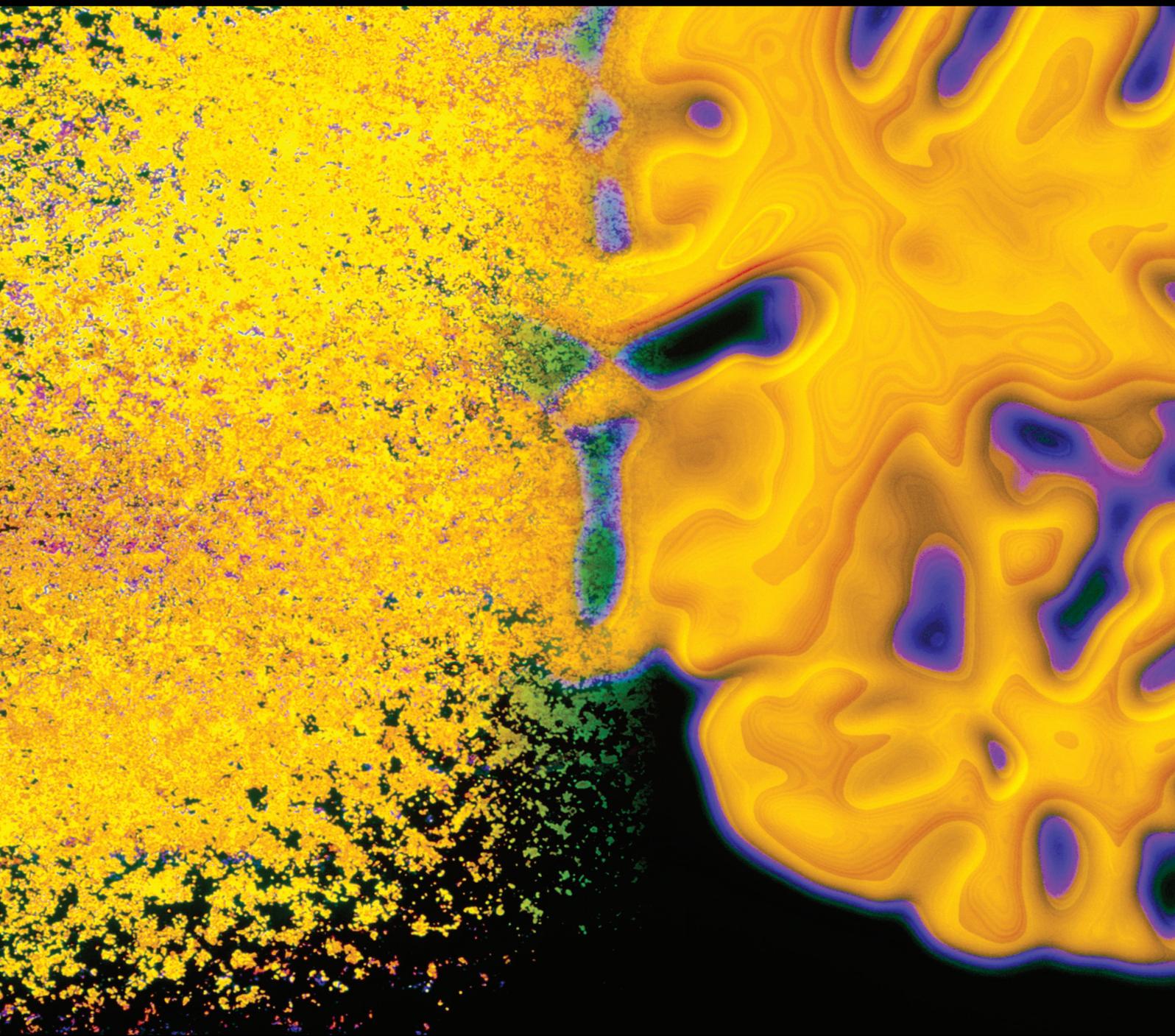


Music in the Brain: From Listening to Playing

Guest Editor: Masayuki Satoh, Stefan Evers, Shinichi Furuya, and Kentaro Ono





Music in the Brain: From Listening to Playing

Behavioural Neurology

Music in the Brain: From Listening to Playing

Guest Editors: Masayuki Satoh, Stefan Evers,
Shinichi Furuya, and Kentaro Ono



Copyright © 2015 Hindawi Publishing Corporation. All rights reserved.

This is a special issue published in “Behavioural Neurology.” All articles are open access articles distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Editorial Board

Jan O. Aasly, Norway
David C. Adamson, USA
Bert Aldenkamp, Netherlands
Andreas A. Argyriou, Greece
Giuliano Avanzini, Italy
Martin Bareš, Czech Republic
Michael E. Behen, USA
Giuseppe Bellelli, Italy
Félix Bermejo-Pareja, Spain
Ahmad Beydoun, Lebanon
Giuseppe Biagini, Italy
Frederic Blanc, France
Vittorio Calabrese, Italy
Lorys Castelli, Italy
Bernard S. Chang, USA
Ming-Jang Chiu, Taiwan
Leungwing Chu, Hong Kong
Steve Chung, USA
C. Constantoyannis, Greece
Gianluca Coppola, Italy
Louise Corben, Australia
Teresa Corona, Mexico
Joao Costa, Portugal
Yves Dauvilliers, France
Alpaslan Dedeoglu, USA
Ahmad R. Dehpour, Iran
Massimiliano Di Filippo, Italy
Hongxin Dong, USA
Dirk Dressler, Germany
Enzo Emanuele, Italy
Luigi Ferini-Strambi, Italy
Pierre O. Fernagut, France
Massimo Filippi, Italy
Elena García-Martín, Spain
Othman Ghribi, USA
Anna R. Giovagnoli, Italy
Cheng-Xin Gong, USA
Sheffali Gulati, India
Roberto Hornero, Spain
Shiro Imagama, Japan

Tauheed Ishrat, USA
Marjan Jahanshahi, UK
Xiaofeng Jia, USA
Zhifan Jia, China
Jiyao Jiang, China
Matthias Kliegel, Switzerland
Péter Klivényi, Hungary
Susanne Knake, Germany
Norbert Kovács, Hungary
Michael J. Larson, USA
Didier Le Gall, France
Cheng-Hsien Lu, Taiwan
Fabio M. Macciardi, USA
Camillo Marra, Italy
Michel Mittelbronn, Germany
Mohammad R. Mohammadi, Iran
Vincent C. Mok, China
Paula I. Moreira, Portugal
Kirsten R. Müller-Vahl, Germany
Shin-ichi Muramatsu, Japan
Ichiro Nakano, USA
Brian Neville, UK
Antonio Orlacchio, Italy
Ivo Paclt, Czech Republic
Deric M. Park, USA
Jesús Pastor, Spain
Nikunj Patel, USA
Paola Piccini, UK
Barbara Picconi, Italy
Olivier Piguet, Australia
Yolande Pijnenburg, Netherlands
Antonio Pisani, Italy
Francesco Pisani, Italy
Angelo Poletti, Italy
Gail J. Pyne-Geithman, USA
João Quevedo, Brazil
Matthias W. Riepe, Germany
Nikolaos K. Robakis, USA
José J. Rodríguez, Spain
Veit Roessner, Germany

Troy T. Rohn, USA
Andrea Romigi, Italy
Alessandro Rossi, Italy
Simone Rossi, Italy
Guido Rubboli, Denmark
Martino Ruggieri, Italy
Evzen Ruzicka, Czech Republic
Hoon Ryu, USA
Martin J. Sadowski, USA
Diego Salas-Gonzalez, Spain
Davide Seripa, Italy
Shyam S. Sharma, India
Riyi Shi, USA
Gabriella Silvestri, Italy
Mahendra P. Singh, India
Tomasz M. Sobów, Poland
Gianfranco Spalletta, Italy
Till Sprenger, USA
Fabrizio Stocchi, Italy
Nicola Tambasco, Italy
Wai-Kwong Tang, Hong Kong
Cristiano Termine, Italy
Masahiro Toda, Japan
Manjari Tripathi, India
Luigi Trojano, Italy
Andrea Truini, Italy
Charles J. Vecht, France
Annalena Venneri, UK
Di Lazzaro Vincenzo, Italy
Divya Vohora, India
Jian Z. Wang, China
Oliver Wirths, Germany
Karsten Witt, Germany
Masahito Yamada, Japan
Mario Zappia, Italy
Theresa A. Zesiewicz, USA
John H. Zhang, USA
Dan Ziegler, Germany

Contents

Music in the Brain: From Listening to Playing, Masayuki Satoh, Stefan Evers, Shinichi Furuya, and Kentaro Ono
Volume 2015, Article ID 927274, 2 pages

Music Engineering as a Novel Strategy for Enhancing Music Enjoyment in the Cochlear Implant Recipient, Gavriel D. Kohlberg, Dean M. Mancuso, Divya A. Chari, and Anil K. Lalwani
Volume 2015, Article ID 829680, 7 pages

Musical Sequence Learning and EEG Correlates of Audiomotor Processing, Matt D. Schalles and Jaime A. Pineda
Volume 2015, Article ID 638202, 11 pages

Sound Richness of Music Might Be Mediated by Color Perception: A PET Study, Masayuki Satoh, Ken Nagata, and Hidekazu Tomimoto
Volume 2015, Article ID 241804, 10 pages

Music Perception Influences Language Acquisition: Melodic and Rhythmic-Melodic Perception in Children with Specific Language Impairment, Stephan Sallat and Sebastian Jentschke
Volume 2015, Article ID 606470, 10 pages

The Influence of Music on Prefrontal Cortex during Episodic Encoding and Retrieval of Verbal Information: A Multichannel fNIRS Study, Laura Ferreri, Emmanuel Bigand, Patrick Bard, and Aurélie Bugaïska
Volume 2015, Article ID 707625, 11 pages

Interaction of Musicianship and Aging: A Comparison of Cortical Auditory Evoked Potentials, Jennifer L. O'Brien, Dee A. Nikjeh, and Jennifer J. Lister
Volume 2015, Article ID 545917, 12 pages

Inferior Frontal Gyrus Activation Underlies the Perception of Emotions, While Precuneus Activation Underlies the Feeling of Emotions during Music Listening, Ken-ichi Tabei
Volume 2015, Article ID 529043, 6 pages

The Mismatch Negativity: An Indicator of Perception of Regularities in Music, Xide Yu, Tao Liu, and Dingguo Gao
Volume 2015, Article ID 469508, 12 pages

A Voxel-Based Morphometry Study of the Brain of University Students Majoring in Music and Nonmusic Disciplines, Kanako Sato, Eiji Kirino, and Shoji Tanaka
Volume 2015, Article ID 274919, 9 pages

Melodic Contour Training and Its Effect on Speech in Noise, Consonant Discrimination, and Prosody Perception for Cochlear Implant Recipients, Chi Yhun Lo, Catherine M. McMahon, Valerie Looi, and William F. Thompson
Volume 2015, Article ID 352869, 10 pages

Editorial

Music in the Brain: From Listening to Playing

Masayuki Satoh,¹ Stefan Evers,² Shinichi Furuya,³ and Kentaro Ono⁴

¹Department of Dementia Prevention and Therapeutics, Graduate School of Medicine at the Mie University, 2-174 Edobashi, Tsu, Mie 514-8507, Japan

²Department of Neurology, University of Münster and Department of Neurology, Krankenhaus Lindenbrunn, Lindenbrunn 1, 31863 Coppenbrügge, Germany

³Department of Informatics and Communication Sciences, Sophia University, 4-4 Chiyoda-ku, Tokyo 1020081, Japan

⁴Human Brain Research Center, Kyoto University, 54 Shogoin Kawaramachi, Sakyo-ku, Kyoto 606-8507, Japan

Correspondence should be addressed to Masayuki Satoh; bruckner@clin.medic.mie-u.ac.jp

Received 11 August 2015; Accepted 11 August 2015

Copyright © 2015 Masayuki Satoh et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Music is one of the most primary abilities of human. The first case of amusia, which meant the impairment of musical ability due to the brain damage, was reported only several years after Broca's first report of aphasia. Since then, case studies were the main strategy to investigate human brain functions for a long time. Compared to remarkable progression of the study about language, music processing in the brain still remains to be clarified. But, over the last few decades, a considerable number of studies have been made on this issue, especially using the neuroimaging techniques. Now is the time to overview and integrate these findings and propose the possible application in clinical and educational situations.

The articles in this special issue include reviews and research studies focusing on cognitive, learning, developmental, and therapeutic aspects of music. The backgrounds of the authors are various: basic neuroscience, clinical neurology, pediatrics, geriatrics, and engineering. The methods include electroencephalography (EEG), positron emission tomography (PET), functional MRI (fMRI), and near infrared spectroscopy (NIRS). Based on the various points of view, the present issue shows some windows through which we can investigate the relationship between music and brain.

Several activation studies were reported using NIRS, PET, and fMRI. Mismatch Negativity (MMN) is a deviation-specific component of the auditory event-related potential (ERP), which detects a deviation between a sound and an internal representation. In the review article of "The Mismatch Negativity (MMN): An Indicator of Perception of Regularities in Music," X. Yu et al. compared the differences

of MMN features between musicians and nonmusicians, followed by a discussion of the potential roles of the training effect and the natural exposure in MMN. They pointed out some open questions and emphasized the importance to combine MMN with other experimental paradigms.

Another article related to MMN was presented by J. O'Brien et al. In the paper entitled "Interaction of Musician-ship and Aging: A Comparison of Cortical Auditory Evoked Potentials," they investigated whether the beneficial auditory neural effects of early music training persist throughout life and influence age-related changes in neurophysiological processing of sound. The results showed that MMN and P3a latencies for harmonic tone deviances were earlier for older musicians than older nonmusicians. These findings support beneficial influences of musicianship on central auditory function.

In the article "Musical Sequence Learning and EEG Correlates of Audiomotor Processing," M. D. Schalles and J. A. Pineda studied the functional connection between motor and auditory systems during musical performance. They recorded EEG of subjects listening to clips from a song they learned to play, a transposed version of that song, and a control song with different melody and notes from the learned song. EEG power of the beta band over sensorimotor scalp showed increased suppression for the learned song, a moderate level of suppression for the transposed song, and no suppression for the control song. They interpret these findings as a support that the motor system not only was active during the covert perception of music one can play but

also showed sensitivity to changes in pitch when the relative sequence of notes is preserved.

Using fMRI, K. Tabei investigated brain areas involved in perceiving and feeling emotions during music listening. In the article of “Inferior Frontal Gyrus Activation Underlies the Perception of Emotions, while Precuneus Activation Underlies the Feeling of Emotions during Music Listening,” he showed that cortical areas including the prefrontal, auditory, cingulate, and posterior parietal cortices were consistently activated by the perceived and felt emotional tasks. The precuneus showed greater activity during the felt emotion task than during a passive listening task. He suggested that the bilateral inferior frontal gyri and the precuneus are important areas for the perception of the emotional content of music as well as for the emotional response evoked in the listener.

Because of the silence during the operation, PET is suitable to study music recognition. In the PET activation study of “Sound Richness of Music Might Be Mediated by Color Perception: A PET Study” by M. Satoh et al., the cognitive processing of the perception of sound richness was investigated. The posterior portion of the inferior temporal gyrus, including the lateral occipital complex (LOC) and fusiform gyrus, was activated; so they concluded that certain association cortices may represent centers of multisensory integration in terms of both vision and audition.

The anatomical difference of the brain between musician and nonmusicians was reported in the article “A Voxel-Based Morphometry Study of the Brain of University Students Majoring in Music and Nonmusic Disciplines” by K. Sato et al. Voxel-based morphometry (VBM) uses a voxel-wise analysis method for determining focal differences in volume. They showed that the music expert group had the largest gray matter volumes in the right inferior frontal gyrus (BA 44), left middle occipital gyrus (BA 18), and bilateral lingual gyrus. These differences are considered to be caused by neuroplasticity during long and continuous musical training periods.

Functional NIRS (fNIRS) shows the activation of brain regions through decrease in oxyhemoglobin (O₂Hb) and the increase in deoxyhemoglobin (HHb) concentration. Using fNIRS, L. Ferreri et al. aimed to extend previous findings, in the article “The Influence of Music on Prefrontal Cortex during Episodic Encoding and Retrieval of Verbal Information: A Multichannel fNIRS Study,” by monitoring the entire lateral prefrontal cortex (PFC) during both encoding and retrieval of verbal material. The results showed that music provided a less-demanding way of modulating both episodic encoding and retrieval, with a general prefrontal decreased activity under the music versus silence condition. This suggests that music-related memory processes rely on specific neural mechanisms and that music can positively influence both episodic encoding and retrieval of verbal information.

Two articles are dedicated to the research of cochlear implant recipients. Cochlear implants (CIs) are surgically implanted hearing devices that enable the perception of sound for most persons diagnosed with severe to profound deafness. In the article entitled “Melodic Contour Training and Its Effect on Speech in Noise, Consonant Discrimination,

and Prosody Perception for Cochlear Implant Recipients,” C. Y. Lo et al. described the effect of two melodic contour training programs and their relative efficacy as measured on a number of speech perception tasks. Results indicated that there were some benefits for speech perception tasks for CI recipients after melodic contour training. Specifically, consonant perception in quiet and question/statement prosody was improved.

In the article of “Music Engineering as a Strategy for Enhancing Music Enjoyment in the Cochlear Implant Recipient”, G. Kohlberg et al. tested whether reengineering music to reduce its complexity could enhance the listening experience for the CI listener. As a result, compared to the original song, modified versions containing only 1–3 instruments were less enjoyable to the normal hearing (NH) listeners but more enjoyable to the CI listeners and the NH listeners with CI simulation. They concluded that, in addition to improvements in software and hardware, engineering music specifically for the CI listener may be an alternative means to enhance their listening experience.

Language and music share many properties, with a particularly strong overlap for prosody. Music perception skills (melodic and rhythmic-melodic perception and melody recognition) in a group of children with Specific Language Impairment (SLI) were shown in the article of “Music Perception Influences Language Acquisition: Melodic and Rhythmic-Melodic Perception in Children with Specific Language Impairment” written by S. Sallat et al. Children with SLI performed in most tasks below their age level, and these data strengthened the view of a strong relation between language acquisition and music processing.

Even though the accomplishment of each study is limited from the viewpoint of the field of neuroscience, we confine that the border between unknown and already known is spreading step by step. We hope that this special issue will shed light on major developments in the area of music processing in the brain and attract attention by the scientific community to pursue further investigations leading to the implementation of music in clinical, educational, and therapeutic situations.

Acknowledgments

We would like to express our appreciation to all the authors for their informative contributions and the reviewers for their support and constructive critiques in making this special issue possible.

*Masayuki Satoh
Stefan Evers
Shinichi Furuya
Kentaro Ono*

Research Article

Music Engineering as a Novel Strategy for Enhancing Music Enjoyment in the Cochlear Implant Recipient

Gavriel D. Kohlberg, Dean M. Mancuso, Divya A. Chari, and Anil K. Lalwani

Columbia University Cochlear Implant Center, Department of Otolaryngology/Head and Neck Surgery, Columbia University College of Physicians and Surgeons, 180 Fort Washington Avenue, Harkness Pavilion 8th Floor, New York, NY 10032, USA

Correspondence should be addressed to Anil K. Lalwani; anil.lalwani@columbia.edu

Received 20 January 2015; Revised 10 March 2015; Accepted 12 March 2015

Academic Editor: Masayuki Satoh

Copyright © 2015 Gavriel D. Kohlberg et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Objective. Enjoyment of music remains an elusive goal following cochlear implantation. We test the hypothesis that reengineering music to reduce its complexity can enhance the listening experience for the cochlear implant (CI) listener. *Methods.* Normal hearing (NH) adults ($N = 16$) and CI listeners ($N = 9$) evaluated a piece of country music on three enjoyment modalities: pleasantness, musicality, and naturalness. Participants listened to the original version along with 20 modified, less complex, versions created by including subsets of the musical instruments from the original song. NH participants listened to the segments both with and without CI simulation processing. *Results.* Compared to the original song, modified versions containing only 1–3 instruments were less enjoyable to the NH listeners but more enjoyable to the CI listeners and the NH listeners with CI simulation. Excluding vocals and including rhythmic instruments improved enjoyment for NH listeners with CI simulation but made no difference for CI listeners. *Conclusions.* Reengineering a piece of music to reduce its complexity has the potential to enhance music enjoyment for the cochlear implantee. Thus, in addition to improvements in software and hardware, engineering music specifically for the CI listener may be an alternative means to enhance their listening experience.

1. Introduction

The cochlear implant (CI) restores hearing to deafened individuals. Speech discrimination among the postlingually deafened CI users usually exceeds 65% [1]. Despite excellent performance on speech discrimination, enjoyment of music among CI listeners remains poor and has been attributed to decreased music perception. Music perception, the perception of pitch, melody, harmony, rhythm, and timbre, is greatly impaired in cochlear implantees [2–5]. Contributing factors responsible for diminished music perception include low resolution and skewed mapping of transmitted frequencies through the CI to the auditory cortex, difficulty perceiving spectral components individually, and deficits with higher perceptual integration tasks such as auditory stream segregation [2]. Limitations of cochlear implant hardware, sound processing software, and auditory nerve degeneration all play a possible role in signal degradation. Due to these factors, complex music signals are poorly perceived and consequently poorly enjoyed by CI listeners.

A large body of literature has shown that CI listeners have poor perception of musical elements [2–5]. In a test of 42 CI users and 10 normal hearing (NH) adults, CI users were found to be significantly worse than NH adults at pitch perception as well as both melody and timbre recognition [5]. In a study of 9 adults who underwent cochlear implantation, pitch perception was found to be worse after implantation than immediately before implantation [3].

Enjoyment of music has also been shown to decrease following cochlear implantation. In approaching the issue of poor musical enjoyment in CI users, studies have taken two main approaches. In the first, attempts have been made to analyze how varying CI devices and strategies affect music enjoyment. In the second, CI users' music enjoyment has been analyzed across different genres and pieces of music with varying complexities. In most studies, music enjoyment has been assessed by variations of a bipolar visual analog scale (VAS) or a discrete 10-point scale [6–9]. Questionnaire studies of CI users have found a significant decrease in music enjoyment as well as time devoted to listening to music

compared with before onset of deafness. Many CI listeners could not enjoy music at all, stating that music did not sound natural. Despite the overall decrease in enjoyment, 38–73.6% of CI users still listened to music and 30.2%–37% stated they would undergo implantation simply to be able to listen to music [10–13].

