Cortical processing of musical consonance: an evoked potential study

Kosuke Itoh, Shugo Suwazono and Tsutomu Nakada^CA

Center for Integrated Human Brain Science, Brain Research Institute, University of Niigata, 1-757 Asahimachi-dori, Niigata, Japan

^CA Corresponding Author: tnakada@bri.niigata-u.ac.jp

Received 26 August 2003; accepted 16 September 2003

DOI: 10.1097/01.wnr.0000099989.54721.75

Cortical processes underlying perception of musical consonance were investigated by long-latency auditory evoked potentials (EPs). Subjects listened to a random sequence of dyadic pure tones paired at various pitch intervals (1, 4, 6, 7, or 9 semitones). Amplitudes of P2 and N2 components of auditory EPs were significantly modulated by pitch interval of the dyads, being most negative for 1 semitone (minor second) and least negative or most positive for 7 semitones (perfect fifth). The results indicate that neural processing of consonance depend not only on peripheral mechanisms in the inner ear but also on higher associative processing of pitch relationships in the cerebral cortex. *NeuroReport* 14:2303–2306 © 2003 Lippincott Williams & Wilkins.

Key words: Auditory cortex; Auditory evoked potentials (AEP); Chord; Consonance; Harmony; Music

INTRODUCTION

Consonance of chords is one of the most salient features of sounds in the experience of music. The empirical rule is that chords comprised of tones with fundamentals that have simple frequency ratios are perceived as being more consonant than those with fundamentals that are related by more complex ratios. Multiple levels of auditory processing are believed to contribute to perception of consonance in music [1–3]. At the peripheral level, consonance can be defined as the absence of roughness, a condition perceived when two slightly-mistuned frequencies cannot be resolved in the cochlear basilar membrane due to its limited bandwidths [4–6]. On the other hand, later central processing is also believed to underlie consonance perception in music [1–3,7], and has been proposed to include associative processing of pitch relationships such as those related to gestalt grouping [2,7].

In contrast to successful psychoacoustical formalization of peripherally-determined aspects of consonance [4–6], central auditory processing remains poorly understood. Accordingly, we investigated cortical activities for associative processing of simultaneous pitches, by recording long-latency ($\geq$100 ms) components of auditory evoked potentials (EPs) while participants listened to dyads comprising two pure tones of various pitch intervals (1, 4, 6, 7, or 9 semitones).

MATERIALS AND METHODS

Subjects: Eight right-handed female subjects, 18–23 years old, participated in the study. None of the subjects were professional musicians or possessed absolute pitch. However, they had 3–10 years of formal training in Western tonal music, which may possibly limit the generality of the current results [8]. Informed consent was obtained from all subjects and studies were carried out in accordance with the human research guidelines of the Internal Review Board of University of Niigata.

Stimuli: Sinusoidal tones (350 ms duration, 100 ms fall-time), tuned to the equal-tempered chromatic scale (A4=440 Hz) in the range G#2 (104 Hz)–E5 (659 Hz), 16 bit, 44.1k Hz, were first created at a fixed amplitude. Subsequently, these simple tones were paired at pitch intervals 1, 4, 6, 7, and 9 semitones (see Table 1 for their conventional names in music) to produce 60 different dyads, 12 dyads per interval. The pairs were constructed so that the constituent simple tones of the 12 dyads were evenly distributed within the above-defined frequency range for each pitch interval condition, which ensured that all conditions had comparable overall frequency characteristics (range and mean). Sound intensity levels were also controlled. The stimuli were presented randomly in sequence with 900–1100 ms stimulus onset asynchrony in a single block (a total of 240 trials, 48 trials/condition). Stim software (Neurosoft, El Paso, TE, USA) running on an IBM compatible computer and air-conducted earphones were utilized for binaural stimulus presentation at a comfortable listening level.

EEG recording and analysis: Subjects were instructed to listen attentively to the stimuli while sitting in a comfortable

Table I. Effect of pitch interval on N1, P2, and N2 amplitudes.

