# Distribution of spectral modulation transfer functions in a young, normal-hearing population (L)

Eric C. Hoover,[a)] Ann C. Eddins, and David A. Eddins

*Department of Communication Sciences and Disorders, University of South Florida, 4202 East Fowler Avenue, PCD 1017, Tampa, Florida 32620, USA*

Spectral modulation transfer functions (SMTFs) were measured in 49 young (18–35 years of age) normal-hearing listeners. Noise carriers spanned six octaves from 200 to 12 800 Hz. Sinusoidal (on a log-amplitude scale) spectral modulation with random starting phase was superimposed on the carrier at spectral modulation frequencies of 0.25, 0.5, 1.0, 2.0, 4.0, and 8.0 cycles/octave. Modulation detection thresholds (in dB) yielded SMTFs that were bandpass in nature, consistent with previous investigations reporting data for only a few subjects. Thresholds were notably consistent across subjects despite minimal practice. Population statistics are reported that may serve as reference data for future studies. © 2018 Acoustical Society of America.
https://doi.org/10.1121/1.5020787

[JJL] Pages: 306–309

## I. INTRODUCTION

Auditory spectral processing can be conceptualized in terms of a continuum bounded by the limits of spectral resolution at one end and global spectral envelope perception at the other. Indices of spectral resolution (also termed frequency selectivity) have been used to quantify the auditory filter bandwidth within the framework of representing peripheral processing as a bank of overlapping bandpass filters. It is known, for example, that spectral resolution quantified via estimates of auditory filter bandwidth varies with center frequency, stimulus level, and degree of hearing impairment (e.g., Moore, 2007). One common method used to estimate the width of the auditory filter is a notched noise masking technique (e.g., Patterson *et al.*, 1976). Moore (1987) described the auditory filter bandwidth characteristics of a large ($n = 93$) group of young, normal-hearing listeners to provide reference data as well as a robust data set for comparison to smaller data sets that might be characterized by variability that cannot be explained easily. To date, there is no equivalent reference for measures of spectral shape perception. Several groups have used spectral shape perception to characterize the impaired auditory function because it involves the integration of stimulus intensity patterns within as well as across auditory filters (e.g., Won *et al.*, 2007; Bernstein *et al.*, 2013; Davies-Venn *et al.*, 2015). Reference data from a large group of young listeners with normal hearing can give an indication of how much variability one should expect on such a task, useful when considering data from various special populations.

A popular method for estimating spectral shape perception that invokes a linear systems approach involves estimates of a spectral modulation transfer function (SMTF) based on spectral modulation detection thresholds (e.g., Supin *et al.*, 1994; Shamma *et al.*, 1993; Eddins and Bero, 2007; Saoji *et al.*, 2009). Spectral envelope perception is an essential aspect of the pattern recognition analysis automatically executed by the auditory system as it processes everyday sounds to accomplish routine identification and discrimination tasks. Such processing is particularly evident in vowel identification (e.g., van Veen and Houtgast, 1985) and sound localization in the vertical plane (e.g., Qian and Eddins, 2008). Processing to support such pattern perception includes spectral resolution and involves encoding the pattern of intensity variations across the array of auditory filter bank outputs (i.e., the excitation pattern). This method is analogous to the procedures commonly used to study auditory temporal processing by mapping out a temporal modulation transfer function. In both cases, modulation thresholds are measured over a relevant range of modulation frequencies to characterize the domain-specific transfer function. Physiological as well as behavioral data support the hypothesis that such transfer functions represent the output of a bank of domain-specific filters tuned to different modulation frequencies (e.g., Dau *et al.*, 1997, Kowalski *et al.*, 1996; Saoji and Eddins, 2007). A number of studies have begun to compare SMTFs and spectro-temporal modulation transfer functions (STMTFs) of young, normal-hearing controls to other populations such as listeners with hearing loss, advanced age, and those using various auditory prostheses (e.g., Saoji *et al.*, 2009; Summers and Leek, 1994). To better compare within and among such populations, here we report as reference data the spectral modulation detection thresholds as a function of spectral modulation frequency for 49 young listeners with normal hearing.