Attempts to improve music enjoyment have examined cochlear implant hardware and software strategies with mixed results [14]. A study in CI users with a MED-EL device on one side and a cochlear nucleus on the other found no significant difference in music enjoyment when either device type was used exclusively [6]. Similarly, two other studies comparing fine structure processing (FSP) strategy to continuous interleaved sampling (CIS) strategy found no difference in music enjoyment among CI listeners [8, 15].

Given the significant reduction in the population of auditory neurons available to relay complex musical signals in severe to profoundly deafened patients undergoing implantation, it may be unrealistic to expect normal music perception despite further enhancement in CI software and hardware. In the absence of significant improvement in perception, music may have to be specially engineered for the CI listener to enhance its enjoyment. In exploration of this concept, we investigate the features of music that impact on its enjoyment (not perception) in normal hearing individuals and cochlear implantees.

Music perception among CI listeners has been studied extensively [3–5]. Investigating perception of music among implantees is important in assessing the ability of CI software and hardware to restore musical perception and may prove beneficial in guiding our understanding of music enjoyment among CI listeners. It is imperative to note, however, that music perception and music enjoyment are fundamentally different and may not necessarily correlate. For example, Alexander et al. in their study of music perception and enjoyment found that, despite significantly poorer performance on music perception, the enjoyment of music among cochlear implantees was comparable to normal hearing individuals [16]. This disconnect between enjoyment and perception is similar to a cochlear implantee performing well on speech perception testing but finding the quality of speech sound unpleasant or unnatural. Thus, music perception cannot be substituted for music enjoyment when investigating the types and characteristics of music that are enjoyable to CI listeners; consequently, the enjoyment of music must be studied directly.

2. Materials and Methods

2.1. Participants. After obtaining Columbia University Medical Center Institutional Review Board approval, we analyzed music enjoyment in 16 normal hearing (NH) individuals and 9 cochlear implantees. After obtaining consent, NH subjects underwent an audiologic evaluation, including evaluation of pure tone thresholds, speech discrimination, and otoscopic evaluation. Inclusion criteria included English speaking, 18 years of age or older, no history of hearing loss, and pure tone audiometric thresholds less than or equal to 20-decibel

hearing loss in both ears at all tested frequencies. Inclusion criteria for cochlear implantees included English speaking, age over 18 years, and postlingual deafness.

2.2. Music Piece. Each subject listened to and evaluated an original and 20 modified versions of a 20-second piece of country music (“Milk Cow Blues” performed by Angela Thomas Wade) recorded in multitrack format. The original song segment included multiple musical elements: female vocals; three melodic instruments (guitar, piano, and fiddle); rhythmic drums including the snare. Each of the 20 modified versions was comprised of a unique subset of musical elements of the original song. Five modified versions were comprised of a single musical element: vocals; piano; guitar; fiddle; and the snare. Seven modified versions were obtained by combining two musical elements: vocals and guitar; vocals and snare; vocals and piano; vocals and fiddle; snare and guitar; snare and piano; and snare and fiddle. Four modified versions were obtained by combining three musical elements: vocals, snare, and guitar; vocals, snare, and piano; vocals, snare, and fiddle; guitar, piano, and fiddle. Two modified versions were comprised of four musical elements: piano, guitar, fiddle, and vocals; snare, kick, overhead, and tom drums. One version was comprised of five musical elements: snare, kick, overhead, tom drums, and vocals. One version with all of the musical elements except for the vocals contained 10 musical elements.

2.3. Music Presentation to NH Listeners. The subjects listened to the musical samples in a sound proof booth (IAC Acoustics, New York, NY). Participants adjusted the volume according to their preference. Participants rated each of the music segments using a visual analog scale implemented using MATLAB version 7.1 (Mathworks, Natick, MA) on a MacBook Air (Apple, Cupertino, CA). Subjects were instructed to rate each sample on a scale from 0 to 10 in each of the following categories: “pleasant and unpleasant,” “sounds like music and does not sound like music,” and “natural and unnatural,” with higher numeric scores corresponding to higher levels of pleasantness, musicality, and naturalness. Subjects were presented the music samples in a random order.

Each of the music segments was then processed through CI simulation software provided by Advanced Bionics Corp., using MATLAB version 7.1 [17]. These music segments with CI simulation were then presented and rated by the NH listeners in the same fashion as the music segments without CI simulation.

2.4. Music Presentation to CI Listeners. Listening took place in a sound proof booth. Music segments were presented in a free field at 60 dB SPL. Participants rated each music segment in a similar fashion as the NH listeners as described above.

2.5. Cochlear Implant Simulation Software. Half of the music segments were presented to the NH listeners after being processed through a CI simulation. The CI simulation software aimed to simulate a CI listening experience by modeling both the sound processing that occurs to an incoming sound

signal in the speech processor of the cochlear implant and the spread of excitation related to electrical stimulation in the cochlea [17]. In particular, CI sound processing was modeled by filtering the sound signal into 15 logarithmically spaced channels covering the range from 350 to 5500 Hz. The envelope signal was computed for each channel and used to modulate a noise band. The noise band for each channel was chosen to have the center frequency corresponding to the center frequency of the channel and to simulate appropriate spread in the cochlea. Litvak et al. [17] varied the amount of spread in the noise band in dB/octave and determined that 20 dB/octave appeared as an accurate model of spread that occurs for the average CI listener. In addition, they showed that this simulation matched NH listeners' speech discrimination on vowel recognition.

2.6. Statistical Analysis. *t*-tests were used to compare enjoyment between subsets of musical segments within the NH listeners and the CI listeners. In addition, analysis of variance (ANOVA) was applied to compare the mean enjoyment of the original version of the song and the mean enjoyment of modified versions comprised of 1, 2, or 3 instruments. Enjoyment comparisons were analyzed separately for pleasantness, musicality, and naturalness scales. Statistical significance was considered for $P < 0.05$ for both the *t*-tests and ANOVA statistical tests. Of note, all 4 of the modified versions of the song with 4 or more musical elements were excluded in the linear regression results. Three of these modified versions contained percussion elements that were not present in any of the other modified versions, compromising the generalizability of any results related to these modified versions. The fourth excluded modified version was the only modified version remaining with 4 musical elements, limiting the relevance of this category. Of note, all statistically significant linear regression findings hold when including these 4 excluded modified versions.

3. Results

The mean age of the 16 NH participants was 29 years (25–33). Eight were female and eight were male. Eighty-eight percent spoke English as their first language. Speech discrimination was 96–100% on 25 spondee words. One participant had a history of pressure equalization tube placement as a child.

The mean age of the cochlear implantees was 54 years (26–74). Six were female and 3 were male. One hundred percent spoke English as their first language. Six were implanted with the Advanced Bionics HiRes 90K implant and 3 with the Advanced Bionics CII implant. All CI listeners used the Fidelity 120 sound processing strategy. The average time since implantation was 7 years (2–13). Three had bilateral cochlear implants. Speech discrimination ranged from 60 to 97% (mean 84.6%) on the AzBio Sentence list.

3.1. Enjoyment of the Original Music Sample. The original version of the song “Milk Cow Blues” was well liked by the NH listeners without CI simulation (pleasant, musical, and natural; 8.3, 8.8, and 8.4). NH listeners rated the original

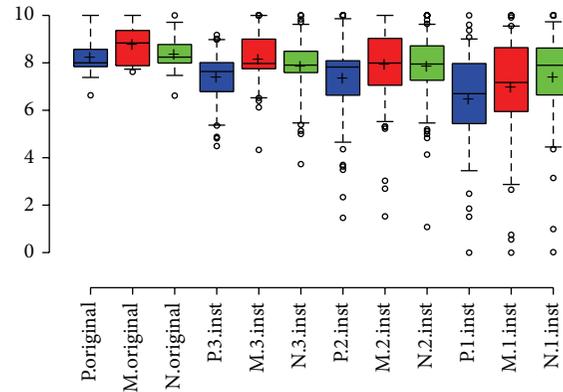


FIGURE 1: NH listeners without CI simulation enjoyment by number of instruments. NH listeners preferred the original music sample compared to modified segments comprised of a single instrument, two instruments, or three instruments. P: pleasant, N: natural, M: sounds like music, and Inst: instruments. Center lines show the medians, box limits indicate the 25th and 75th percentiles, whiskers extend to 5th and 95th percentiles, outliers are represented by dots, and crosses represent sample means.

version of the song as the 3rd most pleasant, 2nd most musical, and 5th most natural out of the 21 total segments (original plus 20 modified versions). There was no modified version of the song that was preferred by the NH listeners on all three enjoyment scales (pleasantness, musicality, and naturalness). With CI simulation, NH listeners rated the original version poorly (pleasant, musical, and natural; 0.69, 1.4, and 0.76) and rated the original version 16th most pleasant, 13th most musical, and 12th most natural. Ten of the 20 modified versions were preferred to the original on all three enjoyment scales by the NH listeners with CI simulation. The CI listeners rated the original version of the song very poorly (pleasant, musical, and natural; 5.4, 6.3, and 5.6) compared to the modified versions of the song ranking it as the least pleasant and the 2nd least musical and natural. CI listeners preferred 19 of the 20 modified versions to the original on all three enjoyment scales.

3.2. Comparison of the Original Music Sample to Modified Version with 1–3 Instruments

3.2.1. Preferences among NH Listeners without CI Simulation. Music enjoyment was universally greater for NH listeners for the original music segment compared to modified versions comprised of a single instrument, two instruments, or three instruments (Figure 1). Compared to the original sample, there was a relative reduction in pleasantness of 21.4%, 9.0%, and 9.5% for modified versions with a single instrument, two instruments, and three instruments, respectively ($P < 0.05$). The relative reduction in enjoyment of the original music sample compared to modified versions with 1–3 instruments for the three scales (pleasant, musical, and natural) is summarized in Table 1. A linear regression analysis of enjoyment with ANOVA found a significant difference in the means

TABLE 1: Relative reduction of enjoyment of the original music sample compared to modified versions containing 1–3 instruments among NH listeners without CI simulation.

Relative reduction in enjoyment	Original versus 1 instrument	Original versus 2 instruments	Original versus 3 instruments
Pleasant	21.4	9	9.5
Musical	18.5	7.5	7.2
Natural	11.6	6.1	6.2

Values in bold are for $P < 0.05$.

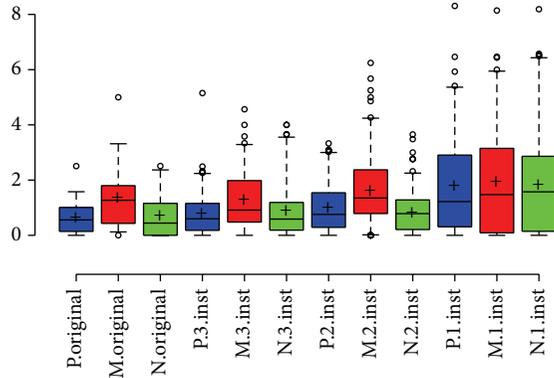


FIGURE 2: NH listeners with CI simulation enjoyment by number of instruments. NH listeners preferred modified segments comprised of a single instrument, two instruments, or three instruments compared to the original music sample. P: pleasant, N: natural, M: sounds like music, and Inst: instruments. Center lines show the medians, box limits indicate the 25th and 75th percentiles, whiskers extend to 5th and 95th percentiles, outliers are represented by dots, and crosses represent sample means.

for the original version and for modified versions with 1–3 instruments ($P < 0.01$ for pleasant, musical, and natural scales).

3.2.2. Preferences among NH Listeners with CI Simulation.

When listening with CI simulation, NH listeners rated the original music segment less enjoyable than modified versions comprised of one to three musical elements (Figure 2). Compared to the original music sample, modified versions with a single instrument were significantly more pleasant by 59.5% ($P < 0.05$). The relative increase in enjoyment for the modified versions with 1–3 instruments compared to the original song for NH listeners with CI simulation is summarized in Table 2. A linear regression analysis of enjoyment with ANOVA found a significant difference in the means for the original version and for modified versions with 1–3 instruments on the pleasant scale ($P = 0.001$) and natural scale ($P = 0.003$) but not the musical scale ($P = 0.09$).

3.2.3. Preferences among CI Listeners. CI listeners rated the original music segment less enjoyable than modified versions comprised of one to three musical elements (Figure 3). Compared to the original music sample, modified versions

TABLE 2: Relative increase of enjoyment of the original music sample compared to modified versions containing 1–3 instruments among NH listeners with CI simulation.

Relative increase in enjoyment	Original versus 1 instrument	Original versus 2 instruments	Original versus 3 instruments
Pleasant	59.5	33	15.6
Musical	15.1	14.4	88.7
Natural	51.7	7.9	15

Values in bold are for $P < 0.05$.

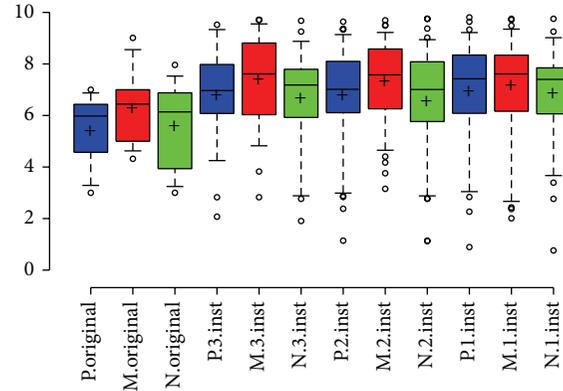


FIGURE 3: CI listeners enjoyment by number of instruments. CI listeners preferred modified segments comprised of a single instrument, two instruments, or three instruments compared to the original music sample. P: pleasant, N: natural, M: sounds like music, and Inst: instruments. Center lines show the medians, box limits indicate the 25th and 75th percentiles, whiskers extend to 5th and 95th percentiles, outliers are represented by dots, and crosses represent sample means.

with a single instrument and 3 instruments were significantly more pleasant by 21.9% and 18.3%, respectively ($P < 0.05$). The relative increase in enjoyment for the modified versions with 1–3 instruments compared to the original song for NH listeners with CI simulation is summarized in Table 3. A linear regression analysis of enjoyment with ANOVA found a significant difference in the means for the original version and for modified versions with 1–3 instruments on the pleasant scale ($P = 0.034$) but not the musical ($P = 0.22$) or natural scales ($P = 0.21$).

3.3. Comparison of Modified Versions with and without Vocals.

The original music sample contained prominent female vocals. Modified versions ($n = 10$) containing the vocal element of the original music sample were compared to modified versions ($n = 10$) that did not include the vocals. NH listeners without CI simulation found versions with the vocals significantly more enjoyable on all three scales, pleasant, musical, and natural ($P < 0.01$). In contrast, with CI simulation, NH listeners found versions with vocals significantly less enjoyable on all three scales ($P < 0.02$). For the CI listeners, there was virtually no difference in enjoyment between versions with and without vocals (mean

TABLE 3: Relative increase of enjoyment of the original music sample compared to modified versions containing 1–3 instruments among CI listeners.

Relative increase in enjoyment	Original versus 1 instrument	Original versus 2 instruments	Original versus 3 instruments
Pleasant	21.9	20.1	18.3
Musical	12.2	14.1	13.5
Natural	17.8	15	15.3

Values in bold are for $P < 0.05$.

pleasant, musical, and natural with and without vocals: 6.84 versus 6.85, 7.24 versus 7.44, and 6.66 versus 6.76, resp.).

3.4. Comparison of Modified Versions with Melodic Elements with and without Rhythmic Elements. Modified versions of the original song with melodic elements only (vocals, piano, fiddle, and guitar) were compared to versions that also included rhythmic drum instruments. All modified versions of the song were used in this comparison except for the version with only the snare and the version with only the four drum instruments, as these segments did not include any melodic elements. There was no difference in enjoyment on all three scales for NH listeners without CI simulation between versions with and without rhythmic instruments. NH listeners with CI simulation found the rhythmic versions significantly more enjoyable on all three scales ($P < 0.05$ for pleasant, musical, and natural). CI listeners found virtually no difference in enjoyment between versions with and without rhythmic instruments (mean pleasant, musical, and natural with and without rhythmic instruments: 6.80 versus 6.89, 7.21 versus 7.40, and 6.73 versus 6.70, resp.).

4. Discussion

In this study, we investigated features of music that impact on its enjoyment and tested the novel hypothesis that an original piece of music could be reengineered to make it more enjoyable for the cochlear implantee. To the best of our knowledge, no prior study has taken the approach of altering a specific piece of music to determine if it can be made more enjoyable for CI listeners. We modified a 20-second segment of a country music song by playing it with various subsets of the original music sample's musical elements. Modified versions contained various combinations of the vocals and instruments.

We found that NH listeners enjoyed the original music segment the most and rated the modified, less complex, versions less enjoyable. On the other hand, for the NH listeners with CI simulation, enjoyment increased significantly when the complexity of the original musical sample was reduced by limiting the number of elements in modified versions of the song to 1–3 elements. The extent of increase in enjoyment of these less musically complex modified music segments was even more pronounced among the CI listeners.

Studies on music perception have shown that CI listeners have severe difficulty identifying pitch, timbre, and melody,

the main qualities in vocals and melodic instruments such as guitar, piano, and fiddle [3, 5]. On the other hand, the musical quality best perceived by CI listeners has been found to be rhythm [4]. In our study we found that NH listeners with CI simulation preferred modified music samples that included rhythmic instruments compared to segments containing only melodic instruments. Interestingly, the CI listeners did not rate samples with and without rhythmic instruments differently. The impact of rhythmic instruments will need to be further studied with musical pieces containing greater numbers of rhythmic instruments.

The effect of vocals on music perception has been studied in CI listeners. A study of 87 CI listeners found that the presence of lyrics in pop and country songs significantly improves the CI listeners' ability to identify a melody [18]. In this study, we evaluated the effect of vocals on music enjoyment. We found that while enjoyment was greater for music with vocals for NH listeners without CI, it was significantly less for NH listeners with CI. CI listeners, on the other hand, rated modified segments with and without vocals exactly the same.

We found that for NH listeners with CI simulation enjoyment was significantly increased for modified segments without vocals and for modified segments with rhythmic instruments. While we did not find that music segments limited to particular qualities, such as only vocals, melodic, or rhythmic elements, were significantly more enjoyable to CI listeners, we did find that reducing the number of musical elements significantly increased music enjoyment both for the NH listeners through CI simulation and for the CI listeners.

Several studies have examined the relationship between music enjoyment and complexity of musical pieces in CI listeners. In a study of 15 CI users and 24 hearing aid (HA) users meeting audiologic criteria for CI, they were asked to rate their enjoyment of music played by a single instrument, solo instruments with background accompaniment, and ensembles. The study found that music played by a single instrument was more enjoyable than music played by multiple instruments to the CI and HA listeners when analyzed together. Of note, the music played by the single instrument, solo instruments with background accompaniment, and ensembles was not specified to be the same music segment [7]. A study of 20 NH listeners with CI simulation processing found that minimizing reverberation time increased music enjoyment [19].

Part of this study involved querying normal hearing listeners' enjoyment of music through a CI simulation validated for speech perception. There are limitations to using CI simulation as a proxy for CI listeners' enjoyment. In our study, relative enjoyment of modified music segments was different between the NH listeners with CI simulation and the CI listeners. A previous study with CI simulation by Wright and Uchanski also found significant differences between music perception and enjoyment in NH listeners with CI simulation and CI listeners [20]. While the study by Wright and Uchanski and this study did not show clear correspondence in music enjoyment between NH listeners with CI and CI listeners, there is reason to continue to attempt

simulation studies. There are advantages to conducting initial or exploratory studies with CI simulation in NH listeners. The CI listener population has significant variation in age, duration of deafness, etiology of deafness, musical background and training, and rehabilitative outcomes. The use of NH listeners of similar age, hearing and musical training avoids these limitations typical of implanted population. Continued study of NH listeners with CI simulation processing may lead to an accurate prediction of CI listener enjoyment, which would be of great benefit.

In this study only a single piece of music was studied. Further analysis of other genres of music and other pieces within the country music genre will need to be examined. Additionally, other methods for altering the complexity of music beyond including and excluding vocals and instruments need to be explored.

5. Conclusion

Musical enjoyment with or without CI is influenced by the complexity of the original music. Our study offers preliminary evidence that engineering of music to reduce the complexity of music has the potential to make listening more enjoyable for the CI listener. Thus, in addition to improvement in software and hardware, engineering of music specifically for a CI listener may be an important way to enhance his or her listening experience.

Conflict of Interests

Anil K. Lalwani serves in the Medical Advisory Board of Advanced Bionics Corporation. The authors have no other funding, financial relationships, or conflict of interests to disclose.

Acknowledgments

The authors would like to thank Leonid Litvak, Abhjilt Kulkarni, and Kanthaiiah Koka of Advanced Bionics Corporation for assistance in the project by providing the cochlear implant simulation software.

