Please cite this article in press as: Bidelman GM. Amplified induced neural oscillatory activity predicts musicians' benefits in categorical speech perception. Neuroscience (2017), http://dx.doi.org/10.1016/j.neuroscience.2017.02.015

Neuroscience xxx (2017) xxx–xxx

AMPLIFIED INDUCED NEURAL OSCILLATORY ACTIVITY PREDICTS MUSICIANS' BENEFITS IN CATEGORICAL SPEECH PERCEPTION

GAVIN M. BIDELMAN *

Institute for Intelligent Systems, University of Memphis, Memphis, TN, USA

School of Communication Sciences & Disorders, University of Memphis, Memphis, TN, USA

Univeristy of Tennessee Health Sciences Center, Department of Anatomy and Neurobiology, Memphis, TN, USA

Abstract—Event-related brain potentials (ERPs) reveal musical experience refines neural encoding and confers stronger categorical perception (CP) and neural organization for speech sounds. In addition to evoked brain activity, the human EEG can be decomposed into induced (non-phase-locked) response whose various frequency bands reflect different mechanisms of perceptual-cognitive processing. Here, we aimed to clarify which spectral properties of these neural oscillations are most prone to music-related neuroplasticity and which are linked to behavioral benefits in the categorization of speech. We recorded electrical brain activity while musicians and nonmusicians rapidly identified speech tokens from a sound continuum. Time-frequency analysis parsed evoked and induced EEG into alpha- (∼10 Hz), beta- (∼20 Hz), and gamma-(>30 Hz) frequency bands. We found that musicians' enhanced behavioral CP was accompanied by improved evoked speech responses across the frequency spectrum, complementing previously observed enhancements in evoked potential studies (i.e., ERPs). Brain-behavior correlations implied differences in the underlying neural mechanisms supporting speech CP in each group: modulations in *induced* gamma power predicted the slope of musicians' speech identification functions whereas early *evoked* alpha activity predicted behavior in nonmusicians. Collectively, findings indicate that musical training tunes speech processing via two complementary mechanisms: (i) strengthening the formation of auditory object representations for speech signals (gamma-band) and (ii) improving network control and/or the matching of sounds to internalized memory templates (alpha/beta-band). Both neurobiological enhancements may be deployed behaviorally and account for musicians' benefits in the perceptual categorization of speech. © 2017 IBRO. Published by Elsevier Ltd. All rights reserved.

Key words: alpha activity, categorical speech perception, experience-dependent plasticity, induced oscillations, gamma activity.

INTRODUCTION

To successfully perceive auditory objects, the human brain must assemble diverse sensory information into common, well-formed groupings, a process known as categorical perception (CP). At its core, CP is known as the "invariance" or "many-to-one mapping" problem whereby an infinite collection of sensory features must be converted into a finite, invariant perceptual space to be acted upon by the perceptual system. CP is particularly evident in speech perception. When presented with a gradually morphed continuum of equidistant acoustic speech sounds, listeners' perception typically shifts abruptly near the midpoint, marking a change in the perceived category. CP is critical to speech-language abilities (Mody et al., 1997) and even though categorical boundaries emerge and are codified in early life (Eimas et al., 1971; Kuhl et al., 1992), whether or not they can be modified with training or experience has remained relatively unexplored (for cross-language differences in CP, see Näätänen et al., 1997; Xu et al., 2006; Bidelman and Lee, 2015).

In this regard, musicians represent an ideal model to investigate how auditory experience alters speech listening skills and its underlying neural substrates. Several recent studies have suggested that musicians have enhanced neural processing in several sensory modalities and benefits in perceptual-cognitive skills including speech and language processing (Alain et al., 2014; Moreno and Bidelman, 2014). Supporting behavioral enhancements, neurophysiological studies have revealed functional differences in brainstem and cortical neuroelectric activity in musicians in the form of larger responsiveness, higher fidelity of neural representations, and more efficient neural encoding of speech signals (Musacchia et al., 2008; Bidelman et al., 2014; Bidelman and Alain, 2015).

Extending prior studies on the benefits of musicianship to speech-language function, we recently demonstrated that musicians' speech listening skills extend to categorical processing, a higher order linguistic operation requiring a comparison between acoustic speech signals and their internalized memory representations (i.e., match to "phonetic template")

*Corresponding author. Address: School of Communication Sciences & Disorders, University of Memphis, 4055 North Park Loop, Memphis, TN 38152, USA. Fax: +1-(901)-525-1282.
E-mail address: g.bidelman@memphis.edu

Abbreviations: CP, categorical perception; ERPs, Event-related brain potentials; ITPC, inter-trial phase coherence; NMs, nonmusicians; PLI, phase-locking index.

http://dx.doi.org/10.1016/j.neuroscience.2017.02.015

(Bidelman et al., 2014). Behaviorally, musicians demonstrated faster classification and more pronounced (i.e., steeper) psychometric identification functions for speech, indicating enhancements in mapping sound objects to their categorical identities. Speech-evoked ERPs further revealed that musicians' behavioral benefits were supported by a concert of neuroplastic effects in speech encoding from brainstem to cortex (Bidelman et al., 2014; Bidelman and Alain, 2015). While these and other studies reveal early *evoked* activity underlying speech coding and how it is shaped by experience, they cannot speak to putative connections between *induced* brain responses that may also underlie skilled categorization (cf. Bidelman, 2015) nor how (if) such intrinsic brain activity is altered in an experience-dependent manner (Trainor et al., 2009).