## II. METHOD

### A. Subjects

A total of 50 listeners with pure-tone air- and bone-conduction thresholds within normal limits participated in the study. Of those, 49 ranged in age from 18 to 35 yr with a

a)Electronic mail: erichoover@usf.edu

mean of 24 yr, a median of 23 yr, and a standard deviation of 6.2 yr. The 50th listener was 56 years old and was excluded from summary statistics and analyses due to age. All listeners had audiometric thresholds less than 25 dB hearing level (HL). Inclusion criteria included negative history of head injury, ear disease, ear surgery, or conductive hearing loss. Additional demographic data included sex (41 female; 9 male), handedness (43 right, 6 left, 1 unspecified), and musical experience (19 had greater than 5 years of formal training). No listeners had prior experience performing psychoacoustic listening tasks.

## B. Stimuli

Stimuli represent a subset of those used in previous studies by Eddins and Bero (2007), Saoji and Eddins (2007), and Liu and Eddins (2008). Gaussian noise carriers were bandpass filtered (Butterworth) with $-3$ dB cutoff frequencies of 200 and 12 800 Hz (6 octaves) and $-32$ dB/octave slopes outside the nominal passband. These gradual slopes were chosen to minimize transients in the spectral modulation domain analogous to rise/fall windowing in the time domain. Sinusoidal spectral modulation was superimposed on the bandpass noise carrier such that the modulation was sinusoidal on a $\log_2$ frequency scale, log (dB) amplitude scale. Modulation frequencies included 0.25, 0.5, 1.0, 2.0, 4.0, and 8.0 cycles/octave. On each trial, the phase of the sinusoidal modulator was chosen randomly from a uniform distribution random ($0-2\pi$ radians), effectively rendering the use of local (in the audio-frequency domain) level cues associated with the signal interval unreliable, thereby encouraging the listener to make judgments based on the broad spectral envelope pattern spanning the audio-frequency range of the carrier. Stimuli presented in the standard (unmodulated) and signal (modulated) intervals were scaled to have an overall level of 87 dB sound pressure level (SPL), equal to a spectrum level of 45 dB. This comfortable but somewhat loud level was chosen to support comparison to future data with persons having mild-to-moderate hearing loss and to support high modulation detection thresholds that might require large modulation depths for detection.

The stimulus duration was 400 ms including a 10-ms cosine-squared rise-fall window. Digital stimuli were output via Tucker-Davis Technologies (TDT; Alachua, FL) hardware including RP2.1 (sampling rate of 48 828 Hz), PA5, HB6, and delivered to the left ear of each listener via Etymotic (Elk Grove, IL) ER-2 insert earphone. Stimuli were calibrated using an ear simulator (Brüel & Kjær model DB-100; Nærum, Denmark), [1/2]-inch pressure microphone (Brüel & Kjær model 4134), preamplifier (Brüel & Kjær model 2669), measurement amplifier (G.R.A.S. model 12AA; Holte, Denmark), and digital multi-meter (Fluke model 45; Everett, WA) with reference values established using a standard calibrator (Brüel & Kjaer model 4230).

## C. Procedure

Testing was conducted in a double-walled sound-attenuating chamber. Spectral modulation detection thresholds were measured via cued, three-interval, two-alternative forced-choice procedure with feedback via an adaptive, 3-down-1-up tracking rule estimating 79.4% correct detection (Levitt, 1971). Unmodulated carrier noises were presented in the cue and the standard intervals while a spectrally modulated carrier was presented in the signal interval. The task was to identify the interval containing spectral modulation. Stimulus intervals were marked by lights on a handheld button box. Responses were indicated by button press and correct/incorrect feedback was provided via the light above the correct interval. The 400-ms stimuli were separated by a 400-ms silent interval. The initial independent variable (modulation depth measured as the peak-to-trough difference in the modulation envelope) value for each block of trials was 15 dB with an initial step size of 3 dB for the first 3 reversals and then reduced to 0.4 dB for the remaining trials of the 75-trial block. Threshold estimates for each condition were based on the last even number of reversals, excluding the first 3 and final estimates were taken as the average of 2, 75-trial blocks. All testing was completed in a single session that consisted of pure-tone audiometry followed by familiarization with the spectral modulation detection task that consisted of threshold estimate for a spectral modulation frequency of 1.5 cycles/octave in a single 60-trial block using the adaptive tracking parameters described above. Subsequently, the order of testing for the six modulation-frequency conditions was randomized across listeners.