References

- [1] L. K. Holden, C. C. Finley, J. B. Firszt et al., "Factors affecting open-set word recognition in adults with cochlear implants," *Ear and Hearing*, vol. 34, no. 3, pp. 342–360, 2013.
- [2] C. J. Limb and A. T. Roy, "Technological, biological, and acoustical constraints to music perception in cochlear implant users," *Hearing Research*, vol. 308, pp. 13–26, 2014.
- [3] V. Looi, H. McDermott, C. McKay, and L. Hickson, "The effect of cochlear implantation on music perception by adults with usable pre-operative acoustic hearing," *International Journal of Audiology*, vol. 47, no. 5, pp. 257–268, 2008.
- [4] Y.-Y. Kong, R. Cruz, J. A. Jones, and F.-G. Zeng, "Music perception with temporal cues in acoustic and electric hearing," *Ear and Hearing*, vol. 25, no. 2, pp. 173–185, 2004.
- [5] R. Kang, G. L. Nimmons, W. Drennan et al., "Development and validation of the University of Washington clinical assessment of music perception test," *Ear and Hearing*, vol. 30, no. 4, pp. 411–418, 2009.
- [6] R. L. Harris, W. P. R. Gibson, M. Johnson, J. Brew, M. Bray, and C. Psarros, "Intra-individual assessment of speech and music perception in cochlear implant users with contralateral Cochlear and MED-EL systems," *Acta Oto-Laryngologica*, vol. 131, no. 12, pp. 1270–1278, 2011.
- [7] V. Looi, H. McDermott, C. McKay, and L. Hickson, "Comparisons of quality ratings for music by cochlear implant and hearing aid users," *Ear and Hearing*, vol. 28, no. 2, supplement, pp. 59S–61S, 2007.
- [8] V. Looi, P. Winter, I. Anderson, and C. Sucher, "A music quality rating test battery for cochlear implant users to compare the FSP and HDCIS strategies for music appreciation," *International Journal of Audiology*, vol. 50, no. 8, pp. 503–518, 2011.
- [9] K. Rosslau, K. N. Spreckelmeyer, H. Saalfeld, and M. Westhofen, "Emotional and analytic music perception in cochlear implant users after optimizing the speech processor," *Acta Oto-Laryngologica*, vol. 132, no. 1, pp. 64–71, 2012.
- [10] B. Philips, B. Vinck, E. De Vel et al., "Characteristics and determinants of music appreciation in adult CI users," *European Archives of Oto-Rhino-Laryngology*, vol. 269, no. 3, pp. 813–821, 2012.
- [11] L. Lassaletta, A. Castro, M. Bastarrica et al., "Does music perception have an impact on quality of life following cochlear implantation?" *Acta Oto-Laryngologica*, vol. 127, no. 7, pp. 682–686, 2007.
- [12] L. Lassaletta, A. Castro, M. Bastarrica et al., "Changes in listening habits and quality of musical sound after cochlear implantation," *Otolaryngology—Head and Neck Surgery*, vol. 138, no. 3, pp. 363–367, 2008.
- [13] G. Kohlberg, J. B. Spitzer, D. Mancuso, and A. K. Lalwani, "Does cochlear implantation restore music appreciation?" *The Laryngoscope*, vol. 124, no. 3, pp. 587–588, 2014.
- [14] W. R. Drennan and J. T. Rubinstein, "Music perception in cochlear implant users and its relationship with psychophysical capabilities," *Journal of Rehabilitation Research and Development*, vol. 45, no. 5, pp. 779–790, 2008.
- [15] L. Magnusson, "Comparison of the fine structure processing (FSP) strategy and the CIS strategy used in the MED-EL cochlear implant system: speech intelligibility and music sound quality," *International Journal of Audiology*, vol. 50, no. 4, pp. 279–287, 2011.
- [16] A. J. Alexander, L. Bartel, L. Friesen, D. Shipp, and J. Chen, "From fragments to the whole: a comparison between cochlear implant users and normal-hearing listeners in music perception and enjoyment," *Journal of Otolaryngology—Head & Neck Surgery*, vol. 40, no. 1, pp. 1–7, 2011.
- [17] L. M. Litvak, A. J. Spahr, A. A. Saoji, and G. Y. Fridman, "Relationship between perception of spectral ripple and speech recognition in cochlear implant and vocoder listeners," vol. 122, no. 2, pp. 982–991, 2007.
- [18] K. Gfeller, D. Jiang, J. J. Oleson et al., "The effects of musical and linguistic components in recognition of real-world musical excerpts by cochlear implant recipients and normal-hearing adults," *Journal of Music Therapy*, vol. 49, no. 1, pp. 68–101, 2012.
- [19] M. V. Certo, G. D. Kohlberg, D. A. Chari, D. M. Mancuso, and A. K. Lalwani, "Reverberation time influences musical enjoyment with cochlear implants," *Otology & neurotology*, vol. 36, no. 2, pp. e46–e50, 2010.

- [20] R. Wright and R. M. Uchanski, "Music perception and appraisal: cochlear implant users and simulated cochlear implant listening," *Journal of the American Academy of Audiology*, vol. 23, no. 5, pp. 350–365, 2012.

Research Article

Musical Sequence Learning and EEG Correlates of Audiomotor Processing

Matt D. Schalles^{1,2} and Jaime A. Pineda^{1,3}

¹Department of Cognitive Science, University of California San Diego, La Jolla, CA 92093, USA

²VA Northern California Health Care System, Martinez, CA 94553, USA

³Department of Neuroscience, University of California San Diego, La Jolla, CA 92093, USA

Correspondence should be addressed to Matt D. Schalles; matt.schalles@gmail.com

Received 24 April 2015; Revised 8 July 2015; Accepted 14 July 2015

Academic Editor: Shinichi Furuya

Copyright © 2015 M. D. Schalles and J. A. Pineda. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Our motor and auditory systems are functionally connected during musical performance, and functional imaging suggests that the association is strong enough that passive music listening can engage the motor system. As predictive coding constrains movement sequence selections, could the motor system contribute to sequential processing of musical passages? If this is the case, then we hypothesized that the motor system should respond preferentially to passages of music that contain similar sequential information, even if other aspects of music, such as the absolute pitch, have been altered. We trained piano naive subjects with a learn-to-play-by-ear paradigm, to play a simple melodic sequence over five days. After training, we recorded EEG of subjects listening to the song they learned to play, a transposed version of that song, and a control song with different notes and sequence from the learned song. Beta band power over sensorimotor scalp showed increased suppression for the learned song, a moderate level of suppression for the transposed song, and no suppression for the control song. As beta power is associated with attention and motor processing, we interpret this as support of the motor system's activity during covert perception of music one can play and similar musical sequences.

1. Introduction

The performance of music recruits and synchronizes many neural systems, integrating motor output through auditory, somatosensory, and oftentimes visual input. Listening to sounds can easily stimulate the motor system to act in the form of head nodding, foot tapping, and dancing. Functional imaging of silent piano performance and passive listening to piano song reveals shared recruitment of auditory and premotor cortices [1–3], for pianist and nonpianist alike. Activity in these cortical regions is also modulated in task dependent measures. Subjects engaged in tapping along with simple beats exhibited simultaneous increases of blood flow to auditory and dorsal premotor cortices when the complexity of tapping beat increased [4]. Zatorre [5] proposed extending the definition of the auditory system to include interactions with other cognitive systems, such as motor and multisensory networks, to account for the distributed nature

of cortical responses during auditory and musical perception. A growing body of evidence suggests that neural representations of movements and sounds may become linked through mechanisms such as Hebbian learning or as an emergent property of an auditory-motor loop [6]. The audiomotor system is adaptable and exhibits associations of novel sound-action pairings on short [7] and intermediate [8] time scales. If the motor system is routinely recruited for musical or auditory processing, what might it contribute as part of a distributed auditory network? Or to put it another way, what aspects of musical experience might be explained by a distributed audiomotor system which are not explained by an auditory system without motor contributions?

The audiomotor system is sensitive to rhythmic aspects of music such as timing and metrical complexity [9, 10] and also melodic information, both in terms of timbre of instruments [11] and pitch of melodic notes associated with motor actions [8]. Tracking metrical and melodic information requires

accumulation of auditory information over time, which is used to generate expectations of future notes at precise time intervals. Neural processes track musical progressions and respond to violations of melodic or harmonic expectations with similar latency as semantic violations in language [12]. This suggests that we might process music as a hierarchically organized sequence of information, similar to language and motor programs [13, 14].

In the present study we asked whether the audiomotor system might be sensitive to sequential ordering of information in a musical passage, such that it could help generate top-down predictions for incoming auditory stimuli. Previous hypotheses posit that the motor system might contribute top-down predictive information in the form of a metrical grid [10, 15] or predicting the occurrence of a rhythmic beat or pulse [6] in music. This model is supported by *in vivo* multiunit recordings in monkeys during a visuomotor task describing initial feedforward communication from sensory to frontal cortices, followed by sustained feedback from frontal to sensory cortices [16]. While visual and auditory systems have different functional and anatomical connections with the motor system, it may be reasonable to assume that this feedforward and feedback dynamic may be representative of a supramodal sensorimotor loop.

As the audiomotor system is sensitive to both rhythmic and melodic aspects of music, we focused our present work on the sequencing of melodic information for two primary reasons. The first is that previous imaging work [8] demonstrated simultaneous activation of superior temporal auditory and premotor cortices during passive listening to a song that subjects had recently learned to play. Using a similar learn-to-play-by-ear task, we hypothesized that we could observe audiomotor engagement during passive listening using EEG measures that are sensitive to visuomotor engagement. And second, a melodic sequence affords an ideal manipulation of pitch information, while controlling for relative sequential information via use of a musical transposition of key. We hypothesized that if the audiomotor system is sensitive to melodic sequential information in the form of relative pitch intervals, then shifting the key or absolute pitch of a melodic sequence would still contain enough information associated with a motor sequence to engage the audiomotor system during passive listening.

Lahav et al. [8] report that the premotor and inferior parietal cortices respond preferentially when listening to a melodic passage that a listener knows how to play on the piano and exhibit little to no response when listening to a song unrelated to the one they learned how to play. Additionally, if subjects listen to a novel song that is composed from the same note set as the melody they learned to play, they exhibit reduced activation of the premotor and inferior parietal regions relative to listening to the learned song. The similarity was explained as a pitch-motion matching system, such that a single key press elicits a single sound and hearing the sound could trigger the association with the single motor act. An alternative explanation could describe the difference between learned song and novel song with learned notes as a difference in sequential information leading to reduced activity in motor planning areas.

In the present study, we asked whether the motor system is sensitive to preservation of the sequential ordering of musical information, even when the pitch information is altered. We used the piano ear-learning task developed by Lahav et al. [8, 17] and tested pitch-recognition-production matching before and after training. For the posttraining audiomotor system engagement, our work builds on past reports by using EEG measures of motor and sensorimotor system engagement [18], which could corroborate fMRI findings, as both BOLD signal increases and mu and beta band suppression in EEG (described below) are thought to index increases in cortical activity. If EEG is a good measure for audiomotor processing over the course of a musical phrase, it could encourage future study building on models of responses to discrete notes [19].

There is a strong history of EEG measures of motor system engagement during movement and the observation of movement. Reports from the late 1940s through the late 1970s describe mu power (8–13 Hz) over sensorimotor cortex as decreasing from a resting state during both actions and observation of actions, as discussed in [20–22]. The mu rhythm is suppressed during movement observation [23], performance of an action [24], and observation of object directed actions [25]. The mu rhythm is also suppressed during motor imagery [26]. The mu rhythm itself shares frequency properties with other brain rhythms, such as occipital alpha, but its intrinsic activity is functionally distinct, and source estimates localize it to the bilateral sensorimotor cortex surrounding the central sulcus [27]. The mu rhythm may be a good index of audiomotor processing as its suppression relative to baseline is associated with movement sounds, and combined sight and sound of actions suppress mu greater than either sensory input alone [28]. Mu rhythms as well as beta rhythms (15–30 Hz) are suppressed prior to a sound action (such as tapping on a drum) and exhibit rebound enhancement immediately after that action, whether the action is performed, observed, or heard [29]. Mu and beta both show increased phase coherence between motor, somatosensory, and auditory cortices when subjects synchronize movements to rhythmic sounds [30]. Additionally, beta rhythms appear to entrain to rhythmic sounds [31, 32] in auditory cortices and also exhibit similar patterns of suppression and enhancement for both listening and tapping [30] over motor areas. The synthesis of these studies indicates a similar neural process that is active during movement and listening, which is observed over the sensorimotor cortex. Additionally, sounds that have no clear movement association exhibit higher amounts of mu desynchronization after watching a video that associates that sound with a clear movement [7]. The mu rhythm also desynchronizes when expert pianists read sheet music [33]. Reading music is an activity that associates visual input, with a particular action, and becomes reinforced with auditory feedback.

We predicted that the mu and beta rhythms recorded from sensorimotor scalp would be suppressed maximally when a subject listens to a melody s/he knows how to play and fail to suppress or even show enhancement when listening to a melody unrelated to the learned song. Additionally, we predicted that listening to a transposed version of the learned

TABLE 1: Self-reported demographics for musical experience in subjects with past musical exposure.

Age	Gender	Instrument(s)	Years of training	Currently playing	Hours per week
18	M	Saxophone	4	No	0
19	F	Guitar, flute	1	No	0
19	M	Guitar, vocals	4	Yes	1.5
20	F	Voice, viola	5	No	0
20	F	Flute	2	No	0
20	F	Clarinet, saxophone	7.5	No	0
20	M	Violin	2	No	0
20	M	Drum set, tabla	6	Yes	2

melody would also elicit suppression; however the response might be attenuated relative to the learned melody itself, as previous findings indicate motor sensitivity to absolute pitch information, which is altered in the transposition.

2. Material and Methods

2.1. Participants. 16 undergraduate students (nine female, mean age 19.9 years, 15 right handed), from the University of California, San Diego, completed the experiment in exchange for a combination of monetary compensation and course credit. One additional subject (female, 22 years old, right handed) completed the training and behavioral experiments, but not the electrophysiological component, and her data are included in behavioral measures. Subjects' handedness was self-reported, and additionally subjects were screened for head trauma, the use of psychiatric medication, and piano experience. In previous use of this task, only nonmusicians were included, but we included subjects with prior musical experience ($n = 8$), as long as all were inexperienced with the piano keyboard. We surveyed the number and type of instruments played, years spent playing, and whether subjects were currently playing music. Results of this survey are available in Table 1. All subjects were able to detect pure tones ranging between 250 Hz and 8 kHz at 30 db in both their right and their left ears. Subjects signed consent for procedures that were approved by the UCSD Institutional Review Board.

2.2. Song Stimuli. The same training song as described in Lahav et al. [8, 17] was implemented for the current study. For all the songs, synthesized backing instruments, guitar, bass, and drums, were composed following the score provided in [17] using Sonar Cakewalk Music Studio v4. The songs were each eight measures long and had a duration of 24 seconds at 80 beats per minute. The melodic line for each song was voiced by synthesized piano. The melody for each song was 15 notes long and comprised of a set of five notes (F-G-A-Bb-C), one for each finger on the playing hand. The transposed melody preserved the relative intervals between the notes in the learned song but shifted them into a set of notes one tritone, or half-octave, higher (B-Db-Eb-E-Gb) that did not overlap with the learned song note set. The control song was comprised of the same note set as the transposed song. The notes in the control song were arranged in a different sequence from the trained and transposed melodies while

preserving the same total length of note duration and total changes in pitch height over the 8 bars. Backing rhythm guitar and bass lines in the transposed and control songs were likewise transposed up a tritone from the learned song.

2.3. Pitch-Recognition-Production Task. Sounds for this and subsequent behavioral tasks described here were presented over ambient speakers (Logitech 2.1 stereo computer speakers) which subjects adjusted to a comfortable level. Subjects were seated in a sound attenuated Faraday cage, positioned approximately 1.5 meters away from the speakers, in the center of the stereo field. Before the first training session and after the last, subjects listened to a series of 30 notes, randomly selected from the 5-note set (F-G-A-Bb-C) of the training melody. After each note, subjects were asked to press the corresponding key on the keyboard. Auditory feedback from the keyboard was disabled to prevent subjects from self-correcting their key presses as they progressed. This test was conducted to measure if subjects created behaviorally significant associations between sounds and discrete motor acts based on the one-note-one-finger training paradigm. Stimuli presentation and response recording were performed with Max/MSP 4.5.

2.4. Musical Training. For five consecutive days subjects practiced playing the melody line on a MIDI piano controller. On the first day, subjects were shown which five keys corresponded to the five notes used in the melody. One finger on the right hand was assigned to each of the five keys. Subjects were minimally supervised while figuring out the melody line by ear with the assistance of a computerized training environment. The song was introduced incrementally, starting with the first two measures. The subject was allowed to listen to and play along with exemplar piano lead over the two measures as many times as s/he desired. When ready, the subject would play the melody line over the backing instruments minus the exemplar piano lead. If the correct sequence of notes was played within 1st/16th note of the correct time, the computer informed the subject that he/she could move on to the next two measures. After a subject completed an additional two measures in the same fashion, the next training step was to play all of the previously learned measures in sequence. Thus they would first practice measures one and two and then measures three and four and then play measures one through four, until they could play

the complete eight-measure melody. A training session was finished for the day when the subject could play the entire song with no mistakes. Time to completion was recorded for each training session. The training environment was coded in Max/MSP 4.5. A training session also involved listening to the transposed and control melodies before and after working through the piano sequence to control for familiarity by presenting each of the three experimental songs in their entirety the same amount of time to each subject each day.

2.5. EEG Task. After completion of training and the pitch-recognition-production task, subjects completed an EEG session where they listened to six-second-long clips (two measures) from the three songs. A pair of probe tones followed each song clip and subjects were asked to respond if the two tones were present in the previous song clip. Ten clips were created from each song, totaling 30 trials across the three conditions. A resting period of two seconds preceded the onset of song stimuli. A moving baseline for mu ratio calculations was collected from this prestimulus window, across all three conditions. Thirty bins of two-second baselines equaled the same number of time points as ten bins of six-second-long stimuli per experimental condition. Stimuli were presented pseudorandomly by Neurobehavioral Systems Presentation v. 13 software. Nineteen channels of EEG and two of ocular EMG were recorded using a Neuroscan Synamps system, according to the 10–20 standards for electrode placement (F3, Fz, F4, F7, F8, Fp1, Fp2, C3, Cz, C4, P3, Pz, P4, T5, T6, O1, O2, T3, T4, VEOG). Recordings were referenced to a digitally linked pair of mastoid electrodes and grounded at Fpz. Recordings were online bandpass filtered between 0.3 and 100 Hz and amplified by a gain factor of 1000.

3. Analysis

3.1. Behavioral. Length of time to complete training was recorded each day, and a training slope variable was calculated by a linear fit of the difference between the first and the second days of training. A series of pairwise correlations were calculated for the time to complete training on the first day, the training slope, years of previous musical experience, and pitch-recognition-production scores. We also calculated the difference between pitch-recognition-production scores from the posttest and the pretest sessions and added this to the correlation matrix. Correlations and analysis of variance were computed with MATLAB v. 7.10.

3.2. EEG

3.2.1. Preprocessing. Offline data were processed in EEGLAB [34]. Data were band-passed between 3 and 40 Hz using the default FIR filter called by EEGLAB v. 12.0.2.4b. Epochs centered around onset of song stimuli were extracted to include the two-second baseline window before sound onset and the six seconds of duration of song stimulus. Independent component analysis (ICA) was performed (infomax algorithm) on the scalp channels, resulting in 19 components. Artifactual components, such as those representing eyeblinks or other head muscles, were visually identified and removed if

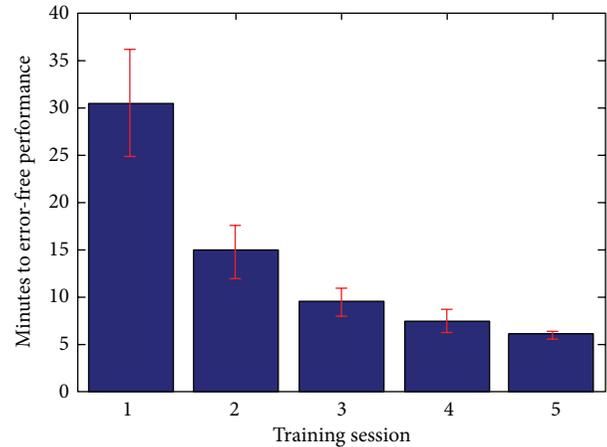


FIGURE 1: Mean amount of time to completion of training by day. Error bars (red) represent the standard error of the mean.

they met the following three criteria: (1) irregular occurrence throughout the session, (2) scalp location indicating facial muscles, and (3) presence of abnormal spectrogram, such as extremely high-power low frequencies (eyeblinks) or disproportionately large power from 20 to 30 Hz (muscle contamination). After artifacts were removed, EEG signals were remixed from source space back into channel space for further analysis.

3.2.2. Frequency Measures. All experimental conditions and baseline epochs were converted to frequency spectra using a fast Fourier transform with 0.5 Hz resolution. Frequency bands were summed with a trapezoid function for mu (8–13 Hz) and beta (20–30 Hz). Given the novel implementation of this behavioral task with EEG measures, it seemed prudent to explore other frequency bands outside of the mu and beta rhythms, such as theta (4–8 Hz) and gamma (30–40). Frequency band suppression was calculated as the log ratio of condition divided by baseline. This baseline ratio accounted for normalizing the differences inherent in spectral power due to interpersonal differences in scalp condition. The log transformation has the effect of turning ratios smaller than 1 into negative numbers, representing suppression below baseline, whereas enhancement is represented by positive values. Pairwise comparisons, correlations, and one-way ANOVAs were calculated in MATLAB v. 7.10 and repeated measure ANOVAs were computed in SPSS v. 20.0.

4. Results

4.1. Behavioral. The amount of time it took subjects to learn the melodic sequence on the first day was highly variable (mean, 30.53 minutes; SD, 22.88). A significant effect of training day ($F(4, 80) = 12.02$; $p = 1.07e - 7$) revealed a decrease in time to error-free performance and reduction in variability across all subjects over the five days of training (see Figure 1). Pearson correlation of length of time to error-free performance on the first day and years playing music shows a significant, negative slope ($r(15) = -0.58$, $p = 0.01$), but the

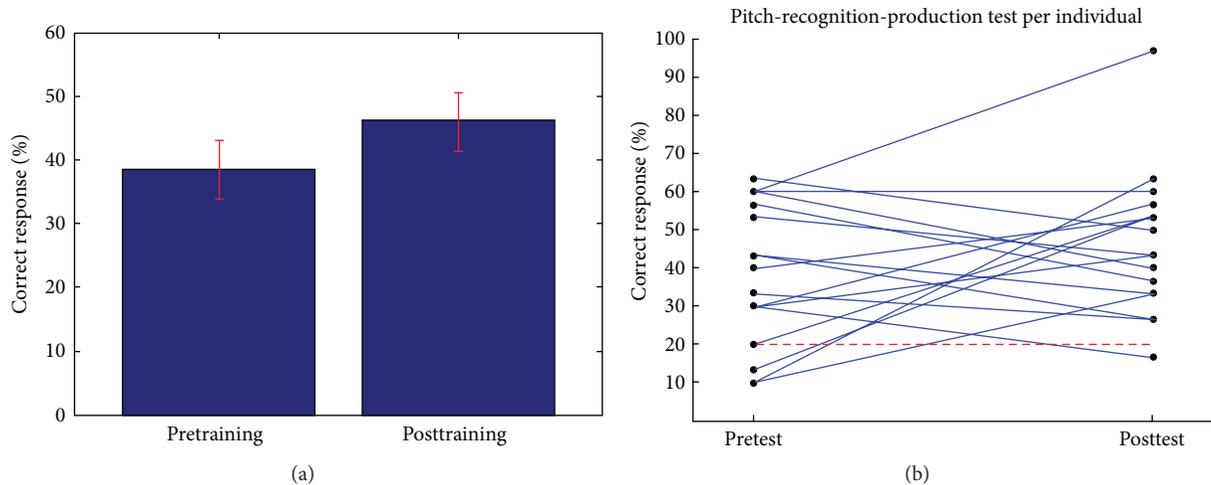


FIGURE 2: (a) Mean performance on the pitch-recognition-production test before and after musical sequence training. Chance performance is 20%. Error bars (red) represent the standard error of the mean (no significant difference between tests). (b) Individual performance with red dashed line denoting chance level performance.

relationship loses significance by the second day of training ($r(15) = -0.39$, $p = 0.12$). Previous musical experience was not correlated with pitch-recognition-production pretraining scores, posttraining scores, or the difference between them.

