeated brief tone bursts at the same frequency substantially reduces the SPE. We would expect the masked thresholds to be approximately 4 dB lower for both Schroeder+ and Schroeder - masking conditions, but the substantially smaller SPE indicates that the multiple burst signals may be affected differently by the maskers.

Multi-burst signals that vary systematically in frequency show SPEs that are comparable to the long duration, single frequency signal. A plausible explanation for the greater masker effectiveness of Schroeder-phase maskers with STI signals may lie in the fact that the components of the STI signals were presented in phase with their respective component in the Schroeder masker. Somewhat surprising is the fact that the direction of frequency change in the STI signal has no apparent effect on the magnitude of the resulting SPE.

When the maskers are FM tone glides, the SPEs are smaller than those for long duration single frequency signals, and there is no substantial difference among masker types or across STI direction.

ACKNOWLEDGMENTS

Research supported by a grant from the Office of Naval Research # N000140911017.

REFERENCES