## III. RESULTS AND DISCUSSION

Spectral modulation detection thresholds (in dB) are shown in Fig. 1 on the ordinate with spectral modulation frequencies (in cycles per octave) on the abscissa. Thick black symbols and lines represent the mean thresholds with bars indicating the standard deviation. Thin lines show thresholds for individual listeners, with a red line showing a single older listener whose data were excluded from the following analyses. The resulting SMTFs reflect a bandpass characteristic that is slightly steeper on the high-frequency side of the function. A Kolmogorov-Smirnov test of normality was positive at all modulation frequencies ($p < 0.001$), consistent with a significant deviation from normality. As a result, tests of skewness and kurtosis were performed. Positive skew was observed at 0.25 [2.49, standard error (SE) = 0.34], 0.5 (0.89, SE = 0.34), 1.0 (1.45, SE = 0.34), 2.0 (1.50, SE = 0.34), and 4.0 (1.22, SE = 0.34) cycles per octave, and positive kurtosis at 0.25 (10.15, SE = 0.66), 1.0 (2.18, SE = 0.66), 2.0 (2.79, SE = 0.66), and 4.0 (1.74, SE = 0.66) cycles per octave. As a result of the deviation from normality at these modulation frequencies, the upper cutoff of normal spectral modulation detection was reported using median and quantile statistics. Table I displays detailed descriptive statistics for each modulation frequency, including cutoffs for both parametric and non-parametric interpretation of the normative data.

To determine whether or not thresholds differed across the six modulation frequencies, a one-way analysis of variance (ANOVA) was computed, indicating a significant effect of modulation frequency ($F_{5,288} = 26.19$, $p < 0.0001$). *Post hoc* testing (Tukey) revealed that modulation frequencies formed two clusters evenly dividing the middle frequencies

J. Acoust. Soc. Am. **143** (1), January 2018
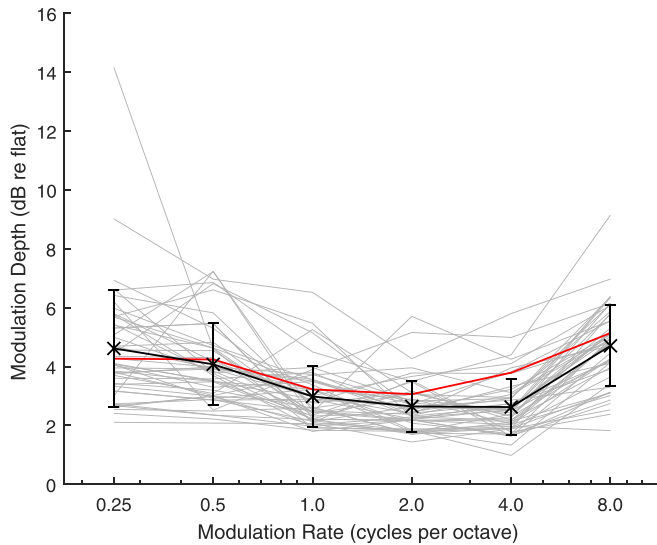
Hoover *et al.* 307

FIG. 1. (Color online) Spectral modulation detection thresholds in dB relative to a flat spectrum used in the standard stimulus as a function of spectral modulation frequency (cycles per octave). (Top) Thin lines depict individual data for 50 listeners with normal hearing. A single older listener (age: 56 yr) is depicted in red. The solid black line shows mean and standard deviation. (Bottom) Median and centile ranges are shown for the normative data. The dark shaded region shows the interquartile range and the light shaded region shows 5%–95%, indicating the limits of normal performance.