The pitch-recognition-production matching test showed little improvement (see Figure 2) from the pretraining percent correct (mean, 38.63% correct; SD, 5.54) to the posttraining score (mean, 46.27% correct; SD, 5.51). The musically naive subject group showed a greater improvement in mean score, from 37.4 (SD = 5.9) percent correct to 48.5 (SD = 3.6) correct after training, compared with the musically experienced group who modestly improved from 40 (SD = 5.4) to 43.8 (SD = 7.3) percent. A mixed two-way ANOVA of within-subject factor of test (pre and post) and between-subject factor of musical experience corroborates this lack of significant difference between subject groups on the P-R-P test ($F(1, 15) = 0.38$; $p = 0.55$).

4.2. EEG. Brain rhythm suppression was calculated at electrodes C3 and C4 following reports of mu activity at these recording sites [22, 35] and scalp projections of mu components are centered under these electrodes [36]. Repeated measures ANOVA with factors of electrode (C3 and C4), condition (control, learned, and transposed songs), and frequency (theta, mu, beta, and gamma) revealed a main effect of frequency ($F(3, 13) = 2988.37$, $p = 1.0e - 3$) and a marginally significant interaction between frequency and electrode ($F(3, 13) = 2.65$, $p = 0.09$). The theta and mu bands revealed consistent enhancement of power relative to baseline across all the conditions, whereas beta and gamma were generally suppressed relative to baseline. Across theta, beta, and gamma frequencies at these electrodes, the relative power was lowest for the learned song, followed next by control and then by transposed. Mu was the only frequency that exhibited higher power during the learned melody relative to the scrambled melody at C3. While we hypothesized the learned song would show greatest suppression, we did

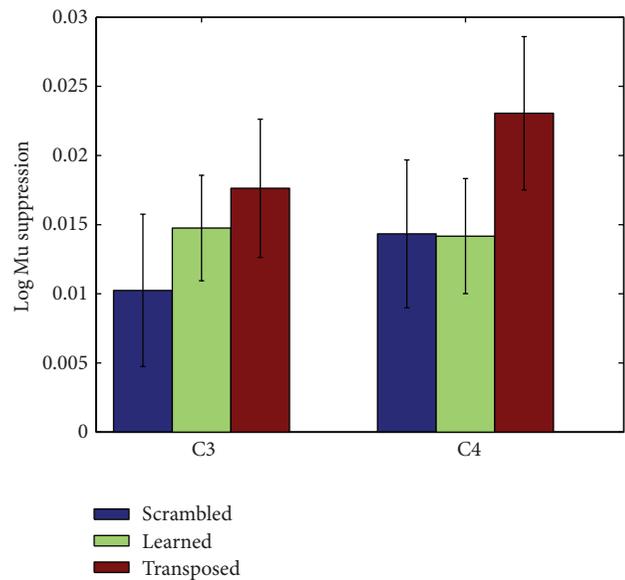


FIGURE 3: Mu suppression at electrodes C3 and C4. Error bars represent standard error of the mean.

not expect the transposed song to elicit the least amount of suppression, even enhancement in some cases. Within-subjects comparisons revealed a significant main effect for frequency ($F(3, 45) = 6247.88$, $p = 1.0e - 4$) and interaction between frequency and electrode ($F(3, 45) = 2.91$, $p = 0.05$).

The only frequency band that demonstrated suppression for all musical conditions was the beta band at electrode C3 (see Figure 4). Beta shows the same pattern of lowest log ratio power for the learned song, followed by control song and then transposed song. The pattern of relative power holds consistent for all three frequency bands at C3 and C4. The predicted results were not observed at electrode C3 or electrode C4 in the mu band (see Figure 3). Power values for each condition by frequency band and electrode

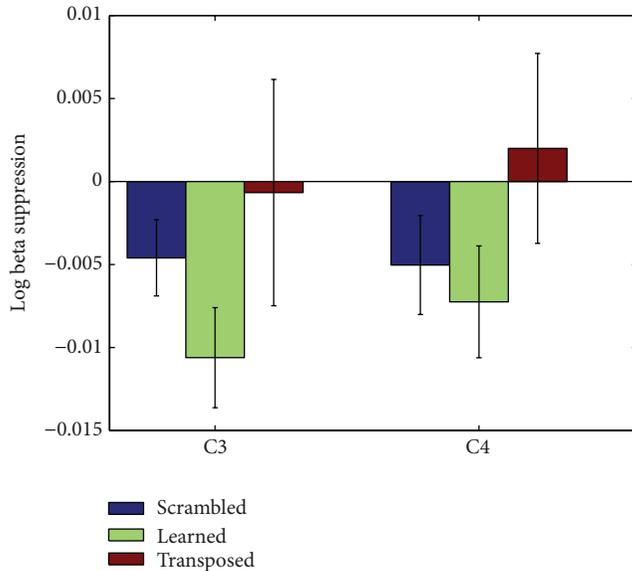


FIGURE 4: Beta suppression at electrodes C3 and C4. Error bars represent standard error of the mean.

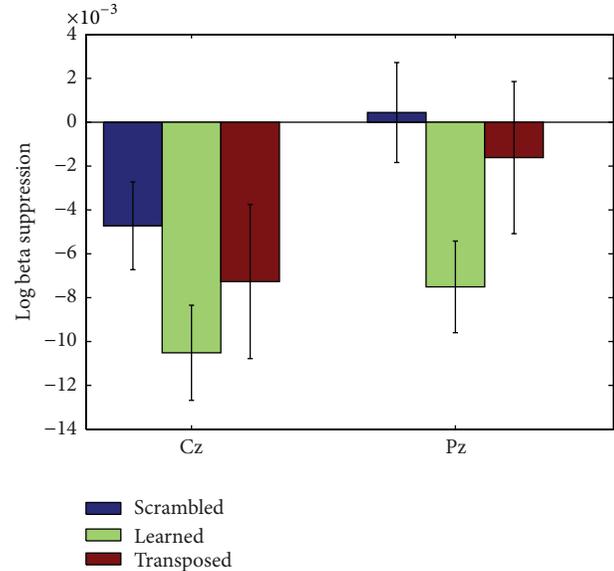


FIGURE 5: Beta suppression at electrodes Cz and Pz. Error bars represent standard error of the mean.

are available in the Supplementary Material available online at <http://dx.doi.org/10.1155/2015/638202>.

Four additional repeated measure ANOVAs were calculated, one each for the theta, mu, beta, and gamma frequency bands, with electrode (19) and condition (3) factors (power tables of condition by electrode for each frequency band are available in Supplementary Materials). Within-subject effects revealed a main effect of electrode for the mu ($F(18, 270) = 1.7, p = 0.04$) and theta bands ($F(18, 270) = 3.508, p = 1.0e-3$). No main effect was observed for condition; however the theta ($F(36, 540) = 1.59, p = 0.2$), beta ($F(36, 540) = 1.59, p = 0.02$), and gamma ($F(36, 540) = 1.656, p = 0.01$) bands all revealed a significant interaction between electrodes and conditions. As seen in Figure 6, the activity across all frequency bands was centered over the midline electrodes. Examining these electrodes, both Cz and Pz revealed a trend in the predicted direction across conditions in the beta band, with learned melody exhibiting greatest suppression, followed next by transposed and lastly by the scrambled control melody. A significant main effect was observed for condition at these two recording sites ($F(2, 14) = 7.12, p = 0.007$). All three musical conditions exhibited suppression with regard to the baseline at Cz, and at Pz the scrambled melody control showed a slight enhancement (Figure 5). Overall Cz showed greater suppression than Pz for all three conditions, though the difference between learned and scrambled melodies was larger at Pz. As the beta band has been shown to play a role in perception of sounds [29, 37] and the effect was only observed over sensorimotor cortex, this finding supports the hypothesis that the motor system may be involved in perception of musical sequences. No significant correlations were found between beta suppression at these sites and years of musical training, pitch-recognition-production difference scores, or length of time to reach error-free performance on the first day of training (statistics reported in Table 2). Lack

of correlation between brain responses and these behavioral measures indicated that prior musical experience or aptitude is not likely influencing the neural physiological responses at the level of individuals. However, when musical experience was included as a between-subjects factor of group (musically experienced, musically naive) and repeated measures ANOVA with factors of central electrode (C3, Cz, C4, and Pz) and song condition (control, learned, and transposed), there was a main effect of electrode ($F(3, 12) = 10.84, p = 0.001$) and an interaction between electrode by group ($F(3, 12) = 8.42, P = 0.003$). The greatest group differences were at C3, with nonmusicians exhibiting greater beta suppression, and C4 where musically experienced participants exhibited greater beta suppression.

As cortical alpha power is variable in its peak frequency across the population [38] we performed an additional multivariate analysis on mu power that was integrated over a band defined by an individual's mu peak frequency within a fixed bandwidth (8–20 Hz, the frequency window between theta and beta cutoffs) [39]. Repeated measures ANOVA with factors of electrode (19) and condition (3) revealed no main effect of condition ($F(2, 30) = 0.43, p = 0.657$) or an interaction between condition and electrode ($F(36, 540) = 0.68, p = 0.92$). Including a between-subjects factor of group (2) in a repeated measure ANOVA with central electrode (C3, Cz, C4, and Pz) and condition (3) factors revealed a within-subjects interaction of electrode by musical experience ($F(3, 42) = 2.94, p = 0.044$). The musician group showed virtually no suppression at electrodes C3 and Pz, relative to nonmusicians who exhibited relatively high levels of suppression. The interaction was not further explained by the addition of condition, as the three-way interaction was not significant ($F(6, 84) = 1.28, p = 0.27$), and the differences between musicians and nonmusicians are greatest at the lateralized electrodes (C3 and C4). Power values are

TABLE 2: Nonparametric correlations between behavioral measures and beta band activity.

	Beta band behavioral correlations											
	C3			C4			CZ			PZ		
	control	learned	transposed	control	learned	transposed	control	learned	transposed	control	learned	transposed
Spearman's rho												
Completion time day 1	.833 16	.914 .029 16	.935 -.022 16	.974 .009 16	.284 .286 16	.766 .081 16	.948 -.018 16	.427 .213 16	.300 .277 16	.884 -.040 16	.152 .375 16	.791 .072 16
Years of musical training	.529 .170 16	.912 .030 16	.899 -.035 16	.419 -.217 16	.412 -.220 16	.335 -.258 16	.325 -.263 16	.296 -.279 16	.091 -.436 16	.958 -.014 16	.072 -.461 16	.402 -.225 16
Pre- and post difference PRP test	.491 -.186 16	.184 -.350 16	.158 -.370 16	.401 -.226 16	.578 -.150 16	.349 -.251 16	.014 -.599 16	.410 -.221 16	.559 -.158 16	.320 -.265 16	.704 -.103 16	.644 -.125 16

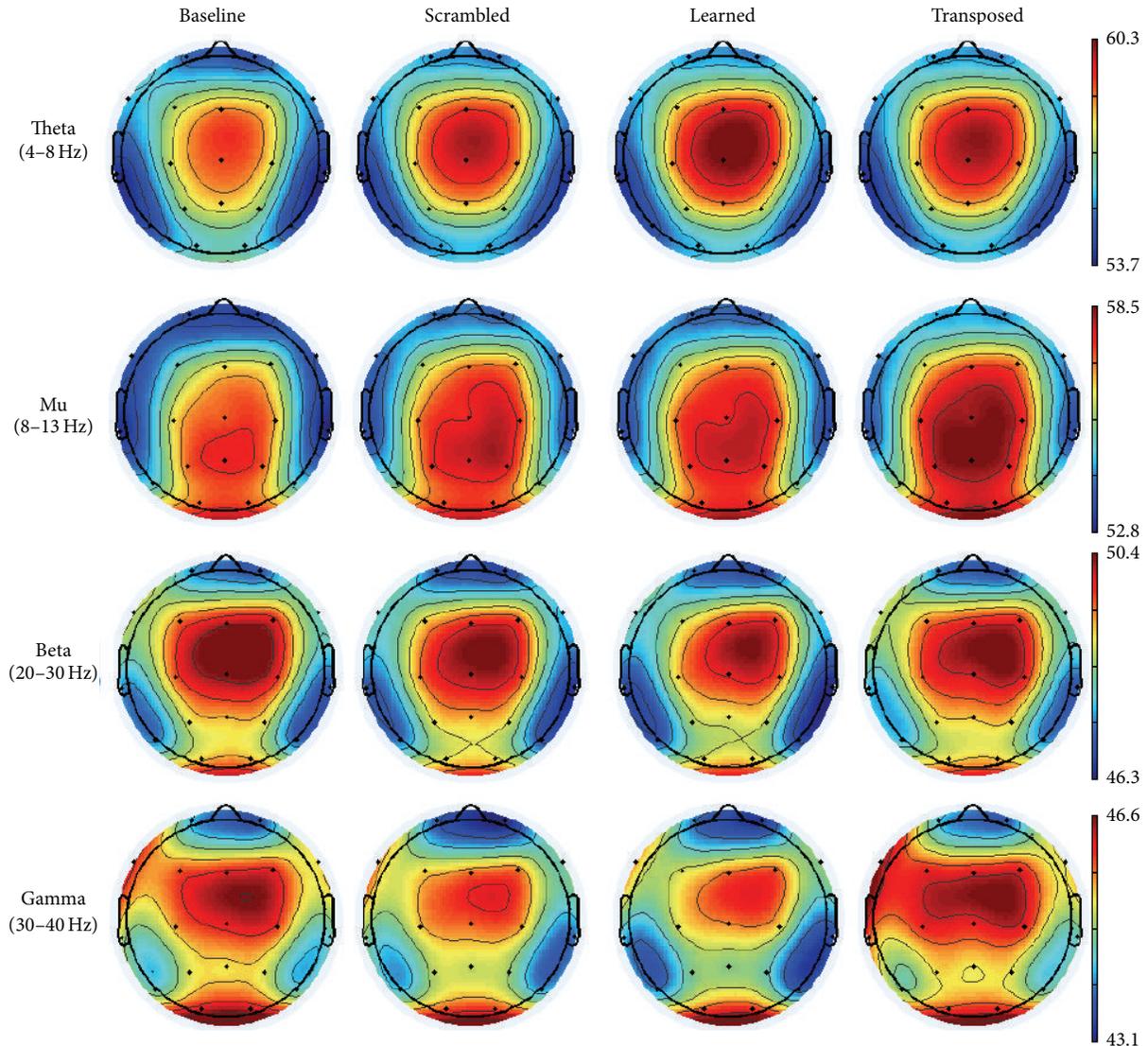


FIGURE 6: Scalp distribution of different frequency band activities. Units for power heat maps are microvolts squared.

available in the Supplementary Materials for individual mu peak power spectra for all electrodes by condition and for central cluster of electrodes by group.

5. Discussion

The present study reports a novel finding that cortical audiomotor system activity, as evidenced by EEG, is sensitive to the effects of auditory sequence manipulations when the sounds are associated with movement. When subjects heard a melody they learned to play and a transposed version of that melody, they exhibited suppression of the beta band relative to baseline and relative to a scrambled melody control. The level of motor system engagement indexed by beta suppression was greater in response to listening to learned melodies than listening to the transposed version of these melodies. Suppression of the beta rhythm while listening to transposed versions of the learned melody indicates a role for motor system associations with the sequential aspects of an

auditory stimulus. The present study builds on previous work by Lahav et al. [8] who reported motor system activity in the form of a blood oxygen level dependent (BOLD) signal when listening to the same learned melody. Lahav et al. further demonstrated partial motor system activation when subjects heard a novel melody composed of notes from the learned melody. They interpreted the partial activation as evidence that the motor system was sensitive to associations formed between single notes and single finger movements (recall one note per finger on the right hand). We hypothesized that the difference between motor system engagement levels for learned melodies and scrambled melodies of the same note set could be explained by an audiomotor system sensitivity to sequences of sounds. The evidence currently reported supports this hypothesis. A logical extension of the work would compare the levels of motor system engagement while listening to the transposed melody and scrambled melody (same notes and different sequence from learned song). If the motor system is sensitive to both single note (pitch)

and sequence information, then the sum of motor activation between these two conditions should approximate the level observed when subjects listen to the melody they learned to play.

Suppression of beta rhythms is associated with motor engagement. Caetano et al. [29] reported that the beta frequency band desynchronizes in anticipation of actions, hearing, and seeing that same action, followed by rebound synchronization after completion of the event. They also reported that the mu rhythm followed a similar time course of suppression and rebound enhancement, however with slight delays in rebound compared to the beta rhythm. During the sound only condition, mu and beta responses exhibited less suppression during the anticipatory phase but rebound at the same latency as visual based stimuli. This suggests that prestimulus suppression is related to movement preparation or planning. Boonstra et al. [30] also report a similar beta suppression during auditory perception and a pretap suppression with rebound enhancement when subjects tapped along with a steady, rhythmic sound. Both beta and mu frequency bands are seen as the two most active bands in terms of phase coherence between a cerebral network engaged during rhythmic sound tap synchronizing [40]. Further evidence to support the relevance of beta band in musical sequencing or rhythmic processing comes from reports of activity centered around 25 Hz in response to missing (expected) rhythmic sounds [41–43] and from its proposed role in modulating perceived beat structure [37].

Rather than predicted mu suppression, mu enhancement was observed across all conditions. The trend was similar across electrodes in the sensorimotor scalp region. Scrambled and learned song conditions had similar low levels of enhancement, while the transposed song enhanced mu significantly greater than the other two conditions. Mu suppression was predicted based on past work identifying mu rhythms as having high power during rest and suppressed during visual, auditory, and audiovisual input sans movement [28]. However, the stimuli used by McGarry et al. [28] were not musical in nature, as the action and sounds were tearing a sheet of paper. Prediction of action sounds in the absence of visual input is difficult. Music and language, conforming to grammatical rules, can build predictions of future sounds based on the relationships between and sequences of past sounds. To assume that the same neural system predicts visual and auditory movement related stimuli may be an error, even though auditory information can facilitate visual processing. While the mu rhythm is sensitive to auditory information [28] and plays a role in visual to motor and audio transformations in terms of reading sheet music [33], it may not play a direct role in audiomotor processing by itself. Caetano et al. [29] report that the mu rhythm responded more robustly to tapping on a drum when there is somatosensory feedback. Pantomiming the same tapping action in the absence of a surface to tap on fails to suppress the mu rhythm in the same way as a tap with tactile feedback. Listening to a melodic sequence that has motor associations may not suppress mu rhythms as the experience does not include the sensation, or perhaps even simulation of a tactile response. However one should draw comparisons cautiously

as previously reported mu responses to tapping indicate a response to a discrete movement, whereas in the present study neural responses were averaged over several discrete sound-action pairings.

Mu enhancement, greatest while listening to the transposed song, may reflect an inhibitory response, rather than motor system preparation simply associated with movement. The inhibition timing hypothesis [44, 45] interprets event related increases of mu power as an inhibition response during intensive activity in other cortical rhythms. For instance, mu may signify the inhibition of motor output, such that changes in beta rhythms sequencing motor commands remain a simulation, unable to affect actual muscles. The hypothesis further proposes a role of mu enhancement as a main source of synchronization in cortical rhythms to synchronize neuronal timing. Given the strong temporal structure of musical sounds, it may be reasonable to expect mu power to increase during perception of more complex sound passages that require additional sequencing resources. It is possible that listening to a transposition requires more cortical resources such that greater demands are placed on the timing mechanisms critical to sound sequencing. If mu enhancement reflects increased difficulty of auditory sequencing, then listening to complex rhythmic patterns may elicit greater enhancement of mu rhythms than simple rhythmic patterns.

We additionally report a novel inclusion of musically experienced subjects with the piano learning ear training task. At the individual subject level, previous music experience was not associated with any other measure, except the length of time it took to complete the first day of ear training. At the group level, previous musical training was associated with lower levels of mu and beta suppression over the left sensorimotor cortex. For instrumentalists who have experience associating right hand movements with pitch perception in music performance, this may represent a more efficient use of cortical resources. The right sensorimotor cortex, contralateral to piano trained hand, exhibited greater beta suppression in musically experienced subjects. This might be explained in part by previous reports indicating a preferential role of the right hemisphere in relative pitch processing [46]. Musical training may be represented in this case by increased recruitment of right hemisphere sensorimotor networks during pitch processing, resulting in increased suppression. Musical experience may explain effects in the present data, but caution is urged in generalization of these findings, as we did not specifically recruit musicians, and only two of our subjects were regularly playing music at the time of their participation.

Subjects in the present study learned to play the melodic sequence by ear, as evidenced by the changes in length of time to complete training across sessions, replicating past use of this behavioral task [17]. The time to complete training followed an exponential decay curve, also showing a collapse of variance across subjects. Our novel inclusion of subjects with previous musical experience had an effect on the length of time to learn the song on the first day of training, but after five days of ear training and piano playing, differences between subjects groups were negligible. The two population

groups, both piano naive, did not have significantly different P-R-P test scores before or after training. As previous reports from Lahav et al. [17] were performed with only musically naive subjects, the present work extends this task as a viable training with little differences for both the musically naive and musically experienced but piano naive subjects.

Taken in light of previous findings, enhanced mu and suppressed beta might indicate greater cortical demands in response to sounds associated with a motor action. As previous authors [8] hypothesized a trained association between discrete musical pitches and discrete finger movements, an extension of this work could make use of the temporal resolution of EEG and design the posttraining assessment to focus on reactions to discrete musical notes or a sequence of multiple notes to test whether cortical oscillations are recruited according to the time intervals related to sequence complexity. Relevant work from functional imaging suggests increases in auditory and premotor cortical activity proportional to difficulty of tapped rhythms [10]. If the motor system contributes to offline processing of sequential or rhythmic sounds at the level of discrete sounds, then one could predict beta desynchronization or mu synchronization in response to heard sounds without movement.

Conflict of Interests

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interests.

Acknowledgments

The authors would like to thank Noah Bresler for help with subject training and data collection, Steven Pillen for consultation with regard to data analysis, and Dronker's lab writing group for feedback on earlier versions of the paper.

