Published in final edited form as:

Proc Meet Acoust. 2013 June ; 19(1): . doi:10.1121/1.4799742.

Effect of musical training on static and dynamic measures of spectral-pattern discrimination

Stanley Sheft*, Kirsten Smayda, Valeriy Shafiro, W. Todd Maddox, and Bharath Chandrasekaran

Abstract

Both behavioral and physiological studies have demonstrated enhanced processing of speech in challenging listening environments attributable to musical training. The relationship, however, of this benefit to auditory abilities as assessed by psychoacoustic measures remains unclear. Using tasks previously shown to relate to speech-in-noise perception, the present study evaluated discrimination ability for static and dynamic spectral patterns by 49 listeners grouped as either musicians or nonmusicians. The two static conditions measured the ability to detect a change in the phase of a logarithmic sinusoidal spectral ripple of wideband noise with ripple densities of 1.5 and 3.0 cycles per octave chosen to emphasize either timbre or pitch distinctions, respectively. The dynamic conditions assessed temporal-pattern discrimination of 1-kHz pure tones frequency modulated by different lowpass noise samples with thresholds estimated in terms of either stimulus duration or signal-to-noise ratio. Musicians performed significantly better than nonmusicians on all four tasks. Discriminant analysis showed that group membership was correctly predicted for 88% of the listeners with the structure coefficient of each measure greater than 0.51. Results suggest that enhanced processing of static and dynamic spectral patterns defined by low-rate modulation may contribute to the relationship between musical training and speech-in-noise perception. [Supported by NIH.]

INTRODUCTION

The effect of musical training on aspects of auditory processing has been studied across a wide range of tasks and procedures. Along with indicating the effect of task-specific training, one goal of this work has been to determine the extent to which the training effects generalize to domains outside music (Kraus and Chandrasekaran, 2010). Both behavioral and physiological studies have demonstrated enhanced processing of speech in challenging listening environments that is attributable to musical training (*e.g.*, Parbery-Clark *et al.*, 2009; for a recent review, see Shahin, 2011). The relationship of this benefit to psychoacoustic measures of auditory abilities at present is unclear. Frequency and pitch discrimination has received the greatest attention in psychoacoustic studies of the effect of musical training (*e.g.*, Spiegel and Watson, 1984; Micheyl *et al.*, 2006; Parbery-Clark *et al.*,

*Corresponding author's address: Communication Disorders and Sciences, Rush University Medical Center, Chicago, IL 60612, ssheft@gmail.com.

The content is solely the responsibility of the authors and does not necessarily represent the official views of NIDCD, NIMH, or the National Institutes of Health.

2009; Bidelman *et al.*, 2011; Deguchi *et al.*, 2012). Compared to nonmusicians, results have shown superior performance by musicians with the effect of psychoacoustic training of nonmusicians also evident when measured. However, past work focusing on the relationship between psychoacoustic measures and speech perception has not demonstrated involvement of frequency discrimination ability. For the related metric of frequency selectivity for which relationship to speech perception is often found, musical training does not have a significant effect on estimates of auditory filter bandwidth (Fine and Moore, 1993; Oxenham *et al.*, 2003).