(1.0, 2.0, and 4.0 cycles per octave) from the extrema (0.25, 0.5, and 8.0 cycles per octave). Within each group, thresholds at each frequency were not significantly different from each other (all $p > 0.05$), but each frequency was significantly different from all in the other group (all $p < 0.001$). Relatively higher thresholds at 8 cycles per octave likely reflects limited frequency resolution as multiple modulation cycles are passed within a single auditory filter, effectively resulting in an averaging across cycles and resulting in a marked reduction in the effective modulation depth. Limited frequency resolution alone should result in a progressive improvement in modulation detection thresholds with decreasing modulation frequency. The plateau between 1 and 4 cycles per octave and the upturn at even lower modulation frequencies reflects the limited ability of the auditory system to compare intensity across progressive wider audio-frequency regions.

Additional analyses included correlations among spectral modulation detection thresholds in each condition and the demographic variables of sex, handedness, and musical

experience as well as conventional pure-tone average audiometric threshold (0.5, 1.0, 2.0 kHz) and high-frequency pure-tone average (2.0, 4.0, 8.0 kHz). None of those correlations reached significance (all $p > 0.05$).

The form of the average SMTF is similar to data reported previously by Eddins and Bero (2007) and Saoji and Eddins (2007) for the same stimulus generation methods and modulation frequencies and by others (e.g., Chi et al., 1999; Bernstein et al., 2013; Davies-Venn et al., 2015) for different stimulus generation methods and modulation frequencies. The only notable difference ($>0.7$ dB) between the current data and the previous studies using the same stimulus generation methods and modulation frequencies occurred for the 0.25 cycles per octave condition, were the average across 49 subjects (4.6 dB) in this study is lower than the average across 3 subjects (7.1 dB; Eddins and Bero, 2007) or 4 subjects (7.4 dB; Saoji and Eddins, 2007). Procedurally, the primary difference was an increase in presentation level in this study (87 dB SPL) versus 72 dB SPL in the other studies. All subjects in the previous studies were within the range of the data reported here, but the presentation level should be considered in the interpretation of comparisons across studies. The generality of these data is bolstered by the fact that narrowing the carrier bandwidth or shifting the center frequency up or down has little effect on modulation detection thresholds (Eddins and Bero, 2007) other than dictating the lower-frequency limit at which greater than one cycle of modulation can be carried within the nominal passband. It is also important to note that this measure of spectral shape perception, unlike traditional studies of spectral profile analysis measuring detection of an increment in a single tone of a multi-tone complex, required no substantial practice and had very consistent across observers (Drennan and Watson, 2001).

Bernstein, J. G., Mehraei, G., Shamma, S., Gallun, F. J., Theodoroff, S. M., and Leek, M. R. (2013). "Spectrotemporal modulation sensitivity as a predictor of speech intelligibility for hearing-impaired listeners," J. Am. Acad. Audiol. 24, 293–306.

Chi, T., Gao, Y., Guyton, M. C., Ru, P., and Shamma, S. (1999). "Spectro-temporal modulation transfer functions and speech intelligibility," J. Acoust. Soc. Am. 106, 2719–2732.

Dau, T., Kollmeier, B., and Kohlrausch, A. (1997). "Modeling auditory processing of amplitude modulation. I. Detection and masking with narrow-band carriers," J. Acoust. Soc. Am. 102, 2892–2905.

Davies-Venn, E., Nelson, P., and Souza, P. (2015). "Comparing auditory filter bandwidths, spectral ripple modulation detection, spectral ripple discrimination, and speech recognition: Normal and impaired hearing," J. Acoust. Soc. Am. 138, 492–503.

Drennan, W. D., and Watson, C. S. (2001). "Sources of variation in profile analysis. I. Individual differences and extended training," J. Acoust. Soc. Am. 110, 2491–2497.