References

- [1] S. Baumann, S. Koeneke, M. Meyer, K. Lutz, and L. Jäncke, "A network for sensory-motor integration: what happens in the auditory cortex during piano playing without acoustic feedback?" *Annals of the New York Academy of Sciences*, vol. 1060, pp. 186–188, 2005.
- [2] M. Bangert, T. Peschel, G. Schlaug et al., "Shared networks for auditory and motor processing in professional pianists: evidence from fMRI conjunction," *NeuroImage*, vol. 30, no. 3, pp. 917–926, 2006.
- [3] S. Baumann, S. Koeneke, C. F. Schmidt, M. Meyer, K. Lutz, and L. Jancke, "A network for audio-motor coordination in skilled pianists and non-musicians," *Brain Research*, vol. 1161, pp. 65–78, 2007.
- [4] J. L. Chen, R. J. Zatorre, and V. B. Penhune, "Interactions between auditory and dorsal premotor cortex during synchronization to musical rhythms," *NeuroImage*, vol. 32, no. 4, pp. 1771–1781, 2006.
- [5] R. J. Zatorre, "There's more to auditory cortex than meets the ear," *Hearing Research*, vol. 229, no. 1-2, pp. 24–30, 2007.
- [6] A. D. Patel and J. R. Iversen, "The evolutionary neuroscience of musical beat perception: the Action Simulation for Auditory Prediction (ASAP) hypothesis," *Frontiers in Systems Neuroscience*, vol. 8, article 57, 2014.
- [7] S. Li, B. Hong, X. Gao, Y. Wang, and S. Gao, "Event-related spectral perturbation induced by action-related sound," *Neuroscience Letters*, vol. 491, no. 3, pp. 165–167, 2011.
- [8] A. Lahav, E. Saltzman, and G. Schlaug, "Action representation of sound: audiomotor recognition network while listening to newly acquired actions," *The Journal of Neuroscience*, vol. 27, no. 2, pp. 308–314, 2007.
- [9] J. L. Chen, V. B. Penhune, and R. J. Zatorre, "Moving on time: brain network for auditory-motor synchronization is modulated by rhythm complexity and musical training," *Journal of Cognitive Neuroscience*, vol. 20, no. 2, pp. 226–239, 2008.
- [10] J. L. Chen, V. B. Penhune, and R. J. Zatorre, "The role of auditory and premotor cortex in sensorimotor transformations," *Annals of the New York Academy of Sciences*, vol. 1169, pp. 15–34, 2009.
- [11] E. H. Margulis, L. M. Mlsna, A. K. Uppunda, T. B. Parrish, and P. C. M. Wong, "Selective neurophysiologic responses to music in instrumentalists with different listening biographies," *Human Brain Mapping*, vol. 30, no. 1, pp. 267–275, 2009.
- [12] A. Patel, "Language, music, syntax and the brain," *Nature Neuroscience*, vol. 6, no. 7, pp. 674–681, 2003.
- [13] I. Molnar-Szakacs and K. Overy, "Music and mirror neurons: from motion to 'emotion,'" *Social Cognitive and Affective Neuroscience*, vol. 1, no. 3, pp. 235–241, 2006.
- [14] K. Overy and I. Molnar-Szakacs, "Being together in time: musical experience and the mirror neuron system," *Music Perception*, vol. 26, no. 5, pp. 489–504, 2009.
- [15] R. J. Zatorre, J. L. Chen, and V. B. Penhune, "When the brain plays music: auditory-motor interactions in music perception and production," *Nature Reviews Neuroscience*, vol. 8, no. 7, pp. 547–558, 2007.
- [16] M. Siegel, T. J. Buschman, and E. K. Miller, "Cortical information flow during flexible sensorimotor decisions," *Science*, vol. 348, no. 6241, pp. 1352–1355, 2015.
- [17] A. Lahav, A. Boulanger, G. Schlaug, and E. Saltzman, "The power of listening: auditory-motor interactions in musical training," *Annals of the New York Academy of Sciences*, vol. 1060, pp. 189–194, 2005.
- [18] J. A. Pineda, "Sensorimotor cortex as a critical component of an 'extended' mirror neuron system: does it solve the development, correspondence, and control problems in mirroring?" *Behavioral and Brain Functions*, vol. 4, article 47, 2008.
- [19] J. Haueisen and T. R. Knösche, "Involuntary motor activity in pianists evoked by music perception," *Journal of Cognitive Neuroscience*, vol. 13, no. 6, pp. 786–792, 2001.
- [20] J. A. Pineda, "The functional significance of mu rhythms: translating 'seeing' and 'hearing' into 'doing,'" *Brain Research Reviews*, vol. 50, no. 1, pp. 57–68, 2005.
- [21] C. Neuper, M. Wörtz, and G. Pfurtscheller, "ERD/ERS patterns reflecting sensorimotor activation and deactivation," *Progress in Brain Research*, vol. 159, pp. 211–222, 2006.
- [22] L. M. Oberman, V. S. Ramachandran, and J. A. Pineda, "Modulation of mu suppression in children with autism spectrum disorders in response to familiar or unfamiliar stimuli: the mirror neuron hypothesis," *Neuropsychologia*, vol. 46, no. 5, pp. 1558–1565, 2008.
- [23] S. Cochin, C. Barthelemy, B. Lejeune, S. Roux, and J. Martineau, "Perception of motion and qEEG activity in human adults," *Electroencephalography and Clinical Neurophysiology*, vol. 107, no. 4, pp. 287–295, 1998.

- [24] J. A. Pineda, B. Z. Allison, and A. Vankov, "The effects of self-movement, observation, and imagination on μ rhythms and readiness potentials (RP'S): toward a brain-computer interface (BCI)," *IEEE Transactions on Rehabilitation Engineering*, vol. 8, no. 2, pp. 219–222, 2000.
- [25] S. D. Muthukumaraswamy, B. W. Johnson, and N. A. McNair, "Mu rhythm modulation during observation of an object-directed grasp," *Cognitive Brain Research*, vol. 19, no. 2, pp. 195–201, 2004.
- [26] G. Pfurtscheller, C. Brunner, A. Schlögl, and F. H. Lopes da Silva, "Mu rhythm (de)synchronization and EEG single-trial classification of different motor imagery tasks," *NeuroImage*, vol. 31, no. 1, pp. 153–159, 2006.
- [27] R. Salmelin and R. Hari, "Characterization of spontaneous MEG rhythms in healthy adults," *Electroencephalography and Clinical Neurophysiology*, vol. 91, no. 4, pp. 237–248, 1994.
- [28] L. M. McGarry, F. A. Russo, M. D. Schalles, and J. A. Pineda, "Audio-visual facilitation of the mu rhythm," *Experimental Brain Research*, vol. 218, no. 4, pp. 527–538, 2012.
- [29] G. Caetano, V. Jousmäki, and R. Hari, "Actor's and observer's primary motor cortices stabilize similarly after seen or heard motor actions," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 104, no. 21, pp. 9058–9062, 2007.
- [30] T. W. Boonstra, A. Daffertshofer, C. E. Peper, and P. J. Beek, "Amplitude and phase dynamics associated with acoustically paced finger tapping," *Brain Research*, vol. 1109, no. 1, pp. 60–69, 2006.
- [31] T. Fujioka, L. J. Trainor, E. W. Large, and B. Ross, "Beta and gamma rhythms in human auditory cortex during musical beat processing," *Annals of the New York Academy of Sciences*, vol. 1169, pp. 89–92, 2009.
- [32] T. Fujioka, L. J. Trainor, E. W. Large, and B. Ross, "Internalized timing of isochronous sounds is represented in neuromagnetic β oscillations," *The Journal of Neuroscience*, vol. 32, no. 5, pp. 1791–1802, 2012.
- [33] L. P. Behmer and K. J. Jantzen, "Reading sheet music facilitates sensorimotor mu-desynchronization in musicians," *Clinical Neurophysiology*, vol. 122, no. 7, pp. 1342–1347, 2011.
- [34] S. Makeig, S. Debener, J. Onton, and A. Delorme, "Mining event-related brain dynamics," *Trends in Cognitive Sciences*, vol. 8, no. 5, pp. 204–210, 2004.
- [35] E. R. Ulloa and J. A. Pineda, "Recognition of point-light biological motion: mu rhythms and mirror neuron activity," *Behavioural Brain Research*, vol. 183, no. 2, pp. 188–194, 2007.
- [36] A. Moore, I. Gorodnitsky, and J. Pineda, "EEG mu component responses to viewing emotional faces," *Behavioural Brain Research*, vol. 226, no. 1, pp. 309–316, 2012.
- [37] J. R. Iversen, B. H. Repp, and A. D. Patel, "Top-down control of rhythm perception modulates early auditory responses," *Annals of the New York Academy of Sciences*, vol. 1169, pp. 58–73, 2009.
- [38] W. Klimesch, "EEG-alpha rhythms and memory processes," *International Journal of Psychophysiology*, vol. 26, no. 1–3, pp. 319–340, 1997.
- [39] O. M. Bazanova and D. Vernon, "Interpreting EEG alpha activity," *Neuroscience & Biobehavioral Reviews*, vol. 44, pp. 94–110, 2014.
- [40] B. Pollok, J. Gross, K. Müller, G. Aschersleben, and A. Schnitzler, "The cerebral oscillatory network associated with auditorily paced finger movements," *NeuroImage*, vol. 24, no. 3, pp. 646–655, 2005.
- [41] T. P. Zanto, E. W. Large, A. Fuchs, and J. A. S. Kelso, "Gamma-band responses to perturbed auditory sequences: evidence for synchronization of perceptual processes," *Music Perception*, vol. 22, no. 3, pp. 531–547, 2005.
- [42] T. P. Zanto, J. S. Snyder, and E. W. Large, "Neural correlates of rhythmic expectancy," *Advances in Cognitive Psychology*, vol. 2, no. 2, pp. 221–231, 2006.
- [43] J. S. Snyder and E. W. Large, "Gamma-band activity reflects the metric structure of rhythmic tone sequences," *Cognitive Brain Research*, vol. 24, no. 1, pp. 117–126, 2005.
- [44] W. Klimesch, P. Sauseng, and S. Hanslmayr, "EEG alpha oscillations: the inhibition-timing hypothesis," *Brain Research Reviews*, vol. 53, no. 1, pp. 63–88, 2007.
- [45] W. Klimesch, "Alpha-band oscillations, attention, and controlled access to stored information," *Trends in Cognitive Sciences*, vol. 16, no. 12, pp. 606–617, 2012.
- [46] N. E. V. Foster and R. J. Zatorre, "Cortical structure predicts success in performing musical transformation judgments," *NeuroImage*, vol. 53, no. 1, pp. 26–36, 2010.

Research Article

Sound Richness of Music Might Be Mediated by Color Perception: A PET Study

Masayuki Satoh,¹ Ken Nagata,² and Hidekazu Tomimoto^{1,3}

¹Department of Dementia Prevention and Therapeutics, Graduate School of Medicine, Mie University, 2-174 Edobashi, Tsu, Mie 514-8507, Japan

²Department of Neurology, Research Institute for Brain and Blood Vessels, 6-10 Senshu-Kubota-Machi, Akita 010-0874, Japan

³Department of Neurology, Mie University Graduate School of Medicine, 2-174 Edobashi, Tsu, Mie 514-8507, Japan

Correspondence should be addressed to Masayuki Satoh; bruckner@clin.medic.mie-u.ac.jp

Received 6 March 2015; Accepted 17 May 2015

Academic Editor: Michael E. Behen

Copyright © 2015 Masayuki Satoh et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Objects. We investigated the role of the fusiform cortex in music processing with the use of PET, focusing on the perception of sound richness. **Method.** Musically naïve subjects listened to familiar melodies with three kinds of accompaniments: (i) an accompaniment composed of only three basic chords (chord condition), (ii) a simple accompaniment typically used in traditional music text books in elementary school (simple condition), and (iii) an accompaniment with rich and flowery sounds composed by a professional composer (complex condition). Using a PET subtraction technique, we studied changes in regional cerebral blood flow (rCBF) in simple minus chord, complex minus simple, and complex minus chord conditions. **Results.** The simple minus chord, complex minus simple, and complex minus chord conditions regularly showed increases in rCBF at the posterior portion of the inferior temporal gyrus, including the LOC and fusiform gyrus. **Conclusions.** We may conclude that certain association cortices such as the LOC and the fusiform cortex may represent centers of multisensory integration, with foreground and background segregation occurring at the LOC level and the recognition of richness and floweriness of stimuli occurring in the fusiform cortex, both in terms of vision and audition.

1. Introduction

Historically, the style of music has developed from simple to complex. Such development was typically classified as a change from monophony, that is, music for a single voice or part [1], to polyphony, in which two or more strands sound simultaneously, or to homophony in which there was a clear distinction between melody and accompanying harmony [1]. In music with a monophonic style, only the melody is produced and there is no accompaniment. In homophony, to which most nursery and folk songs of western music belong, music consists of melody and its accompaniment. As music with homophonic or polyphonic styles has developed, harmonies have become more complex. For example, music of Mozart or Haydn in the 18th century rarely utilized dissonant chords, while the 20th century music of Ravel or Debussy had several kinds of chords including dissonant ones. Listening to homophonic music is different

from listening to monophonic music according to the following. First, with homophonic music, listeners discriminate melody and its accompaniment. Even if the melody and the accompaniment are played by the same instrument (i.e., with the identical timbre), we can easily and instantaneously perceive the melody and the accompaniment. The neural basis of this is still unknown, but we previously reported in a positron emission tomography (PET) activation study that the lateral occipital complex (LOC), which participates in foreground and background segregation in vision, plays an important role in the discrimination between melody and its accompaniment [2]. The melody and the accompaniment could be regarded, in auditory terms, as the foreground and background, respectively. We suggested that the same neural substrates carried out similar functions beyond the simple discrimination of sensory modalities. Second, the sounds of homophonic music could be richer than monophonic music. The quality of sound is generally called “timbre.” The timbre

is operationally defined as the attribute that distinguishes sounds of equal pitch, loudness, location, and duration [3]. The term “timbre” not only relates to the individual musical instrument, but also relates to expressing the characteristics of the sound of musical pieces. For example, it is generally considered that the timbre of impressionist music of Ravel or Debussy is richer and more flowery than the classical music of Mozart or Haydn. In the above-mentioned PET study, the melody with accompaniment also activated the fusiform cortex (in addition to the LOC) compared to the melody without the accompaniment [2]. We interpreted the activation of the fusiform cortex to reflect the rich sound from the accompaniment, but much still remains to be done to identify the role of that area in listening to music.

Over the past few decades, a considerable number of PET activation studies have been made on various aspects of music, sounds, and the brain [4–6], not only in healthy subjects [4, 6] but also in patients with tinnitus [5]. Based on our previous researches, we performed another PET study that investigated brain region activity while subjects listened to melodies with various kinds of accompaniments. Musically naïve subjects listened to melodies of familiar nursery songs with various degrees of sound richness of the accompaniment. According to a visual analogue scale (VAS), for each piece of music we also ascertained to what extent the subjects felt the sound was rich. Using a PET subtraction technique, brain regions that were significantly activated by sound richness were identified.

2. Subjects and Methods

2.1. Subjects. Ten right-handed male volunteers (mean age 21.7 ± 0.95 years; range 20–24) participated in the study. All were students at the Schools of Engineering or Mining, Akita University, and met criteria for Grison’s second level of musical culture [7]. None had received any formal or private musical education, and none had any signs or history of neurological, cardiovascular, or psychiatric disease. All subjects gave written informed consent after the purpose and procedure of the examination had been fully explained. The study was approved by the Ethics Committee of the Research Institute for Brain and Blood Vessels, Akita, Japan, and all experiments were conducted in accordance with the Declaration of Helsinki.

2.2. Task Procedures. The stimuli in this experiment were six melodies of well-known Japanese nursery songs. All subjects were very familiar with these melodies. For each melody, the following three kinds of accompaniment were composed: (i) an accompaniment composed by using only three basic chords (tonic, dominant, and subdominant chord), one of which was set on each bar (chord condition), (ii) a simple accompaniment that is typically used in the traditional music text books in Japanese elementary schools (simple condition), and (iii) an accompaniment with rich and flowery sounds composed by a professional composer [8] (complex condition). The (i) chord and (ii) simple condition accompaniments were composed by one of the authors (Masayuki Satoh). The accompaniment of simple condition consisted of

quarter tones of a chord on a whole note of fundamental tone. The first beat of each cord in the bar was rest, so only the fundamental tone was played at the first beat. All musical stimuli were played using the “FINALE” software [9]. The author Masayuki Satoh wrote musical scores of musical pieces used in this experiment on the “FINALE,” and the software played each piece with piano timbre. Each performance was recorded on a compact disc. Melodies with the three types of accompaniments were randomly presented. Subjects were instructed to listen to each melody, and PET measurements were obtained while listening to these melodies (procedures described below). Subjects were required to make a sign with the index finger of the right hand as the melody of each song finished. All stimuli were presented binaurally via inset stereo earphones.

The instruction to the subjects was as follows: Close your eyes. You will listen to a melody of a familiar nursery song. If you feel that the melody has finished, please make a sign with the index finger of your right hand.

2.3. Positron Emission Tomography Measurements. The protocol used in this study has been previously described in detail [2, 10–12]. Briefly, PET data were acquired in 3D acquisition mode using Headtome V (Shimadzu, Kyoto, Japan). Scans were performed in a darkened room with subjects lying supine with eyes closed. Nine CBF measurements were determined for each subject, three during the chord, three during the simple, and three during the complex condition. Employing ^{15}O -labeled water (H_2^{15}O) intravenous bolus technique [13], emission data were collected for 90 seconds for each measurement following intravenous bolus injection of about 15 mL (40 mCi) H_2^{15}O . A musical piece was initiated 15 seconds prior to data acquisition, followed by another musical piece, and this in total continued for about 120 seconds. Emission data were corrected for attenuation by acquiring 10 minutes of transmission data utilizing ^{68}Ge orbiting rod source performed prior to the activation scans. A wash-out period of approximately 10 minutes was allowed between successive scans. For anatomic reference, all subjects underwent axial T1-weighted imaging (T1WI) and T2-weighted imaging (T2WI) using a 1.5 T magnetic resonance system (Vision, Siemens, Germany). T1WI (TR/TE = 665/14 ms) and T2WI (TR/TE = 3600/96 ms) were obtained using a slice thickness of 5 mm with an interslice gap of 1 mm.

2.4. Data Analysis. PET data analysis was performed on a SGI Indy running IRIX 6.5 (Silicon Graphics, California), using an automated PET activation analysis package [14] composed of six main processing stages which has been previously described in detail [2, 10–12]. The six main stages consisted of intrasubject coregistration, intrasubject normalization, automatic detection of the AC-PC line, detection of multiple stretching points and surface landmarks on intrasubject averaged image sets, intersubject summation and statistical analyses, and superimposition of statistical results onto the stereotactic MRI. Deformation of individual brains to correspond with the standard atlas brain was achieved by spatially matching individual landmarks to the corresponding predefined standard surface landmarks and minimizing

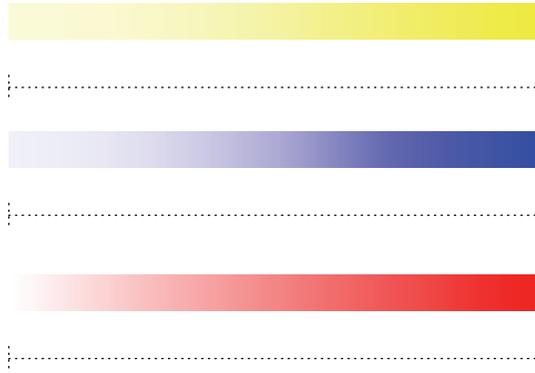


FIGURE 1: Visual analogue scale (VAS) for the assessment of subjective impression of sound richness of each musical stimulus. The length of color bar is 140 mm.

correlation coefficients of regional profile curves between the stretching centers. Activation foci were considered to be significantly activated if the corresponding p value was less than a predetermined threshold ($p < 0.001$, Bonferroni correction for multiple comparisons). Anatomical identification of activation foci was achieved by referring the stereotactic coordinates of the peak activated pixels to the standard Talairach brain atlas [15].

2.5. Visual Analogue Scale (VAS) of Sound Richness. After the PET measurement, the degree of sound richness of each melody with the three types of accompaniments was investigated in each subject. In a quiet room, each subject listened to the stimuli and was required to subjectively mark the VAS (Figure 1) according to the degree of sound richness the subject felt. Three colors (yellow, blue, and red) were used because the lyrics of some songs had a relationship with a specific color, for example, the sea related to blue and the sunset to red. Subjects marked to the right to the degree that they felt that the sound of the music was rich. We measured the distance from the left end to the marked position (mm) and, using the Wilcoxon signed rank test, statistically compared the distance between the three kinds of accompaniments, namely, chord, simple, and complex condition.

3. Results

Regarding the VAS of sound richness, the mean distance from the left end was significantly longer as the accompaniment became more complex (Figure 2): chord condition 54.2 ± 34.2 ; simple condition 71.3 ± 30.1 ; complex condition 101.4 ± 31.1 mm (mean \pm standard deviation (sd)). We can reasonably conclude that, as expected, the more complex the accompaniment became, the richer the subjects reported the sound.

The results of subtractions providing significant regions activated as the sound became more complex are given in Tables 1, 2, and 3 and Figures 3, 4, and 5. The regions activated during the simple condition but not during the chord condition are listed in Table 1 together with stereotactic coordinates based on the brain atlas of Talairach and

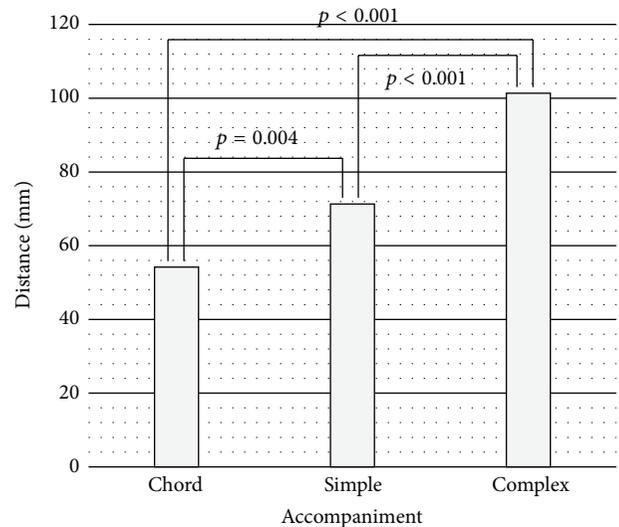


FIGURE 2: Results of VAS of each accompaniment condition. As the accompaniment became more complex, the subjects regarded the sound of musical pieces as being richer.