Pitch interval (semitones)	Minor second (1)	Major third (4)	Tritone (6)	Perfect fifth (7)	Major sixth (9)
N1	-6.4 ± 2.7	-6.9 ± 2.8	-6.2 ± 2.8	-5.5 ± 2.9	-6.0 ± 3.3
P2^{a}	-2.6 ± 3.5^{b}	-1.0 ± 3.7	-0.2 ± 4.0	1.8 ± 3.4^{d}	-0.2 ± 3.1
N2^{a}	-7.0 ± 2.4^{c}	-6.1 ± 1.8	-5.6 ± 2.8	-4.2 ± 2.4^{e}	-6.1 ± 2.0

Values represent mean ± s.d. (μV).
[a] Significant effect of pitch interval, $p < 0.05$.
[b] More negative than 6, 7, and 9 semitones, $p < 0.05$.
[c] More negative than 7 semitones, $p < 0.05$.
[d] More positive than 1 and 4 semitones, $p < 0.05$.
[e] More positive than 1 semitone, $p < 0.05$.

chair in a temperature controlled, electrically shielded, sound attenuated room. Behavioral responses were not required. EEG was recorded using twenty-one silver electrodes applied according to the international 10-20 system [9], positioned at Fpz, Fp1, Fp2, Fz, F3, F4, F7, F8, Cz, C3, C4, T3, T4, T5, T6, Pz, P3, P4, Oz, O1, and O2. Horizontal (hEOG) and vertical (vEOG) electro-oculograms (EOG) were recorded simultaneously. All channels were recorded against linked earlobe electrodes. EEG and EOG were amplified by a SynAmp amplifier (Neuroscan Labs, El Paso, US.) at 16-bit resolution, gain of 500, and at an AD conversion rate of 1 kHz, band-passed between 0.05 and 100 Hz. Electrode impedance was kept below 5 KΩ throughout the experiment.

EEG data from 100 ms prior to and 800 ms following the onset of each stimulus were segmented, low-pass filtered at 30 Hz (48 dB/oct), and baseline corrected by the pre-stimulus period average. The data were then artifact rejected at ±100 μV using electrodes Fpz, Fp1, Fp2, F7, F8, T3, T4, hEOG, and vEOG. To obtain EPs, the non-rejected data segments, time-locked to the stimulus onset, were averaged separately for each pitch interval condition. Finally, the EPs were grand-averaged across all subjects. Peak amplitudes of the N1, P2, and N2 components were defined as the averaged amplitudes for time windows 95–105 ms, 160–180 ms, and 270–290 ms, respectively.

RESULTS

Typical auditory N1, P2, and N2 components were observed, peaking around 100 ms, 160–180 ms, and 270–290 ms, respectively. These waves were modulated by pitch interval of the dyads after N1 in latency (Fig. 1, Table 1). Effects of pitch interval on the amplitudes of N1, P2, and N2 components were evaluated by repeated-measures ANOVA at the vertex (Cz electrode). As a result, the main effect of pitch interval was significant for the P2 ($F(4,28)=7.9$, $p < 0.0005$) and N2 ($F(4,28)=2.8$, $p < 0.05$) components but not for the N1 component ($F(4,28)=1.3$, $p > 0.05$). Tukey-Kramer's post-hoc test revealed that the waves were most negative for the minor second (1 semitone) and least negative (or most positive) for the perfect fifth (7 semitones) for both P2 and N2 (Table 1).

DISCUSSION

Amplitudes of the P2 and N2 components of auditory evoked potentials were significantly modulated by pitch interval of simple-tone dyads. Because intensity levels and overall frequency characteristics (mean and range) were controlled between the pitch intervals, the findings indicated differential brain processing of the dyads based on their pitch-pitch relationships. As long-latency EPs were affected by pitch interval and the sources of P2 have been estimated to be located in associative auditory cortices [10–12], the results provide electrophysiological evidence that associative processing of pitches in cerebral cortex contributes to the perception of consonance and, possibly, experience of emotions [13,14], in music.

This experiment was designed to investigate cortical processing of consonance of individual sounds with special attention to minimizing effects of other cortical functions such as those related to musical expectancy and decision-making. To this end the stimuli were presented randomly in contrast to previous event-related potential (ERP) studies of consonance [15–18] in which chords were presented in specific harmonic contexts. Although deliberate construction of chord sequences may allow recording of cortical responses to target chords with or without the effect of harmonic context [15,19], priming chords can nevertheless affect the perceptual quality of the target chords in unknown ways. Random presentation would alleviate this potential problem by minimizing systematic confounds regarding musical expectancy. In addition, no behavioral responses were imposed to obviate cortical processes related to, e.g. decision making and motor execution. For these reasons the present EPs should reflect consonance processing of individual pitch intervals in a form less confounded by other cortical functions than the ERPs in previous studies.