The EEG can be described as being either "evoked" or "induced" to task manipulations (Pfurtscheller and Lopes da Silva, 1999; Shahin et al., 2009). Evoked, phase-locked responses are obtained by cross-trial averaging to derive the conventional ERP. Additionally, ongoing neural oscillatory "rhythms" not phase-locked to stimuli, can still be induced by task processing or timed stimulus events. Induced spectral measures (e.g., α, β, γ-band) are only observed via time–frequency analysis (Tallon-Baudry and Bertrand, 1999). This technique acts to circumvent the temporal jitter that normally precludes their visibility in conventional ERPs. Induced activity complements the evoked ERP by providing a window into the dynamics of speech-language function and its neural mechanisms from the perspective of intrinsic brain function. Indeed, heightened power in the γ frequency range (30–120 Hz) has been linked to synchronization of nearby brain regions (Giraud and Poeppel, 2012), auditory object construction (Tallon-Baudry and Bertrand, 1999), and semantic processing (Shahin et al., 2009). β-band (15–30 Hz) has been linked to operations related to template matching (Shahin et al., 2009) and working memory (Bashivan et al., 2014; Rose et al., 2016). More recently, we have shown that induced β and high-frequency γ rhythms accompany perceptual confusions when categorizing speech sounds, indicating a role of induced oscillations in CP (Bidelman, 2015)—at least in nonmusicians. Lastly, lower frequencies, including α-band (9–13 Hz) oscillations, have been linked to the intelligibility of speech (Becker et al., 2013) and selective inhibition of irrelevant cues (Strauß et al., 2014) when objects need to be ignored or selected against within the attentional spotlight (Foxe and Snyder, 2011). Given that various frequency components of the EEG relate to different mechanisms in speech perception, the current study aimed to further elaborate the dynamics of these brain rhythms, ascertain which spectral characteristics are associated with speech categorization, and identify which are susceptible to the neuroplastic effects of musical training.

It is well established that musicians show stronger neural encoding of musically relevant sounds as indexed by the auditory ERPs (e.g., Shahin et al., 2003; Baumann et al., 2008; Bidelman, 2013). However, recent studies have also shown that musicians have enhanced neural encoding speech (e.g., Musacchia et al., 2008; Bidelman and Krishnan, 2010; Parbery-Clark et al., 2011; Bidelman et al., 2014; Tierney et al., 2015). In the current study, we extend these previous ERP results by evaluating time–frequency measures of the EEG during speech categorization, a process requiring a listener to reconcile the acoustic speech signal with a phonetic template. Time-frequency differences between musicians and nonmusicians have been reported in response to non-speech sounds (e.g., pure tones and musical sounds; Shahin et al., 2008, 2010; Trainor et al., 2009). Yet, to our knowledge, the effects of musicianship on oscillatory brain activity have not been reported for speech, particularly during CP. Consequently, spectral analyses were expected to provide new insight into the physiological correlates and experience-dependent plasticity of CP by revealing group differences in *induced* brain responses and band-specific oscillations. We hypothesized that (i) musicians would show stronger auditory induced responses than musically naïve participants; and (ii) musicians' benefits in speech CP observed in previous ERP studies would be linked to improved neural coding, particularly in higher frequency induced activity (e.g., β and/or γ oscillations). Our predictions were based on previous studies showing that musical training strengthens induced responses to non-speech auditory stimuli (Shahin et al., 2008, 2010) and that successful speech CP is dominated by coding in β and γ oscillations, EEG bands though to carry information regarding auditory object formation and matching to internal speech templates (Shahin et al., 2009; Bidelman, 2015).

EXPERIMENTAL PROCEDURES

The current report represents a reanalysis of data from our previous ERP study (Bidelman et al., 2014) to examine experience-dependent plasticity in the *induced* components in musicians and nonmusicians.

Participants

Twenty-four young adults participated in the experiment: 12 English-speaking musicians (8 female) and 12 nonmusicians (8 female). Briefly, musicians (Ms) had received ≥7 years of private instruction on their principal instrument (13.6 ± 4.5 yrs) before age 13 (7.7 ± 3.5 yrs). Nonmusicians (NMs) had less than two years of musical instruction in their lifetime (0.4 ± 0.7 yrs). Each listener was right-handed (Oldfield, 1971) and had normal (< 25 dBHL) audiometric thresholds through 4000 Hz. Groups were otherwise matched in age (M: 23.8 ± 0.2 yrs, NM: 24.8 ± 2.7 yrs) and education (for complete demographic details, see Bidelman et al., 2014).

Stimuli and behavioral paradigm

We recorded speech ERPs to tokens along a 5-step vowel continuum where each step varied in first formant frequency (see Fig 1. in Bidelman et al., 2014). A total of 200 trials of each speech stimulus were presented to each listeners. They were required to label each token with a binary response ("u" or "a"). Interstimulus interval