Tournoux [15]. These results show areas of relative blood flow changes that emphasize differences between the two conditions and minimize areas that are common to both conditions. Significant increases in relative cortical blood flow were found in the posterior portion of the left inferior temporal gyrus, bilateral fusiform gyri, the medial surface of the bilateral frontal lobes, the right superior parietal lobule, and the left orbital frontal cortex (Table 1, Figure 3). Compared to the chord condition, the complex condition produced significant activation at the posterior portion of the left inferior temporal gyrus, left fusiform gyrus, right medial surface of the occipital lobe, the lateral surface of the left occipital lobe, and the anterior portion of the left middle temporal gyrus (Table 2, Figure 4). Between the complex and simple condition, the former condition significantly activated the posterior portion of the left inferior temporal gyrus, the left fusiform gyrus, the left retrosplenial region, the anterior portion of the right middle temporal gyrus, the right cingulate gyrus, and the bilateral cerebellum (Table 3, Figure 5). The important point to note is that the activation of the posterior portion of the inferior temporal gyrus and the fusiform gyrus was observed in all results after every subtraction, that is, simple minus chord, complex minus chord, and complex minus simple condition. The opposite subtraction of chord minus simple, chord minus complex, and simple minus complex conditions revealed almost the same activation pattern. The activation was observed at the bilateral orbital frontal cortex, the bilateral or left superior frontal gyrus, and the right superior temporal gyrus (Tables 4–6, Figures 6–8).

4. Discussion

The findings of this experiment are summarized as follows: as an accompaniment became more complex, (i) the subjects felt that the sound of music was richer and (ii) the fusiform

TABLE 1: Regions showing significant changes in rCBF by the subtraction of simple minus chord condition.

Anatomical structures	Brodmann area	Talairach coordinate			z-score
		x	y	z	
Posterior portion of inferior temporal gyrus	37				
L		-48	-55	-16	5.20
Fusiform gyrus					
L	18/19	-19	-58	-9	4.38
R	19	30	-67	-7	4.32
Medial surface of occipital lobe	17/18				
L		-17	-87	-4	3.83
R		10	-87	11	4.024
Superior parietal lobule	7				
R		33	-53	56	3.32
Orbital frontal cortex	11				
L		-17	50	-16	3.08

Coordinates x , y , and z are in millimetres corresponding to the atlas of Talairach and Tournoux. The x -coordinate refers to medial-lateral position relative to midline (negative = left); y -coordinate refers to anterior-posterior position relative to the anterior commissure (positive = anterior); z -coordinate refers to superior-inferior position relative to the anterior commissure-posterior commissure line (positive = superior). z -score refers to the maximum pixel of the region. L and R refer to the left and right hemisphere, respectively.

TABLE 2: Regions showing significant changes in rCBF by the subtraction of complex minus chord condition.

Anatomical structures	Brodmann area	Talairach coordinate			z-score
		x	y	z	
Posterior portion of inferior temporal gyrus	37				
L		-53	-58	-11	4.97
Fusiform gyrus					
L	19/37	-30	-49	-11	3.58
Medial surface of occipital lobe					
R	17	17	-96	2	3.10
Lateral surface of occipital lobe					
L	18	-39	-73	-2	3.00
Anterior portion of middle temporal gyrus					
R	38	35	8	-40	3.96

Details as for Table 1.

TABLE 3: Regions showing significant changes in rCBF by the subtraction of complex minus simple condition.

Anatomical structures	Brodmann area	Talairach coordinate			z-score
		x	y	z	
Posterior portion of inferior temporal gyrus	37				
L		-60	-58	-7	3.81
Fusiform gyrus					
L	36	-33	-26	-25	4.27
Retrosplenial region					
L	29	-6	-51	18	3.91
Anterior portion of middle temporal gyrus					
R	38	37	8	-40	3.44
Cingulate gyrus					
R	31	10	-28	40	3.29
Cerebellum					
L		-51	-49	-38	3.34
R		39	-64	-32	3.34

Details as for Table 1.

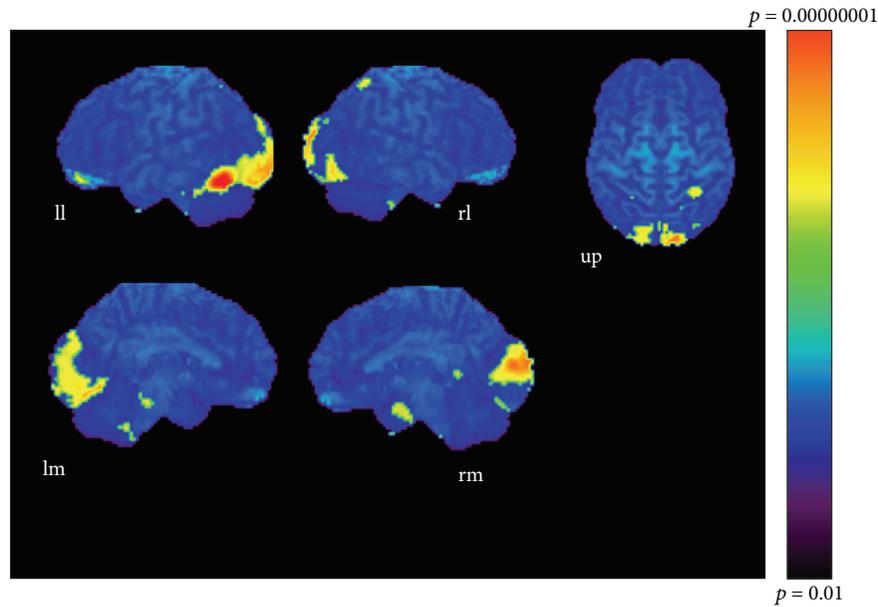


FIGURE 3: Simple-chord condition ($p < 0.001$). ll: lateral surface of left hemisphere; lm: medial surface of left hemisphere; rl: lateral surface of right hemisphere; rm: medial surface of right hemisphere; up: upper surface.

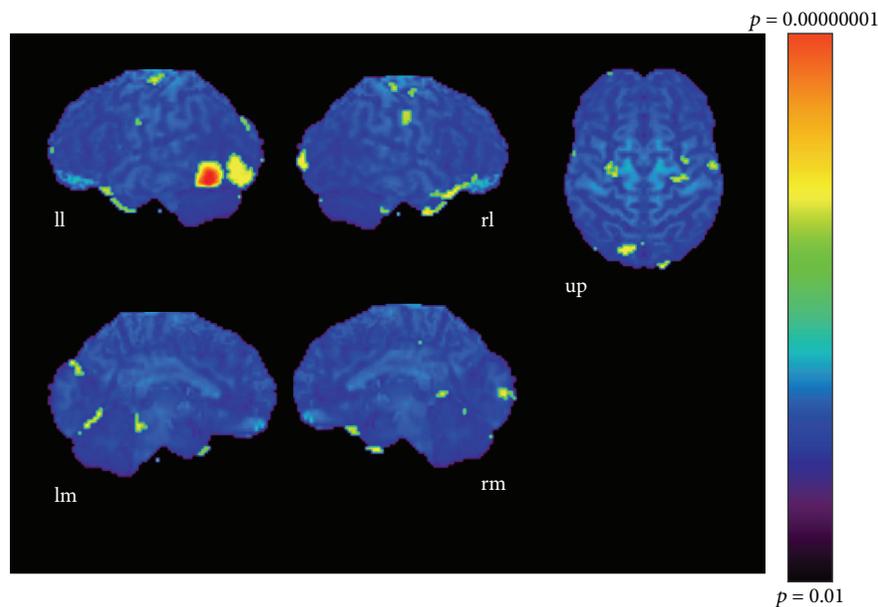


FIGURE 4: Complex-chord condition ($p < 0.001$). ll: lateral surface of left hemisphere; lm: medial surface of left hemisphere; rl: lateral surface of right hemisphere; rm: medial surface of right hemisphere; up: upper surface.

cortex and the posterior portion of the inferior temporal gyrus were activated. In the following paragraphs, we discuss the functional significance of these activated brain regions.

The fusiform cortex might participate in the perception of sound richness. The present study showed that, as the sound became richer, the activation of the fusiform cortex increased. This finding revealed that the degree of the activation of the fusiform cortex was different depending on the degree of the sound richness of the accompaniment in the identical

melodies. It is generally accepted that the fusiform cortex processes color recognition, based on the results of a case [16] and a PET activation study [17]. The findings of the present study and previous reports suggest that color information in vision and sound richness in audition might be similarly registered in the brain. In other words, it is possible that similar information from different sensory modalities might be processed within the same brain region and that the visual association cortex might not only be involved in visual

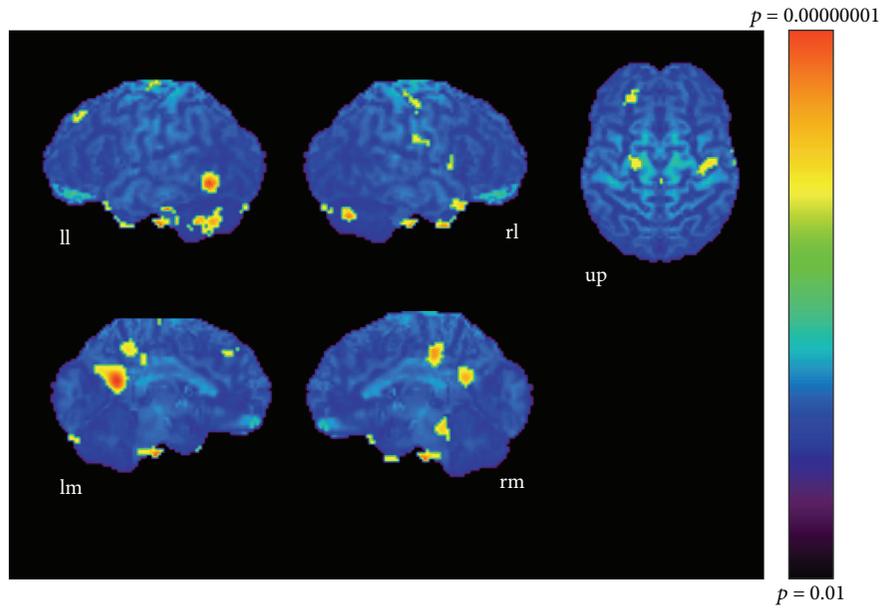


FIGURE 5: Complex-simple condition ($p < 0.001$). ll: lateral surface of left hemisphere; lm: medial surface of left hemisphere; rl: lateral surface of right hemisphere; rm: medial surface of right hemisphere; up: upper surface.

TABLE 4: Regions showing significant changes in rCBF by the subtraction of chord minus simple condition.

Anatomical structures	Brodmann area	Talairach coordinate			z-score
		x	y	z	
Orbital frontal cortex	11				
L		-6	26	-16	2.97
R		5	24	-14	2.73
Superior frontal gyrus	6/8				
L		-24	30	43	3.13
R		21	19	58	3.29
Superior temporal gyrus					
R	22	51	3	2	3.06
Cerebellum					
L		-17	-62	-36	3.82
R		51	-46	-36	3.43

Details as for Table 1.

TABLE 5: Regions showing significant changes in rCBF by the subtraction of chord minus complex condition.

Anatomical structures	Brodmann area	Talairach coordinate			z-score
		x	y	z	
Orbital frontal cortex	10/11				
L		-3	26	-18	5.39
R		12	64	-11	3.20
Superior frontal gyrus					
L	8	-24	28	50	3.05
Superior temporal gyrus					
R	22	62	-37	7	3.75

Details as for Table 1.

TABLE 6: Regions showing significant changes in rCBF by the subtraction of simple minus complex condition 3.

Anatomical structures	Brodmann area	Talairach coordinate			z-score
		x	y	z	
Orbital frontal cortex	11				
L		-1	28	-20	4.79
R		12	32	-18	4.06
Anterolateral portion of superior frontal gyrus					
L	21/22	-48	5	-14	2.97
Superior temporal gyrus					
R	22	62	-37	7	2.90
Cerebellum					
L		-51	-37	-25	3.17

Details as for Table 1.

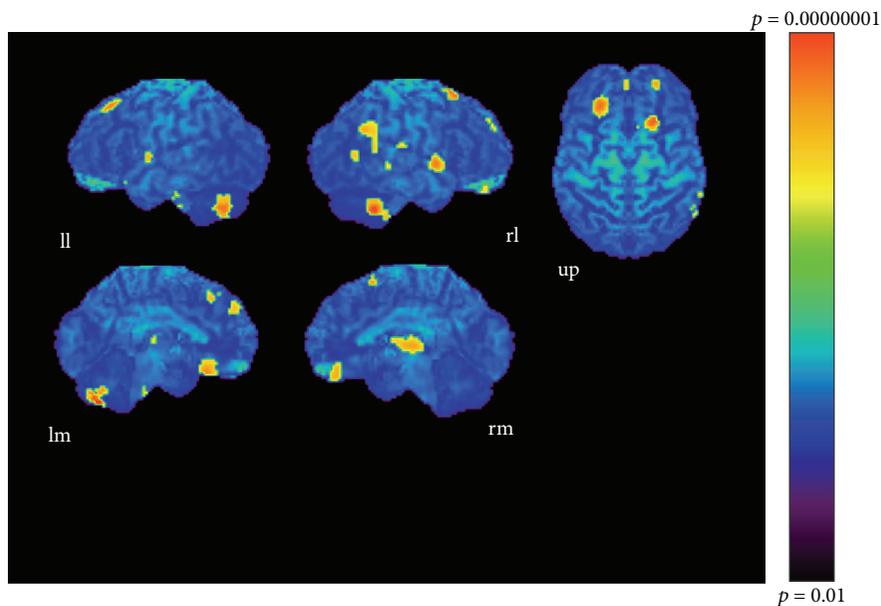


FIGURE 6: Chord-simple condition ($p < 0.001$). ll: lateral surface of left hemisphere; lm: medial surface of left hemisphere; rl: lateral surface of right hemisphere; rm: medial surface of right hemisphere; up: upper surface.

processing. Recent studies have revealed that some sensory modalities are related to each other. This phenomenon is called “cross-modal integration” and was observed between taste and audition [18], taste and smell [19–22], taste and color [23], odor and color [24], taste and music [25], pitch and visual size [26, 27], brightness and frequency of vibrotactile stimuli [28], sound and color [29, 30], and vision and audition [31]. It was reported that cross-modal associations are ubiquitously present in normal mental function [25, 32, 33]. Recent research suggests that cortical auditory processing is divided into separate processing streams [31, 34]. Posterior temporoparietal regions, labeled the “where” or “how” stream, may be specialized for processing sound motion and location [31]. Regions anterior and ventral to primary auditory cortex, labeled the “what” stream, may be specialized for processing characteristic auditory features [31]. Neurons in “what” stream respond directly to auditory and visual sensory stimuli and are important for forming the association

between auditory and visual objects [31]. Therefore, we may conclude that cross-modal integration also occurs at the fusiform cortex between color and sound richness when listening to music.

In the present study, the posterior portion of the inferior temporal gyrus was also activated. This area is called the lateral occipital complex (LOC) and is known to participate in foreground and background segregation in vision [35]. It was suggested that the LOC also participates in the discrimination between melody and its accompaniment [2]. In our previous study, we considered that the LOC might play a similar role of foreground and background segregation in both vision and audition. This finding reinforced the hypothesis that some association cortices carry out a similar function beyond the differences in sensory modalities (Figure 9). After the perception of sounds at the auditory cortex level, the information might be sent to the LOC and fusiform cortex. The former and the latter might participate in the foreground

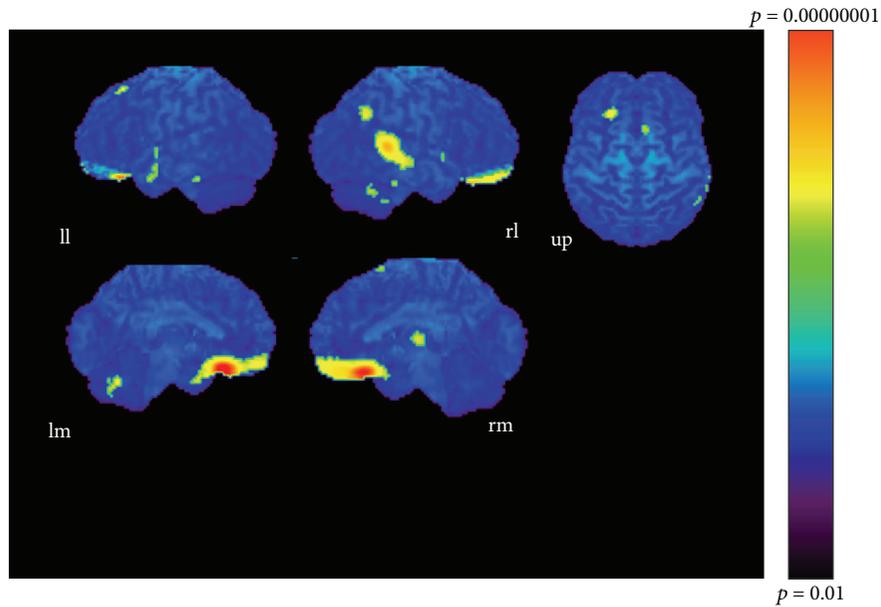


FIGURE 7: Chord-complex condition ($p < 0.001$). ll: lateral surface of left hemisphere; lm: medial surface of left hemisphere; rl: lateral surface of right hemisphere; rm: medial surface of right hemisphere; up: upper surface.

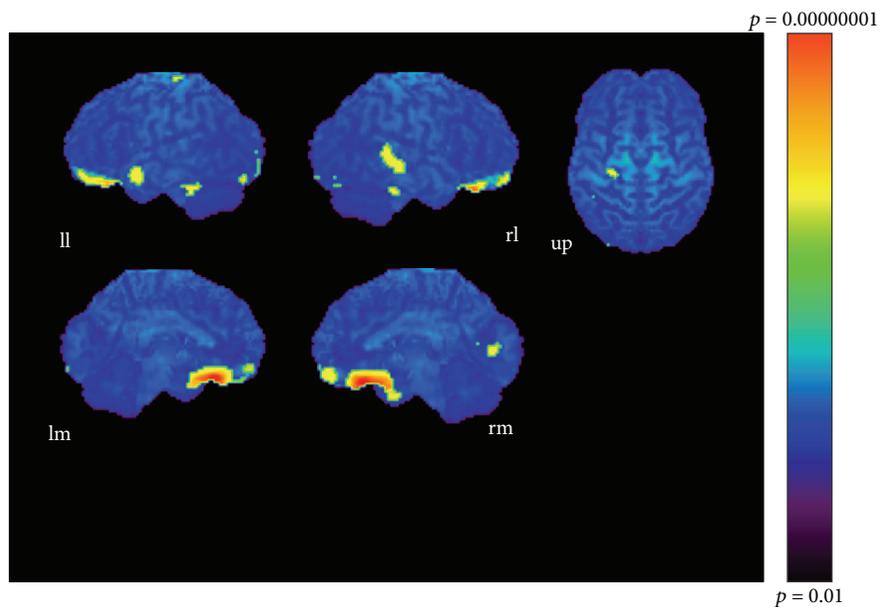


FIGURE 8: Simple-complex condition ($p < 0.001$). ll: lateral surface of left hemisphere; lm: medial surface of left hemisphere; rl: lateral surface of right hemisphere; rm: medial surface of right hemisphere; up: upper surface.

and background segregation and the recognition of sound richness, respectively, both in vision and audition.

The opposite subtraction, namely, chord minus simple, chord minus complex, and simple minus complex condition, all produced an activation of the bilateral orbital frontal cortex. The functional significance of this region in this experiment is unclear. However, this region is known as a structure within Yakovlev's circuit that participates in emotion and memory. Damage to this region often results in

disinhibition, impairment in control over impulsive behavior based on instinct and emotion. It is possible that activation of the orbital frontal cortex was caused by the comfortable and pleasant feeling of listening to familiar nursery songs or by inhibiting the desire to sing along with these familiar melodies.