Collins *et al.* (1994) suggested that use of stimulus uncertainty in psychoacoustic paradigms increases predictive power regarding listener performance in speech tasks. Past work has shown that introduction of stimulus uncertainty enhances the effect of musical training in measures of frequency discrimination (Spiegel and Watson, 1984) and signal detection (Oxenham *et al.*, 2003). Using tasks previously shown to relate to speech-in-noise perception (Sheft *et al.*, 2012a, 2012b), the present study evaluated discrimination ability for static and dynamic spectral patterns by listeners grouped as either musicians or nonmusicians, with both procedures incorporating stimulus uncertainty. The static conditions measured the ability to detect a change in the phase of a low-rate logarithmic sinusoidal spectral ripple of wideband noise with starting phase of the ripple randomized across trials. The dynamic conditions assessed temporal-pattern discrimination of 1-kHz pure tones frequency modulated by different lowpass noise samples. This discrimination of stochastic frequency modulation (FM) introduces uncertainty into the procedure through the sampling of independent noise modulators. Across conditions, parameters were chosen to coincide with aspects of musical and speech stimuli. In the static conditions, spectral-ripple densities of 1.5 and 3.0 cycles per octave (cpo) were chosen to emphasize either timbre or pitch distinctions, respectively. In the dynamic conditions, use of stochastic FM allows for discrimination by distinctive pattern features, a central aspect of speech perception both for conveying specific speech phonemes and properly segmenting and segregating the speech stream. The present study represents initial conditions of a larger project that will evaluate effects of musical training on the relationships among psychoacoustic, speech, and electrophysiological measures of auditory processing. The goal of the current work was to determine if the selected psychoacoustic tasks are appropriate for distinguishing the effect of musical training with spectrally and temporally complex stimuli relevant to a subset of signal characteristics of music and speech.

METHOD

Subjects

Data were collected from 49 normal-hearing subjects. The 23 participants classified as musicians had on average 13.4 years of musical training (S.D. = 6.8) and were practicing an instrument on average 16 hours a week (S.D. = 11). The 26 nonmusicians had on average less than 1 year of musical training (range: 0 – 3). The average age of the musicians was 25.4 years old (range: 20 – 41, 16 females), with the average age of the nonmusicians 22.1 years old (range: 19 – 30, 16 females).

Stimuli and Psychoacoustic Procedure

Spectral-Ripple Conditions—Discrimination of static spectral patterns was assessed using wideband stimuli (0.2 – 8.0 kHz) whose amplitude spectra were sinusoidally rippled in terms of the logarithms of both frequency and amplitude. Ripple density was either 1.5 or 3.0 cpo with a peak-to-trough difference of 30 dB. In the cued two-interval forced-choice (2IFC) task, the phase of the sinusoidal spectral ripple of the standard stimulus was randomized each trial with the task to detect in which of the two observation intervals the starting phase of the ripple was changed (Fig. 1). Thresholds were the just-detectable change in phase. The 500-ms rippled stimuli were shaped with a 50-ms rise/fall time, passed through a speech-shape filter emphasizing the mid frequencies, and presented to listeners at 70 dB SPL.

FM Conditions—Dynamic spectral-pattern discrimination was evaluated in terms of the ability to discriminate 1-kHz pure tones frequency modulated by different samples of lowpass noise. A consequence of the modulation is that the instantaneous frequency of the stimulus follows the amplitude pattern of the noise modulator. The maximum frequency excursion (3F) was fixed at 400 Hz to approximate formant characteristics of speech. With 3F fixed and a common sampling distribution of noise modulators minimizing long-term spectral differences, discrimination of modulated stimuli can rely on only the temporal pattern of frequency deviation (Fig. 2), rather than change in average or peak stimulus statistics.

Using a cued 2IFC procedure, discrimination thresholds were estimated with two metrics. In the first, stimulus duration was varied between trials to determine the duration at which listeners could just discriminate the fluctuating temporal patterns of a trial. Starting with a 450-ms stimulus duration and a 5-Hz modulator bandwidth, the product of duration times bandwidth was constant across change in duration. To avoid cues based on change in the long-term amplitude spectrum, on each trial, the modulator of the signal-interval stimulus was the negative of the time-reversal of the cue modulator, resulting in time-reversal of the instantaneous frequency-function. Stimulus level was 75 dB SPL.

The second metric estimated threshold as the signal-to-noise ratio (SNR) needed to just discriminate the pattern of frequency fluctuation with modulator bandwidth constant at 5 Hz and independent sampling of contrasting modulators. The 500-ms modulated stimuli were temporally centered in 1000-ms maskers. To have modulation characteristics similar to speech, maskers were speech-shaped wideband noise which was processed to include slow random variations in local fine-structure periodicities and loudness. The fine-structure periodicities were introduced through an iterative delay-add process in which delay time was dynamically varied between 0.75–3.0 ms by the time structure of 15-Hz lowpass noise. The loudness variations were achieved by comodulating the maskers with 2.5-Hz lowpass noise. Both signals and maskers were shaped with a 50-ms rise/fall time. In the task, masker level was fixed at 80 dB SPL with the level of the FM tones varied to estimate the threshold SNR.