The current results could not be fully accounted for by the peripheral roughness theory of sensory consonance [4–6]. According to this theory, sensory consonance monotonically increases with pitch distance for pure-tone dyads whose pitch intervals are greater than about 1 or 2 semitones. Therefore, if roughness were the only feature of the stimuli that determined their perceptual character, the P2 and N2 amplitudes were expected to vary in a monotonic relationship with pitch distance. On the contrary, greatest P2 and smallest N2 were elicited not by the greatest pitch interval (9 semitone) but by the perfect fifth (7 semitones), which had the simplest frequency ratio of 2:3. The results therefore indicated that cortical responses to dyads were affected not only by sensory roughness but also on other features of the stimuli concerning pitch–pitch relationships.

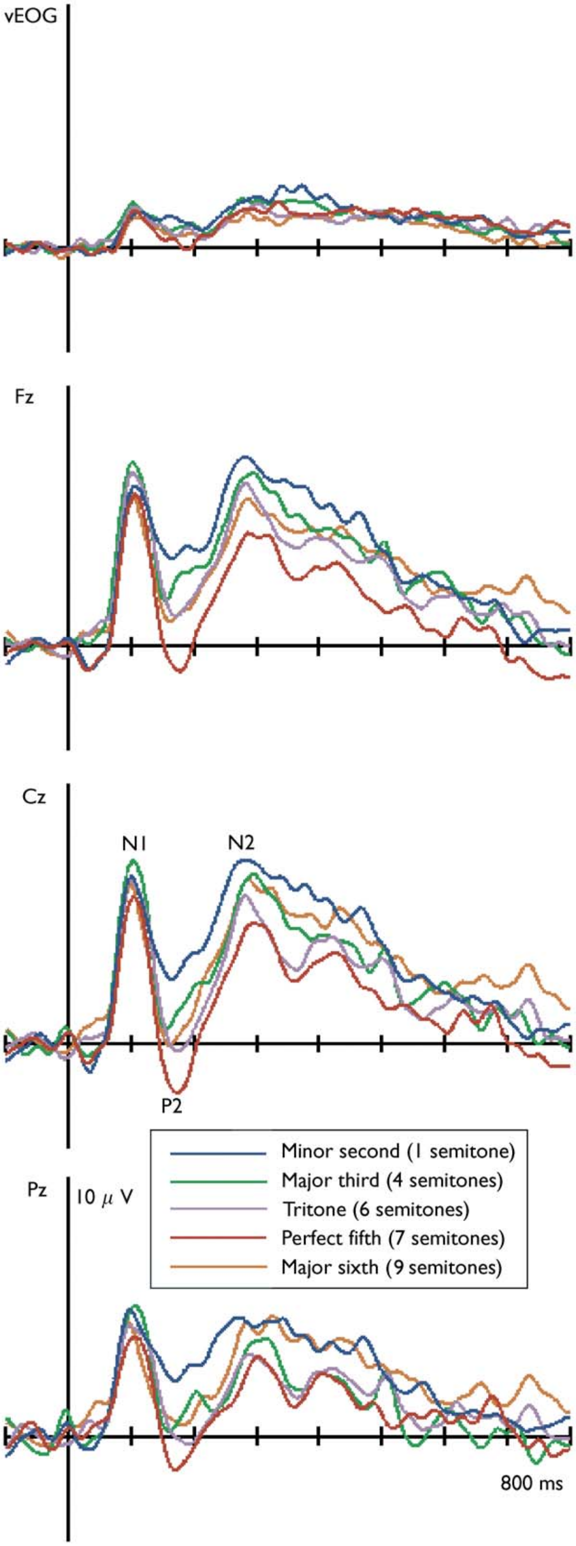

Fig. 1. Auditory evoked potentials elicited by dyadic pure tones. Whereas N1 amplitude was comparable across conditions, P2 and N2 amplitudes were significantly modulated by pitch interval of the stimuli. The elicited waves were most negative for the minor second (1 semitone) and least negative for the perfect fifth (7 semitones) in the P2 and N2 latency ranges. Statistical results are shown in Table 1.

An aspect of associative processing of sounds that may be relevant to the perception of consonance is the concept of holistic or gestalt grouping [2,7]: sounds that are easily grouped together to represent a single entity are thought to be perceived as being consonant. An electrophysiological measure related to these concepts has been found as a fronto-central negativity peaking around 180 ms, termed object-related negativity or ORN [20]. ORN increased in amplitude with greater mistuning of one of the twelve harmonics constituting the stimuli. This manipulation had the effect of making unitary grouping difficult as well as making the sound more dissonant. The current results are consistent with this finding, as the minor second (1 semitone), a dissonant interval, elicited significantly greater fronto-central negativity than the perfect fifth (7 semitones), a consonant interval, at the P2 latency (160–180 ms). Though yet to be delineated, some associative processing in the cerebral cortex may group together simultaneous pitches more readily when their frequencies are related by simpler ratios.