In summary, the fusiform cortex and the LOC might have a similar function in vision and audition. The fusiform cortex recognizes color and sound richness, and the LOC

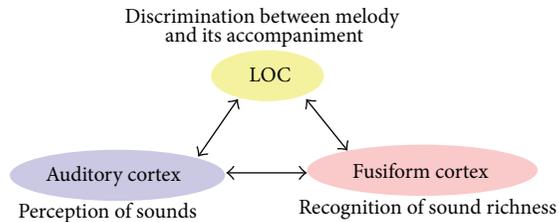


FIGURE 9: Diagram of cognitive processing during listening to music with accompaniment.

participates in foreground and background segregation. We may conclude that the association cortices might play a similar role across multiple sensory modalities. Further studies are needed to clarify the multimodal integration of association cortices.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

References

- [1] S. Sadie, *The Grove Concise Dictionary of Music*, The Macmillan Press, London, UK, 1994.
- [2] M. Satoh, K. Takeda, K. Nagata, and H. Tomimoto, "The lateral occipital complex is activated by melody with accompaniment: foreground and background segregation in auditory processing," *Journal of Behavioral and Brain Science*, vol. 1, no. 3, pp. 94–101, 2011.
- [3] S. M. Town and J. K. Bizley, "Neural and behavioral investigations into timbre perception," *Frontiers in Systems Neuroscience*, vol. 7, article 88, 2013.
- [4] R. J. Zatorre, A. R. Halpern, D. W. Perry, E. Meyer, and A. C. Evans, "Hearing in the mind's ear: a PET investigation of musical imagery and perception," *Journal of Cognitive Neuroscience*, vol. 8, no. 1, pp. 29–46, 1996.
- [5] F. Mirz, C. B. Pedersen, K. Ishizu et al., "Positron emission tomography of cortical centers of tinnitus," *Hearing Research*, vol. 134, no. 1-2, pp. 133–144, 1999.
- [6] D. Klein, R. J. Zatorre, B. Milner, and V. Zhao, "A cross-linguistic PET study of tone perception in Mandarin Chinese and English speakers," *NeuroImage*, vol. 13, no. 4, pp. 646–653, 2001.
- [7] B. Grison, *Une etude sur les alterations musicales au cours des lesions hemispheriques [M.S. thesis]*, 1972, (Cited by A. L. Benton, "The amusias" in: M. Critchley and R. A. Henson Eds, *Music and the Brain*, William Heinemann Medical Books Limited, London, UK, 378–397, 1977).
- [8] H. Hoshina and H. Tabata, *The Enjoyment of the Application of Harmony*, Ongaku-no-Tomo-sha, Tokyo, Japan, 1985, (Japanese).
- [9] FINALE 2004 FOR WINDOWS, MakeMusic, Coda Music Technology, CAMEO Interactive, 2004.
- [10] M. Satoh, K. Takeda, K. Nagata, J. Hatazawa, and S. Kuzuhara, "Activated brain regions in musicians during an ensemble: a PET study," *Cognitive Brain Research*, vol. 12, no. 1, pp. 101–108, 2001.
- [11] M. Satoh, K. Takeda, K. Nagata, J. Hatazawa, and S. Kuzuhara, "The anterior portion of the bilateral temporal lobes participates in music perception: a positron emission tomography study," *American Journal of Neuroradiology*, vol. 24, no. 9, pp. 1843–1848, 2003.
- [12] M. Satoh, K. Takeda, K. Nagata, E. Shimosegawa, and S. Kuzuhara, "Positron-emission tomography of brain regions activated by recognition of familiar music," *American Journal of Neuroradiology*, vol. 27, no. 5, pp. 1101–1106, 2006.
- [13] I. Kanno, H. Iida, S. Miura et al., "A system for cerebral blood flow measurement using an $H_2^{15}O$ autoradiographic method and positron emission tomography," *Journal of Cerebral Blood Flow and Metabolism*, vol. 7, no. 2, pp. 143–153, 1987.
- [14] S. Minoshima, R. A. Koeppe, J. A. Fessler et al., "Integrated and automated data analysis method for neuronal activation studies using [O-15] water PET," in *Quantification of Brain Function, Tracer Kinetics and Image Analysis in Brain PET*, K. Uemura, N. A. Lassen, T. Jones, and I. Kannno, Eds., pp. 409–417, Excerpta Medica/Elsevier, Amsterdam, The Netherlands, 1993.
- [15] J. Talairach and P. Tournoux, *Co-Planar Stereotaxic Atlas of the Human Brain*, Thieme, New York, NY, USA, 1988.
- [16] A. Damasio, T. Yamada, H. Damasio, J. Corbett, and J. McKeel, "Central achromatopsia: behavioral, anatomic, and physiologic aspects," *Neurology*, vol. 30, no. 10, pp. 1064–1071, 1980.
- [17] M. Corbetta, F. M. Miezin, S. Dobmeyer, G. L. Shulman, and S. E. Petersen, "Attentional modulation of neural processing of shape, color, and velocity in humans," *Science*, vol. 248, no. 4962, pp. 1556–1559, 1990.
- [18] J. Simner, C. Cuskley, and S. Kirby, "What sound does that taste? Cross-modal mappings across gustation and audition," *Perception*, vol. 39, no. 4, pp. 553–569, 2010.
- [19] M. Auvray and C. Spence, "The multisensory perception of flavor," *Consciousness and Cognition*, vol. 17, no. 3, pp. 1016–1031, 2008.
- [20] J. Djordjevic, R. J. Zatorre, and M. Jones-Gotman, "Odor-induced changes in taste perception," *Experimental Brain Research*, vol. 159, no. 3, pp. 405–408, 2004.
- [21] D. M. Small and J. Prescott, "Odor/taste integration and the perception of flavor," *Experimental Brain Research*, vol. 166, no. 3-4, pp. 345–357, 2005.
- [22] R. J. Stevenson and C. Tomiczek, "Olfactory-induced synesthasias: a review and model," *Psychological Bulletin*, vol. 133, no. 2, pp. 294–309, 2007.
- [23] M. O'Mahony, "Adapting short cut signal detection measures to the problem of multiple difference testing: the R-Index," in *Sensory Quality in Foods and Beverages*, A. A. Williams and R. K. Atkin, Eds., pp. 69–80, Chapman, London, UK, 1983.
- [24] M. L. Demattè, D. Sanabria, and C. Spence, "Cross-modal associations between odors and colors," *Chemical Senses*, vol. 31, no. 6, pp. 531–538, 2006.
- [25] B. Mesz, M. A. Trevisan, and M. Sigman, "The taste of music," *Perception*, vol. 40, no. 2, pp. 209–219, 2011.
- [26] K. K. Evans and A. Treisman, "Natural cross-modal mappings between visual and auditory features," *Journal of Vision*, vol. 10, no. 1, 2010.
- [27] C. Parise and C. Spence, "Synesthetic congruency modulates the temporal ventriloquism effect," *Neuroscience Letters*, vol. 442, no. 3, pp. 257–261, 2008.
- [28] G. Martino and L. E. Marks, "Cross-modal interaction between vision and touch: the role of synesthetic correspondence," *Perception*, vol. 29, no. 6, pp. 745–754, 2000.

- [29] V. S. Ramachandran and E. M. Hubbard, "Hearing colors, tasting shapes," *Scientific American*, vol. 288, no. 5, pp. 52–59, 2003.
- [30] J. Ward, B. Huckstep, and E. Tsakanikos, "Sound-colour synaesthesia: to what extent does it use cross-modal mechanisms common to us all?" *Cortex*, vol. 42, no. 2, pp. 264–280, 2006.
- [31] M. S. Beauchamp, K. E. Lee, B. D. Argall, and A. Martin, "Integration of auditory and visual information about objects in superior temporal sulcus," *Neuron*, vol. 41, no. 5, pp. 809–823, 2004.
- [32] E. M. Hubbard and V. S. Ramachandran, "Neurocognitive mechanisms of synesthesia," *Neuron*, vol. 48, no. 3, pp. 509–520, 2005.
- [33] R. E. Cytowic and D. M. Eagleman, *Wednesday is Indigo Blue: Discovering the Brain of Synesthesia*, MIT Press, Cambridge, Mass, USA, 2009.
- [34] J. P. Rauschecker and B. Tian, "Mechanisms and streams for processing of 'what' and 'where' in auditory cortex," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 97, no. 22, pp. 11800–11806, 2000.
- [35] K. Grill-Spector, T. Kushnir, S. Edelman, Y. Itzhak, and R. Malach, "Cue-invariant activation in object-related areas of the human occipital lobe," *Neuron*, vol. 21, no. 1, pp. 191–202, 1998.

Research Article

Music Perception Influences Language Acquisition: Melodic and Rhythmic-Melodic Perception in Children with Specific Language Impairment

Stephan Sallat^{1,2} and Sebastian Jentschke³

¹Justus-Liebig-Universität, Gießen, Germany

²Speech and Language Pedagogy and Pathology, Department of Special and Inclusive Education, Faculty of Education, University of Erfurt, Nordhäuser Strasse 63, 99089 Erfurt, Germany

³Freie Universität Berlin, Cluster “Languages of Emotion”, Habelschwerdter Allee 45, 14195 Berlin, Germany

Correspondence should be addressed to Stephan Sallat; stephan.sallat@uni-erfurt.de

Received 24 April 2015; Revised 3 August 2015; Accepted 6 August 2015

Academic Editor: Enzo Emanuele

Copyright © 2015 S. Sallat and S. Jentschke. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Language and music share many properties, with a particularly strong overlap for prosody. Prosodic cues are generally regarded as crucial for language acquisition. Previous research has indicated that children with SLI fail to make use of these cues. As processing of prosodic information involves similar skills to those required in music perception, we compared music perception skills (melodic and rhythmic-melodic perception and melody recognition) in a group of children with SLI ($N = 29$, five-year-olds) to two groups of controls, either of comparable age ($N = 39$, five-year-olds) or of age closer to the children with SLI in their language skills and about one year younger ($N = 13$, four-year-olds). Children with SLI performed in most tasks below their age level, closer matching the performance level of younger controls with similar language skills. These data strengthen the view of a strong relation between language acquisition and music processing. This might open a perspective for the possible use of musical material in early diagnosis of SLI and of music in SLI therapy.

1. Introduction

Despite the complexity of language, most children successfully acquire the capacity to perceive and comprehend it, as well as produce spoken utterances. However, a considerable number of two-year-old children are delayed in crucial aspects of language acquisition such as vocabulary, grammar, or correct articulation. Whereas half of them make up this delay by about three years of age, approximately seven percent of an age cohort continue to have difficulties to acquire their native language without any obvious primary dysfunctions (such as mental, neurological, sensorial, oral-motor disorders; cf. [1]). These children, diagnosed with Specific Language Impairment (SLI), show deficits mainly in grammar processing (e.g., morphosyntax), phonology, and word learning. Because of their impaired language, they form a high-risk group for problems in school, as well as in other cognitive or social-emotional areas [2–5].

There is discussion regarding risk factors and possible causes for SLI in the literature [6–8]; for an overview, see [3]. However, none of these accounts cover all aspects of impaired linguistic and nonlinguistic functions in people with SLI. An interdisciplinary perspective might help to gain further insight into the aetiology of SLI. Such an approach, a comparison of speech and music perception in children with SLI and those with typical language development, is presented in this paper.

Several theoretical accounts proposed that, particularly during early language acquisition, language is rather perceived as music. For example, Koelsch ([9], p. 16) hypothesized that “the human brain, particularly at an early age, does not treat language and music as strictly separate domains, but rather treats language as a special case of music.” Brandt et al. ([10], p. 5) denoted “that music has a privileged status that enables us to acquire not only the musical conventions

of our native culture, but also enables us to learn our native language.” In addition, music and language share a number of similarities (for overviews, see [9–12]). Both are based on acoustic information, involving a limited number of categorical elements or classes (phonemes and tones) that are organized in structured sequences according to specific regularities. These regularities are acquired using similar learning mechanisms [13]. There are indicators for common evolutionary origins [14, 15]. Electrophysiological evidence suggests shared cognitive resources and similar underlying neural substrates for processing semantics [16], syntax [17, 18], and prosody [19, 20].

Prosody is presumably the area with the strongest overlap: Prosodic or suprasegmental features can be regarded as “musical” aspects of the speech signal. Prosody has many functions, such as indicating the emotional state of the speaker, indicating syntactic structure, or indicating cues for turn-taking in conversations. Components of prosody (such as speech rhythm, speech melody, contour, timbre, pauses, and stress) emerge from a combination of acoustic features such as pitch/frequency, loudness/intensity, duration, and timbre [21]. During language acquisition, these prosodic components help the infant to detect word and phrase boundaries. This enables them to acquire regularities about the arrangement of linguistic patterns like phonemes, words, and phrases (for reviews, see [22, 23]). Unlike typically developing children, children with SLI appear not to profit from exaggerated prosody (contour, stress, and pauses) while learning words and grammatical rules [24, 25]. Related to music perception and prosody is a group of theories which propose deficient auditory processing (especially fine-grained temporal processing of auditory information) to account for the problems of children with SLI [26–29]. However, other authors failed to observe such problems [30–32]. More recently, the focus shifted towards auditory features which are crucial for the processing prosodic cues: Corriveau et al. [33] proposed that the accurateness of prosodic processing in SLI is impaired because the children fail to use auditory cues required for the perception of rhythm and stress, which has consequences for their language development. Przybylski et al. [34] provided evidence that children with SLI, and dyslexia, as well as controls with typical language development show better performance in grammaticality judgements after rhythmically regular than after irregular prime sequences. Although the performance level of the clinical groups was generally lower than that of the controls, they still profit from the metrical structure of the regular prime. Recently, Cumming et al. [35] proposed in their “prosodic phrasing” hypothesis that problems with processing certain acoustic properties (particularly amplitude rise time and duration), relevant for both language and music perception, may be responsible for morphosyntactical problems in children with SLI. They also observed that children with SLI were less sensitive to all auditory measures explored in the study.

In order to perceive music, children have to acquire implicit knowledge about musical structure. Two key aspects are pitch and temporal organization (cf. [18, 19]). In addition, children acquire explicit knowledge, for example, the tune of a particular song and its lyrics. A number of studies explored

music perception skills in children with SLI: Jentschke et al. [36] demonstrated that five-year-old children with SLI are impaired in certain aspects of music perception, namely, in that they lack a neurophysiological marker of music-syntactic processing whereas this marker can be observed in children with typical language development. Other studies investigated music production skills in SLI. Corriveau and Goswami [37] showed that ten-year-old children with SLI were impaired in rhythmic tapping to an externally paced source but less to an internally generated rhythm and that the severity of impairment was linked to language and literacy outcomes. Recently, Clément et al. [38] investigated singing in eleven-year-old children with SLI. Compared to children of the same age but with typical language development, children with SLI were poorer in reproducing similar tunes (pitch matching) and in reproducing familiar and unfamiliar melodies. Based upon their findings, a general auditory-motor dysfunction in the children with SLI was proposed.

The present study aimed to add knowledge of whether children with SLI differ from typically developing children with regard to music perception: Skills in pitch organization were explored in a melodic perception task, those of temporal organization were explored in a rhythmic-melodic perception task, and the recognition of musical sequences stored in long-term memory was explored in a melody recognition task. Familiarity and features like tempo, sound, and pitch of these melodies were manipulated. We expected that children with SLI would, similar to their performance in the language domain, lag behind their age-matched peers and perform rather like younger children with typical language development. Such pattern would indicate a relation of music perception and linguistic skills in children with SLI and provide further evidence for a privileged status of music perception skills during language acquisition.

2. Methods

2.1. Participants. Children with specific language impairment (SLI) were compared to a control group with typical language development of the same age (controls with comparable age (CA)) and a group of younger children, whose linguistic abilities were comparable to the SLI group (younger controls with comparable language skill (CL)). Written informed consent was obtained from the parents of all participating children. A questionnaire, providing information about language development, social background, and musical environment of their children, revealed that groups did not differ in variables reflecting social (e.g., number of books or CDs in the household) or musical (e.g., family members playing an instrument; amount of singing with children) background. To determine the socioeconomic background of the children’s families, the occupation of the parents was classified in terms of the “International Standard Classification of Occupation 1988” [39]. This classification was then transformed into “International Socio-Economic Index of Occupational Status” values (ISEI [40]) to provide a status measure for each occupation. There was neither a group difference for duration of school education (the vast majority of all parents attended school for 10 years) nor a group

TABLE 1: Summary of characteristics of the three groups of participants (SLI: Children with Specific Language Impairment; CA: children of comparable age; CL: children with comparable linguistic abilities) encompassing age and gender distribution, performance in psychometric tests (language screening and general cognitive skills), and variables reflecting the socioeconomic background. (*T*): *T*-scores, (*S*): standard scores, (*R*): raw scores.

	SLI (<i>N</i> = 29)	CA (<i>N</i> = 39)	CL (<i>N</i> = 13)
Age (in months)	64.2 (56–71)	63.6 (57–71)	51.3 (48–55)
Gender (male/female)	19/10	22/17	8/5
Psychometric tests			
Language comprehension (<i>T</i>)	38.9 (20–59)	51.8 (39–72)	54.5 (40–74)
Language production (<i>T</i>)	39.5 (21–79)	56.3 (40–79)	58.2 (45–72)
Nonword repetition (<i>T</i>)	31.5 (20–61)	53.0 (35–68)	56.4 (35–70)
Sentence repetition (<i>T</i>)	37.3 (20–58)	55.0 (41–74)	55.3 (39–63)
Phoneme discrimination (<i>R</i>)	9.8 (3–17)	12.6 (7–17)	11.8 (6–16)
Nonverbal IQ (<i>S</i>)	91.8 (81–115)	101.8 (84–120)	103.1 (85–125)
Socioeconomic status			
Mother's occupation (ISEI)	34.2 (16–52)	42.8 (25–66)	50.4 (29–73)
Father's occupation (ISEI)	34.5 (23–71)	37.4 (19–69)	39.4 (29–73)
Children's books	45.4 (1–207)	46.4 (4–138)	42.5 (1–102)
Children's CDs	25.2 (0–88)	21.6 (5–55)	23.3 (7–70)

difference for professional qualification. However, whereas fathers' occupation was similar among groups ($p > 0.50$), the status value of mothers' occupation was lower in the children with SLI compared to both control groups ($p < 0.01$). An overview is given in Table 1.

Children with SLI were recruited from a kindergarten for special education (for children with language and speech disorders); children with typical language development (CA, CL) came from four public kindergartens. All kindergartens were located in Leipzig, Germany. Children (of any group) were excluded from the study when [a] their parents or teachers reported defective hearing or a history of hearing disease, [b] they did not grow up in monolingual families, [c] they had any other speech or language disorder such as oral fluency disorder (e.g., stuttering), [d] they had any other condition (mutism, autism, etc.), or [e] they had a nonverbal IQ below the low average range (i.e., less than 80 IQ points). The data of 29 children with SLI (4;8 to 5;11 years old, $M = 5;4$ years), 39 CA children (4;9 to 5;11 years old, $M = 5;3$ years), and 13 CL children (4;0 to 4;7 years old, $M = 4;3$ years) were evaluated. Like many previous studies (see, e.g., [3]), a higher incidence of SLI was observed in boys (65.5%). We tried to match this proportion in the control groups (CA: 56.4%; CL: 61.5%). The results for the language screening within the SLI group corresponded to the norms of an age group between 3;0 and 3;6 years. However, younger controls with comparable language skills (CL) were only about one year younger than the SLI children, because for younger children the music perception tasks would have been too difficult. The performance of these controls (i.e., their raw scores) in all subtests of the language screening was significantly above the SLI children ($p < 0.030$), indicating a higher semantic-lexical and morphosyntactic knowledge. Whereas the linguistic performance within the CL group lies between the SLI and the CA group, the SLI and the CL group had

similar performance levels at phonological analysis (assessed by phoneme discrimination skills).

2.2. Stimuli and Paradigm. All measurements were acquired in either four (CA group) or five (SLI and CL group) experimental sessions of approximately 20- to 25-minute length. One session contained the language screening; in another nonverbal intelligence was tested. Linguistic skills were investigated with a standardized German language screening for three- to five-year-old children (SETK 3–5; [41]). It contained three parts: language comprehension (manipulation tasks with different objects), speech production (applying morphological rules to words and nonwords), and working memory (repetition of nonwords and sentences). In addition, phonemic discrimination (e.g., “Kanne” [pitcher] versus “Tanne” [fir]) was tested. Nonverbal intelligence was assessed using the Kaufman Assessment Battery for Children (K-ABC [42]).

In the remaining sessions, musical skills were evaluated, using tasks developed by the authors of this study. The tasks explored melodic perception, rhythmic-melodic perception, and melody recognition. Stimuli were created as MIDI files containing the beginning phrases of nursery rhymes (proposed by the kindergarten teachers and well known to all participating children). The MIDI files were exported into WAV files with a piano sound (using Steinberg Cubase SX and The Grand, Steinberg Media Technology, Hamburg, Germany). These were presented on a laptop, using Presentation 0.76 (Neurobehavioral Systems, Inc., Albany, CA), which was also used to record the answers. The suitability and age-appropriateness of stimuli and procedure were checked in a pilot test with 10 typically developing children. The pilot test also served to determine an optimal speed to present the stimuli sequences (135 beats per minute for melodic perception; 120 beats per minute for rhythmic-melodic perception).

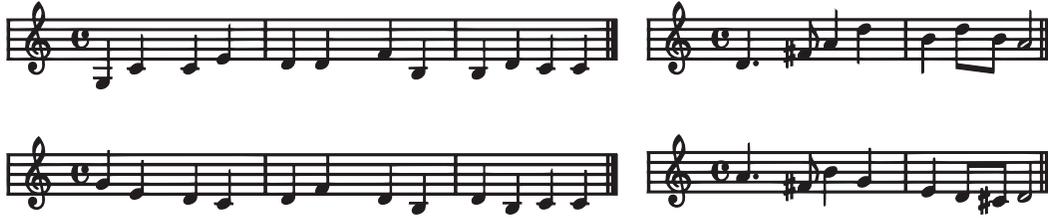


FIGURE 1: Overview of the stimuli used in the melodic perception (left) and rhythmic-melodic perception (right) tasks. The nursery rhyme used in the melodic perception tasks had the title “Es war eine Mutter, die hatte vier Kinder,” the one for the rhythmic-melodic perception tasks had the title “Alle Vögel sind schon da.” Each task had three conditions: standard (first line), transposed (first line, but played in another key), and comparable (bottom line).

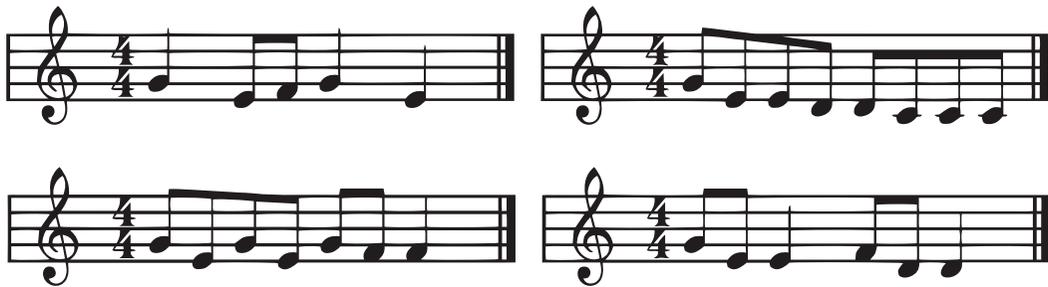


FIGURE 2: Overview of the stimuli (nursery rhymes) used in the melody recognition task. The stimuli were the nursery rhymes “Hänsel und Gretel” (first row, left), “Der Kuckuck und der Esel” (second row, left), “Weil heute Dein Geburtstag ist” (first row, right), and “Hänschen klein” (second row, right). These stimuli were known to all children.

Children sat in front of a laptop with two external speakers and listened to the stimuli. Before the experiment, children listened to the beginnings of the nursery rhymes four to six times and sang it two to four times with the experimenter. All tasks of the experiment were integrated in a game: Children had to help a cuddly toy (Paul, the forgetful rabbit) who played musical phrases, but could not remember if he played it correctly or incorrectly. Then they voted using two different buttons, whether the phrase was correct (unmodified) or incorrect (modified). In the melody recognition experiment, they detected which nursery rhyme was played. They gave their response with buttons placed next to a two-by-two array containing pictures representing the four different songs used in this task (see right part of Figure 2).

2.2.1. Melodic Perception. To test melodic perception, the beginning phrase of a well-known nursery rhyme with 12-tone length and a constant rhythm (containing only quarter notes) was used (see Figure 1, left panel). There were three different blocked conditions: standard, transposed, and comparable melody (described below). Within each condition, there were 20 stimuli: In 10 stimuli the melody was not modified; in another 10 stimuli it was changed: in 5 stimuli the tone height was altered while the contour was preserved and in another 5 both tone height and contour were changed. A response was counted correct, if the children detected whether the phrase was unmodified or changed (these two choices were represented by different buttons). Each block lasted about 5 minutes and the whole experiment about 20 minutes. In the standard melody block, the phrase was presented in the

original key. In the transposed melody block the melody phrase was presented in five different keys (either original [as in the first block] or one or two half tones up and down). Like in language acquisition, where words and sentences have to be recognized as invariant although they are spoken by different speakers, this task aimed to test the ability to recognize preservation or violation of a melody regardless of whether the key was altered or not. In the comparable melody block, a new melody, previously unknown to the children but similar in harmonic structure and length to the original melody, was introduced in order to test melodic perception while removing the opportunity to subvocally speak the text of the nursery rhyme. The task aimed at testing melodic perception and short-term memory for melodies and was similar to methods used in other tests on musical abilities or musical aptitude (e.g., [43–45]), where children have to compare melodies of which they do not have long-term representations. The novel melody for this task was learned before the experiment by singing it on the syllable [na].