Procedure—In the cued 2IFC procedure, the cue was the second stimulus presentation with listeners indicating their selection of which observation interval differed from the cue.

The test procedure used a modified descending method of limits. Thresholds were derived from performance on a 42-trial block, cycling through six levels of the independent variable, either ripple phase, FM duration, or FM SNR (for additional details, see Sheft *et al.*, 2012a). In all conditions, stimuli were fixed with all participants tested with the same signals presented diotically over headphones.

RESULTS

Mean thresholds from the spectral-ripple and FM conditions are shown in Figs. 3 and 4, respectively. In all conditions, on average musicians performed better than nonmusicians. In independent-samples T tests, significant group differences were obtained in the spectral-ripple task with a 1.5 cpo [$t(42.4) = 4.74, p < .001$] and 3.0 cpo [$t(47) = 4.25, p < .001$], the FM duration condition [$t(47) = 3.90, p < .001$], and in the FM SNR condition [$t(37.8) = 4.07, p < .001$]. In the 1.5-cpo ripple and FM SNR analyses, equal variances were not assumed.

Discriminant analysis was used to evaluate the efficacy of the four threshold metrics at correctly classifying participants as a musician or nonmusician. Wilks' lambda (0.453) indicated significance ($p < .001$) of the discriminant function. Analysis showed that group membership was correctly predicted for 88% of the listeners with the structure coefficient of each measure greater than 0.51.

DISCUSSION

Stimulus parameters were selected for relevance to signal characteristics of music and speech. In the spectral-ripple conditions, the ripple density of 3.0 cpo coincides with an augmented chord in music (*e.g.*, C E G#) with rippled stimuli evoking a strong musical chord percept as if played on an organ. Thresholds can then be considered as the just-detectable change in a randomly selected fundamental of a musical chord. With unit conversion of the 3.0-cpo results, mean thresholds were 0.14 and 0.25 semitones for musicians and nonmusicians, respectively. Lowering ripple density to 1.5 cpo eliminates simple musical relationships among spectral peaks so that phase discrimination is more reliant on timbre change. The better performance of musicians than nonmusicians in this condition is consistent with results from Pitt (1994) who found an effect of musical training on timbre perception. With a ripple density of 1.5 cpo, the contour of the spectral profile shows a general similarity to the formant structure of vowels, allowing for involvement of ripple-phase discrimination ability in vowel classification.

The superior performance of musicians in the FM conditions is in agreement with past work showing a beneficial effect of musical training on discrimination of monotonic pitch contours (Chandrasekaran *et al.*, 2009; Bidelman *et al.*, 2011) and detection of a single incongruity in the fundamental-frequency contour of a spoken phrase (Deguchi *et al.*, 2012). The fluctuation of the stochastic FM of the current work is nonmonotonic, extending past findings to more complex pattern discrimination. The mean FM duration thresholds of both subject groups are longer than typical formant transitions of speech, again indicating effect of pattern complexity. The mean FM SNR thresholds of roughly -13 and -9 dB for musicians and nonmusicians, respectively, show that FM pattern discrimination can be quite

robust at low sensation levels. As suggested by Sheft *et al.* (2012a, 2012b), this finding is consistent with possible involvement of FM processing in source segregation, and could possibly enhance the speech-in-noise perception of musicians.

When classifying between two groups, the canonical discriminant function coefficients correspond to the beta weights of multiple regression. These coefficients were 0.656, 0.326, 0.206, and 0.592 for the 1.5-cpo ripple, 3.0-cpo ripple, FM duration, and FM SNR conditions, respectively. The stronger associations of the 1.5-cpo ripple and FM SNR conditions to musical training was confirmed by a stepwise multiple regression in which model fit did not significantly improve with inclusion of further variables. For the aspects of psychoacoustic abilities considered in the present work, this finding indicates that timbre perception and source segregation are important factors in defining effects of musical training.