TABLE I. Spectral modulation detection thresholds, including mean, standard deviation (S.D.), mean + 2 standard deviations, median, and first (Q1) and third (Q3) quartiles.

| Modulation frequency (Hz) | Mean (dB) | S.D. (dB) | +2 S.D. (dB) | Median (dB) | Q1 (dB) | Q3 (dB) |
|---|---|---|---|---|---|---|
| 0.25 | 4.62 | 2.00 | 8.61 | 4.11 | 3.39 | 5.44 |
| 0.5 | 4.08 | 1.40 | 6.87 | 3.79 | 3.00 | 4.63 |
| 1.0 | 2.98 | 1.03 | 5.06 | 2.63 | 2.22 | 3.51 |
| 2.0 | 2.65 | 0.88 | 4.40 | 2.42 | 2.16 | 3.06 |
| 4.0 | 2.63 | 0.96 | 4.55 | 2.43 | 1.98 | 3.05 |
| 8.0 | 4.71 | 1.38 | 7.47 | 4.84 | 3.92 | 5.55 |

Eddins, D. A., and Bero, E. M. (**2007**). "Spectral modulation detection as a function of modulation frequency, carrier bandwidth, and carrier frequency region," J. Acoust. Soc. Am. **121**, 363–372.

Kowalski, N., Depireux, D., and Shamma, S. (**1996**). "Analysis of dynamic spectra in ferret primary auditory cortex: Characteristics of single unit responses to moving ripple spectra," J. Neurophysiol. **76**, 3503–3523.

Levitt, H. (**1971**). "Transformed up-down methods in psychoacoustics," J. Acoust. Soc. Am. **49**, 467–477.

Liu, C., and Eddins, D. A. (**2008**). "Effects of spectral modulation filtering on vowel identification," J. Acoust. Soc. Am. **124**, 1704–1715.

Moore, B. C. J. (**1987**). "Distribution of auditory filter bandwidths at 2 kHz in young normal listeners," J. Acoust. Soc. Am. **81**, 1633–1635.

Moore, B. C. J. (**2007**). *Cochlear Hearing Loss: Physiological, Psychological, and Technical Issues*, 2nd ed. (Wiley, Chichester, UK), pp. 1–332.

Patterson, R. D. (**1976**). "Auditory filter shapes derived with noise stimuli," J. Acoust. Soc. Am. **59**, 640–654.

Qian, J., and Eddins, D. A. (**2008**). "The role of spectral modulation cues in virtual sound localization," J. Acoust. Soc. Am. **123**, 302–314.

Saoji, A. A., and Eddins, D. A. (**2007**). "Spectral modulation masking patterns reveal tuning to spectral envelope frequency," J. Acoust. Soc. Am. **122**, 1004–1013.

Saoji, A. A., Litvak, L., Spahr, A. J., and Eddins, D. A. (**2009**). "Spectral envelope detection and vowel and consonant identifications in cochlear implant listeners," J. Acoust. Soc. Am. **126**, 955–958.

Shamma, S. A., Fleshman, J. W., Wiser, P. R., and Versnel, H. (**1993**). "Organization of response areas in ferret primary auditory cortex," J. Neurophysiol. **69**, 367–383.

Summers, V., and Leek, M. R. (**1994**). "The internal representation of spectral contrast in hearing-impaired listeners," J. Acoust. Soc. Am. **95**, 3518–3528.

Supin, A. Y., Popov, V. V., Milekhina, O. N., and Tarakanov, M. B. (**1994**). "Frequency resolving power measured by rippled noise," Hear. Res. **78**, 31–40.

van Venn, T. M., and Houtgast, T. (**1985**). "Spectral sharpness and vowel dissimilarity," J. Acoust. Soc. Am. **77**, 628–634.

Won, J. H., Drennan, W. R., and Rubinstein, J. T. (**2007**). "Spectral-ripple resolution correlates with speech reception in noise in cochlear implant users," J. Assoc. Res. Otolaryngol. **8**, 384–392.

J. Acoust. Soc. Am. **143** (1), January 2018

Hoover *et al.* 309