CONCLUSION

The study clearly demonstrated that the late components of auditory EPs can be modulated by pitch interval of dyads in a manner not completely accounted for by the roughness theory of sensory consonance. Our results provided electrophysiological evidence that, in addition to peripheral mechanisms, associative processing of pitch relationships in cerebral cortex may underlie the perception of harmony in music.

REFERENCES

1. Terhardt E. Pitch, consonance, and harmony. *J Acoust Soc Am* **55**, 1061–1069 (1974).
2. Terhardt E. The concept of musical consonance: a link between music and psychoacoustics. *Music Perception* **1**, 276–295 (1984).
3. Tramo MJ, Cariani PA, Delgutte B and Braida LD. Neurobiological foundations for the theory of harmony in Western tonal music. *Ann NY Acad Sci* **930**, 92–116 (2001).
4. Kameoka A and Kuriyagawa M. Consonance theory part I: Consonance of dyads. *J Acoust Soc Am* **45**, 1451–1459 (1969).
5. Kameoka A and Kuriyagawa M. Consonance theory part II: Consonance of complex tones and its calculation method. *J Acoust Soc Am* **45**, 1460–1469 (1969).
6. Plomp R and Levelt WJM. Tonal consonance and critical bandwidth. *J Acoust Soc Am* **38**, 548–560 (1965).
7. Bregman AS. *Auditory Scene Analysis: The Perceptual Organization of Sound*. Cambridge, MA: MIT Press, 1990.
8. Lin KL, Kobayashi M and Pascual-Leone A. Effects of musical training on speech-induced modulation in corticospinal excitability. *Neuroreport* **13**, 899–902 (2002).
9. Jasper HH. Report of the committee on methods of clinical examination in electroencephalography. *Electroencephalogr Clin Neurophysiol* **10**, 370–375 (1958).
10. Hari R, Pelizzone M, Makela JP, Hallstrom J, Leinonen L and Lounasmaa OV. Neuromagnetic responses of the human auditory cortex to on- and offsets of noise bursts. *Audiology* **26**, 31–43 (1987).
11. Papanicolaou AC, Rogers RL, Baumann S, Saydjari C and Eisenberg HM. Source localization of two evoked magnetic field components using two alternative procedures. *Exp Brain Res* **80**, 44–48 (1990).
12. Godey B, Schwartz D, de Graaf JB, Chauvel P and Liegeois-Chauvel C. Neuromagnetic source localization of auditory evoked fields and intracerebral evoked potentials: a comparison of data in the same patients. *Clin Neurophysiol* **112**, 1850–1859 (2001).

13. Altenmüller E, Schürmann K, Lim VK and Parlitz D. Hits to the left, flops to the right: different emotions during listening to music are reflected in cortical lateralisation patterns. *Neuropsychologia* **40**, 2242–2256 (2002).
14. Blood AJ, Zatorre RJ, Bermudez P and Evans AC. Emotional responses to pleasant and unpleasant music correlate with activity in paralimbic brain regions. *Nature Neurosci* **2**, 382–387 (1999).
15. Regnault P, Bigand E and Besson M. Different brain mechanisms mediate sensitivity to sensory consonance and harmonic context: evidence from auditory event-related brain potentials. *J Cogn Neurosci* **13**, 241–255 (2001).
16. Koelsch S, Gunter TC, Schroger E, Tervaniemi M, Sammler D and Friederici AD. Differentiating ERAN and MMN: an ERP study. *Neuroreport* **12**, 1385–1389 (2001).
17. Koelsch S, Maess B, Grossmann T and Friederici AD. Electric brain responses reveal gender differences in music processing. *Neuroreport* **14**, 709–713 (2003).
18. Patel AD, Gibson E, Ratner J, Besson M and Holcomb PJ. Processing syntactic relations in language and music: an event-related potential study. *J Cogn Neurosci* **10**, 717–733 (1998).
19. Bigand E and Pineau M. Global context effects on musical expectancy. *Percept Psychophys* **59**, 1098–1107 (1997).
20. Alain C, Arnott SR and Picton TW. Bottom-up and top-down influences on auditory scene analysis: evidence from event-related potentials. *J Exp Psychol: Hum Percept Perform* **27**, 1072–1089 (2001).

Acknowledgements: We thank Dr A. Izumi at Kyoto University for his helpful comments. The work was supported by grants from the Ministry of Education, Culture, Sports, Science, and Technology (Japan).