Acknowledgments

This work was supported by NIDCD (R03 DC008676 to VS and R15 DC011916 to SS) and NIMH (MH077708 and DA032457 to WTM). We thank the Maddox and the SoundBrain Lab research assistants for data collection.

References

- Bidelman GM, Gandour JT, Krishnan A. Musicians and tone-language speakers share enhanced brainstem encoding but not perceptual benefits for musical pitch. *Brain Cog.* 2011; 77:1–10.
- Chandrasekaran B, Krishnan A, Gandour JT. Relative influence of musical and linguistic experience on early cortical processing of pitch contours. *Brain Lang.* 2009; 108:1–9. [PubMed: 18343493]
- Collins LM, Wakefield GH, Feinman GR. Temporal pattern discrimination and speech recognition under electrical stimulation. *J Acoust Soc Am.* 1994; 96:2731–2737. [PubMed: 7983278]
- Deguchia C, Boureux M, Sarloa M, Besson M, Grassi M, Schön D, Colombo L. Sentence pitch change detection in the native and unfamiliar language in musicians and non-musicians: Behavioral, electrophysiological and psychoacoustic study. *Brain Res.* 2012; 1455:75–89. [PubMed: 22498174]
- Fine PA, Moore BCJ. Frequency analysis and musical ability. *Music Percept.* 1993; 11:39–54.
- Kraus N, Chandrasekaran B. Music training for the development of auditory skills. *Nature Rev Neurosci.* 2010; 11:599–605. [PubMed: 20648064]
- Micheyl C, Delhommeau K, Perrot X, Oxenham AJ. Influence of musical and psychoacoustical training on pitch discrimination. *Hear Res.* 2006; 219:36–47. [PubMed: 16839723]
- Oxenham AJ, Fligor BJ, Mason CR, Kidd G. Informational masking and musical training. *J Acoust Soc Am.* 2003; 114:1543–1549. [PubMed: 14514207]
- Parbery-Clark A, Skoe E, Lam C, Kraus N. Musician enhancement for speech-in-noise. *Ear Hear.* 2009; 30:653–661. [PubMed: 19734788]
- Pitt MA. Perception of pitch and timbre by musically trained and untrained listeners. *J Exp Psych Human Percept Perform.* 1994; 20:976–986.
- Shahin AJ. Neurophysiological influence of musical training on speech perception. *Frontiers Psych Aud Cog Neurosci* 2, article. 2011; 126:1–10.
- Sheft, S.; Risley, R.; Shafiro, V. Clinical measures of static and dynamic spectral-pattern discrimination in relationship to speech perception. In: Dau, T.; Jepsen, ML.; Poulsen, T.; Dalsgaard, JC., editors. *Speech Perception and Auditory Disorders. The Danavox Jubilee Foundation; Ballerup, Denmark: 2012a.* p. 481–488.
- Sheft S, Shafiro V, Lorenzi C, McMullen R, Farrell C. Effects of age and hearing loss on the relationship between discrimination of stochastic frequency modulation and speech perception. *Ear Hear.* 2012b; 33:709–720. [PubMed: 22790319]

Spiegel MF, Watson CS. Performance on frequency-discrimination tasks by musicians and nonmusicians. *J Acoust Soc Am*. 1984; 76:1690–1695.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

FIGURE 1.
Schematic illustration of the contrasting amplitude spectra of a discrimination trial with difference due to change in spectral-ripple starting phase; ripple density is 1.5 cpo.

FIGURE 2.

Schematic illustration of stochastic FM showing the contrasting instantaneous frequency functions of two stimuli of a discrimination trial with a 5-Hz modulator bandwidth.

FIGURE 3. Mean results from the spectral-ripple task with panels (a) and (b) for the 1.5- and 3.0-cpo conditions, respectively. Error bars represent 1 S.D.

FIGURE 4. Mean results from the FM tasks with panels (a) and (b) for the duration and SNR conditions, respectively. Error bars represent 1 S.D.