2.2.2. Rhythmic-Melodic Perception. The beginning of another nursery rhyme was used in this task (see Figure 1, right panel). In contrast to the stimuli used in [1], this phrase had a complex rhythmic structure with eighth, quarter, and punctuated quarter notes. In 10 stimuli the original rhythm was kept, while in another 10 stimuli the rhythm was changed at different positions within the phrase (whereas the pitch height was kept constant). This was accomplished, for example, by changing two quarter notes into an eighth and a punctuated quarter note. Comparable to the methodology

in the melodic perception part, the stimuli were presented in three block conditions, standard, transposed, and comparable rhythm (similar to those described above). The task took the same amount of time as the first one (20 minutes).

2.2.3. Melody Recognition. In the melody recognition task, children had to distinguish the starting phrase of 4 different nursery rhymes (see Figure 2), which were familiar to all children. However, two members of the SLI group did not know all songs and were excluded from the melody recognition task. Each rhyme was represented by a picture arranged in a two-by-two array. Children indicated which song they heard by pressing a button placed next to the picture representing it. Stimuli were modified to create four conditions: They were either played with piano sound in the same key (original); with a piano sound, but in a different key (transposed); with different instrumental timbre (timbre-change); or at a faster tempo (faster). Using these parametric manipulations, we aimed to explore several acoustic features which constitute building blocks of prosodic components. In every condition, each song was played three times, resulting in a total of 12 stimuli in each condition (48 stimuli altogether). To familiarize the children with the task, the experimenter spoke the lyrics of the beginning phrase and children had to press the button placed next to the picture representing the rhyme.

2.3. Data Analysis. Using Kolmogorov-Smirnov tests, we ensured that the analysed variables conformed with a standard normal distribution (SLI: $M = 0.47$; $0.07 \leq p \leq 0.84$; CA: $M = 0.58$; $0.07 \leq p \leq 0.93$; CL: $M = 0.88$; $0.61 \leq p \leq 1.00$). However, given that variance was unequal in the three groups (with a relatively broad range of performance, particularly in children with SLI), we decided to use nonparametric tests to compare the three groups (Mann-Whitney U tests) and to explore relations between music perception skills and other variables (Spearman's rank correlations).

First, we compared the results of children with SLI in the language screening and their nonverbal intelligence to those of typically developing children. We also determined, using one-sample t -tests, whether the response probabilities were significantly above chance level. Then, we explored whether the experimental groups showed differences in music perception, comparable to the differences in their linguistic abilities. Finally, the performance in music perception tasks was related to that in the language comprehension and working memory subtests of the language screening and variables reflecting their socioeconomic background (given that the status values of their mother's profession could not be perfectly matched).

3. Results

Children with SLI performed significantly below their age-equivalent mean scores in all subtests of the language screening (see Table 1; language comprehension: $z = -1.11$; speech production: $z = -1.05$; nonword repetition: $z = -1.85$; sentence repetition: $z = -1.27$; $p < 0.001$). In contrast, the results of the two control groups were slightly above the mean

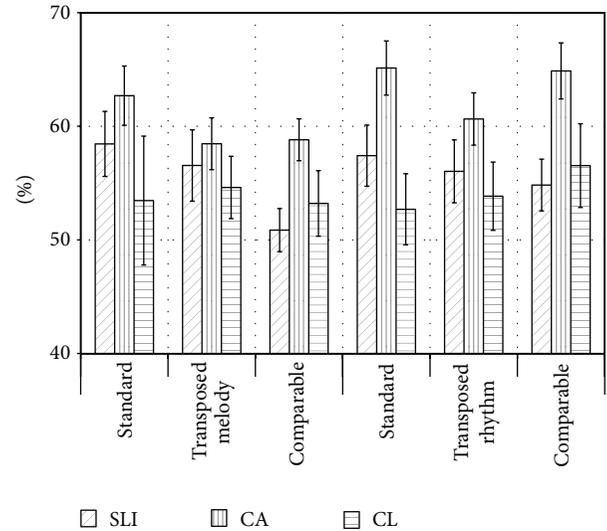


FIGURE 3: Mean percentage of correct responses (and standard error of mean) for the subtests of the melodic (left panel) and the rhythmic-melodic perception tasks (right panel).

of their population norm ($0.18 \leq z_{CA} \leq 0.63$, $0.45 \leq z_{CL} \leq 0.82$). Phonemic discrimination in children with SLI differed from those of comparable age ($p = 0.006$) but not from those with comparable language skills ($p = 0.154$). Although children with SLI had a similar distribution in the range of their IQ scores, nonverbal intelligence scores in children with SLI were below the two control groups ($M_{SLI} = 92.8$; $M_{CA} = 101.8$; $M_{CL} = 103.1$; $p < 0.01$).

In the melodic perception tasks, children with SLI performed nominally poorer than the age-matched controls (CA) in all conditions (see Figure 3 and Table 2). For the comparable melody condition and the sum score of all three conditions a significant difference between the SLI and the CA group was obtained. The scores of children with SLI were generally in the same range as those of the younger controls (CL). They were nominally slightly higher for the standard and the transposed melody condition. Neither between the children with SLI and the younger controls (CL; $p > 0.39$) nor between the two control groups ($p > 0.08$) significant differences were observed for any condition. The performance of children in the CA group was significantly above chance level in all conditions ($p \leq 0.001$), the performance of children with SLI did significantly exceed chance in all ($0.006 \leq p \leq 0.046$), but the comparable melody condition ($p = 0.654$), and in the CL group performance did not reach above chance levels in any condition ($p > 0.118$; presumably because of the small sample size; as their mean performance was relatively similar to the SLI children).

For the rhythmic-melodic perception tasks (see Figure 3 and Table 2), the CA group showed higher performance levels than the children with SLI in all conditions. These differences were significant for the standard and the comparable rhythm condition as well as for the sum score of all three conditions. Similarly, the CA group showed higher performance levels than the language controls (CL), and significant differences

TABLE 2: The first three columns contain mean percentages of correct responses for children with Specific Language Impairment (SLI), children of comparable age (CA), and children with comparable linguistic abilities (CL). In the last three columns significance levels of the group comparisons (using the Mann-Whitney U tests) are given. Results for melodic perception are in the top section, rhythmic-melodic perception in the middle, and melody recognition at the bottom. Significant group differences are written bold.

	Correct responses			Group comparison		
	SLI	CA	CL	SLI-CA	SLI-CL	CA-CL
Melodic perception						
Standard	58.45%	62.70%	53.45%	0.306	0.519	0.161
Transposed	56.55%	58.45%	54.60%	0.578	0.648	0.259
Comparable	50.86%	58.81%	53.21%	0.003	0.389	0.156
Sum score	53.85%	59.23%	52.82%	0.040	0.979	0.081
Rhythmic-melodic perception						
Standard	57.40%	65.15%	52.70%	0.014	0.451	0.006
Transposed	56.05%	60.65%	53.85%	0.067	0.872	0.135
Comparable	54.85%	64.85%	56.55%	0.004	0.519	0.106
Sum score	56.10%	63.55%	54.37%	0.003	0.591	0.010
Melody recognition						
Standard	44.17%	55.75%	59.58%	0.033	0.073	0.678
Faster	39.83%	46.33%	53.17%	0.356	0.187	0.558
Timbre-change	56.75%	65.17%	53.17%	0.205	0.732	0.148
Transposed	46.33%	51.25%	59.58%	0.420	0.089	0.293
Sum score	46.75%	54.65%	56.42%	0.110	0.177	0.866
Spoken (control)	93.83%	98.08%	96.17%	0.113	0.955	0.130

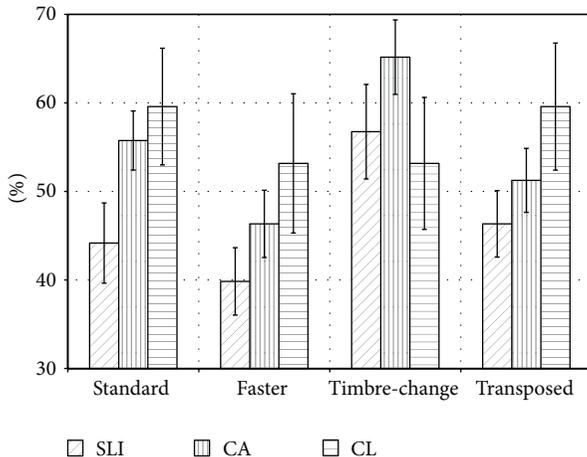


FIGURE 4: Mean percentage of correct responses (and standard error of mean) for the four subtests of the melody recognition task.

between those two groups were observed for the standard condition and the sum score of all conditions. The performance was significantly above chance for all conditions in children with SLI ($0.010 \leq p \leq 0.043$) and the CA group ($p \leq 0.001$) but not for any condition in the CL group ($p > 0.101$; presumably due to the small sample size).

In the melody recognition tasks (see Figure 4 and Table 2), for all conditions except the timbre-change condition, the younger controls (CL) achieved the highest levels of performance, followed by the age-matched controls whereas children with SLI achieved the lowest scores. For the standard

condition there was a significant group difference between children with SLI and those of comparable age. For no other condition did a group comparison reach significance. The performance in all groups was above chance level for all conditions ($p \leq 0.004$).

Using correlation analyses we explored the relationship between the “language comprehension” subtest of the language screening and music perception skills (see Table 3). When the whole group (see left column) is considered, all melody recognition tasks as well as the rhythmic-melodic perception tasks except from the standard condition were significantly correlated with the language comprehension subtest. None of the melody perception tasks were significantly correlated with the language comprehension performance. To ensure that the observed correlations were not simply due to the different skill levels in the three groups, additional correlation analyses were carried out for each group separately. Children with SLI (second column) had significant correlations for the transposed subtest of the rhythmic-melodic perception task, for the sum score of the rhythmic-melodic perception task, and for the faster, and the timbre-change condition as well as the sum score of the melody recognition task. For the control groups (third and fourth column), no correlations between musical and linguistic skills were significant.

Furthermore, we explored correlations among the sum scores of the three music perception tasks and between those sum scores and the results of the psychometric tests and indicators of the socioeconomic status of the family. We observed (see topmost part of Table 4) that the music perception skills

TABLE 3: Correlations of all music perception subtests with the language comprehension subtest of the language screening (significance level in parentheses) for all groups of children (first column) as well as each subgroup separately (second to fourth column). Significant correlations are indicated by bold typeface.

	All children ($N = 81/79$)		SLI ($N = 29/27$)		CA ($N = 39$)		CL ($N = 13$)	
Melodic perception								
Standard	0.10	(0.357)	0.06	(0.764)	0.03	(0.840)	0.17	(0.579)
Transposed	0.13	(0.242)	0.26	(0.169)	0.07	(0.684)	-0.03	(0.931)
Comparable	0.19	(0.097)	0.07	(0.715)	0.10	(0.559)	-0.14	(0.655)
Sum score	0.20	(0.075)	0.14	(0.474)	0.08	(0.628)	0.00	(0.993)
Rhythmic-melodic perception								
Standard	0.16	(0.166)	0.11	(0.578)	0.02	(0.908)	-0.06	(0.845)
Transposed	0.26	(0.017)	0.47	(0.010)	-0.03	(0.855)	0.23	(0.448)
Comparable	0.32	(0.004)	0.18	(0.354)	0.22	(0.189)	-0.06	(0.853)
Sum score	0.31	(0.005)	0.38	(0.044)	0.11	(0.511)	0.01	(0.964)
Melody recognition								
Standard	0.28	(0.014)	0.33	(0.091)	0.08	(0.628)	0.36	(0.227)
Faster	0.29	(0.011)	0.39	(0.043)	0.26	(0.113)	0.26	(0.383)
Timbre-change	0.40	(<0.001)	0.52	(0.005)	0.25	(0.118)	0.53	(0.065)
Transposed	0.25	(0.029)	0.28	(0.153)	0.22	(0.171)	0.35	(0.239)
Sum score	0.33	(0.003)	0.45	(0.019)	0.19	(0.236)	0.39	(0.188)

TABLE 4: Intercorrelations of the music perception subtests (top part), as well as correlations of the music perception subtests with the subtests of the language screening (middle part) and with indicators of the socioeconomic background of the participants. Correlations are reported for the whole group (SLI, CA, and CL), their significance level is shown in parentheses, and significant correlations are indicated by bold typeface.

	Melodic perception		Rhythmic-melodic perception		Melody recognition	
Melodic perception			0.64	(<0.001)	0.36	(0.001)
Rhythmic perception					0.45	(<0.001)
Language comprehension	0.20	(0.075)	0.31	(0.005)	0.33	(0.003)
Language production	0.11	(0.313)	0.24	(0.033)	0.29	(0.009)
Nonword repetition	0.32	(0.004)	0.42	(<0.001)	0.37	(0.001)
Sentence repetition	0.31	(0.002)	0.49	(<0.001)	0.41	(<0.001)
Hand movements	0.17	(0.139)	0.32	(0.004)	0.24	(0.034)
Spatial memory	0.20	(0.068)	0.24	(0.030)	-0.01	(0.944)
Mother's ISEI score	0.05	(0.692)	0.00	(0.972)	0.14	(0.261)
Father's ISEI score	-0.12	(0.374)	-0.04	(0.740)	0.04	(0.750)
Books in household	0.04	(0.731)	0.04	(0.735)	0.12	(0.315)
CDs in household	0.18	(0.121)	0.16	(0.183)	0.16	(0.190)

were moderately to strongly intercorrelated, indicating that they presumably measure common underlying skills.

Moreover, there are (mainly moderate) correlations between music perception skills and the short-term/working memory subtests of the language screening. No correlations were observed between either the "language comprehension" or the "language production" subtests and melodic perception, whereas those subtests correlated with the rhythmic-melodic perception and the melody recognition tasks. Both music perception tasks also correlated with the hand movements subtest, and rhythmic-melodic perception correlated with the spatial memory subtest of the intelligence test (which assess short-term/working memory). In contrast, no significant correlations were found between music perception skills

and measures of the socioeconomic status of the families. That is, regardless of the difference in the status value of the mother's profession, measures of socioeconomic status did not account for the difference in music perception skills in the examined population.

4. Discussion

The present study explored music perception in children with SLI and with typical language development in order to determine whether there is a link between speech perception and different aspects of music perception (melodic and rhythmic-melodic perception, as well as melody recognition).

As expected, the groups differed in their linguistic abilities: Children with SLI performed at least 1 SD below their age norm, whereas the control groups performed at the expected age level. For phonemic discrimination, the SLI group had a speech perception level comparable to the CL group, whereas both groups performed below the CA group. Although children with a nonverbal IQ below 80 were excluded from the analysis, the score of children with SLI was still below that of controls.

Notably, children with SLI and those from the control groups differed significantly in their music perception skills. For the previously unknown melodies (comparable condition) and the sum score of all conditions of the melodic perception task, as well as for all subtests except the transposed condition of the rhythmic-melodic perception tasks, the performance level of children with SLI was significantly below that of the age-matched controls and rather similar to that of children with comparable linguistic abilities (CL). Processing of pitch appeared to be easier than that of rhythm: Whereas the differences between the SLI and the CL group were at least approaching significance ($p \leq 0.067$) for all rhythmic-melodic perception tasks, only for one melodic perception task (comparable) and the sum score a significant group difference was observed.

In the melody recognition tasks, children with SLI performed nominally below both control groups in all but the timbre-change subtest of the melody perception task. However, a significant difference between children with SLI and the CA group was obtained only for the standard condition. While in the melodic and rhythmic-melodic perception tasks (reported above) the performance of children with SLI was similar to the CL group, it is lower (at least nominally) than in either control group for most melody recognition tasks.

Compared to previous evidence, the observation that children with SLI are poorer in detecting violations in melody (pitch) or rhythm (duration) than age-matched controls is similar to results from previous studies [33, 35, 38]. Our data provide a slightly different focus compared to previous studies: Corriveau et al. [33] and Cumming et al. [35] had a stronger focus on rhythm perception and the processing of basic auditory properties (such as amplitude rise time and duration). Clément et al. [38] were primarily interested in the music production skills although they also provided data on music perception skills. Taken together, the data from those studies indicate that the difficulties to process certain acoustic features underlying prosody (amplitude rise time, tempo, stress, and duration) also influence music perception. Importantly, compared to previous studies our data explore a much younger age cohort (four- and five-year-olds as compared to ten- and eleven-year-olds), complementing the knowledge about relations between music perception and linguistic skills at earlier stages of language acquisition.

Significant correlations between the language comprehension subtest and music perception were found when all subgroups are pooled in all conditions of the melody recognition task and for transposed and comparable condition as well for the sum score of the rhythmic-melodic perception tasks but not for the standard condition. Likewise, the sum scores of the music perception tasks were correlated

with almost all subtests of the language screening. When exploring the correlations within the group of children with SLI, significant correlations were primarily observed for the more complex conditions where different musical parameters were varied (i.e., the transposed conditions in the rhythmic-melodic perception task, as well as the faster and the timbre-change condition in the melody recognition task). Given that similar correlations can not be observed within either control group and assuming common underlying processing skills during music perception and for prosodic cues (cf., e.g., [22, 23]), these results provide further evidence for the prosodic phrasing hypothesis which suggests that auditory impairments contribute via perceptual difficulties with global prosodic structure to the grammatical difficulties observed in children with SLI. It appears that those skills are more predictive of linguistic proficiency during early language development and in children who are delayed because of impaired language development.

In addition, the correlations of the rhythmic-melodic perception and the melody recognition tasks with those subtests of the intelligence test which assess short-term/working memory suggest that a common underlying general cognitive skill, such as working memory capacity, may account for the correlations between the different music and language perception tasks. However, the difference can not be accounted for by differences in memory load among the stimuli as those had similar acoustic characteristics and length. Baddeley [46] proposed that the capacity of working memory can be increased by automatizing underlying subprocesses. Lack of automaticity in processing musical elements may limit the storage capacity of working memory and hence impede binding the elements of musical phrases together. Assuming that children with SLI have not yet established automatic, preattentive analyses for certain music-specific parameters (melody, rhythm, timbre, etc.) might account for their poor performance, especially when processing complex musical sequences. This assumption has to be further tested in future studies.

One limitation of the current study is that the performance of most children was relatively close to chance level for the melodic and the rhythmic-melodic perception tasks. Particularly the younger controls did not exceed chance level for most tasks, presumably because of the substantial variance in their performance and the rather small sample size. The development of tasks that can be accomplished by younger children would thus be desirable.

Our results suggest that exploring music perception skills can inform theories about typical and impaired language acquisition. For this reason, the tasks used in the present study should be further developed: Musical material can be a useful indicator of language processing difficulties, because it allows exploring skills that are prerequisites of successful language perception. Parameters or features like pitch, timbre, tempo, and their complexity can be easily manipulated using musical material. A fine-grained assessment of musical skills and a detailed description of their relation to linguistic abilities can provide important information about aetiology of SLI and may open new perspectives for diagnosis and therapy of those children.

5. Conclusion

The current study demonstrated difficulties of children with SLI to perceive changes in the pitch and rhythm of musical phrases. It complements and extends evidence from previous studies that observed difficulties with processing musical syntax [36], perceiving rhythm [33, 35], and producing music [37, 38]. This suggests that children with SLI also show difficulties with aspects of nonlinguistic processing which are potentially a crucial phenomenon of SLI. The observed relations between language and music perception strengthen assumptions about the importance of musical parameters during language acquisition [9, 10].

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

The authors would like to thank their participants and their families who made this research possible and the kindergartens for their help in recruiting children and providing rooms for testing them. The Konrad Adenauer Foundation supported the dissertation of Stephan Sallat.

References

- [1] J. B. Tomblin, N. L. Records, P. Buckwalter, X. Zhang, E. Smith, and M. O'Brien, "Prevalence of specific language impairment in kindergarten children," *Journal of Speech, Language, and Hearing Research*, vol. 40, no. 6, pp. 1245–1260, 1997.
- [2] N. Botting, "Non-verbal cognitive development and language impairment," *Journal of Child Psychology and Psychiatry and Allied Disciplines*, vol. 46, no. 3, pp. 317–326, 2005.
- [3] L. B. Leonard, *Children with Specific Language Impairment*, MIT Press, Cambridge, Mass, USA, 2014.
- [4] L. B. Leonard, "Specific language impairment across languages," in *Speech and Language Impairments in Children: Causes, Characteristics, Intervention and Outcome*, D. V. M. Bishop and L. B. Leonard, Eds., pp. 115–130, Psychology Press, Hove, UK, 2000.
- [5] M. J. Snowling, D. V. M. Bishop, and S. E. Stothard, "Is preschool language impairment a risk factor for dyslexia in adolescence?" *The Journal of Child Psychology and Psychiatry*, vol. 41, no. 5, pp. 587–600, 2000.
- [6] M. Ors, E. Ryding, M. Lindgren, P. Gustafsson, G. Blennow, and I. Rosén, "SPECT findings in children with specific language impairment," *Cortex*, vol. 41, no. 3, pp. 316–326, 2005.
- [7] J. B. Tomblin, L. L. Hafeman, and M. O'Brien, "Autism and autism risk in siblings of children with specific language impairment," *International Journal of Language & Communication Disorders*, vol. 38, no. 3, pp. 235–250, 2003.
- [8] D. F. Newbury, D. V. M. Bishop, and A. P. Monaco, "Genetic influences on language impairment and phonological short-term memory," *Trends in Cognitive Sciences*, vol. 9, no. 11, pp. 528–534, 2005.
- [9] S. Koelsch, "Toward a neural basis of music perception—a review and updated model," *Frontiers in Psychology*, vol. 2, article 110, 2011.
- [10] A. Brandt, M. Gebrian, and L. R. Slevc, "Music and early language acquisition," *Frontiers in Psychology*, vol. 3, article 327, 2012.
- [11] S. Jentschke, "The relationship between music and language," in *The Oxford Handbook of Music Psychology*, S. Hallam, I. Cross, and M. Thaut, Eds., Oxford University Press, Oxford, UK, 2nd edition, 2015.
- [12] A. D. Patel, *Music, Language, and the Brain*, Oxford University Press, New York, NY, USA, 2008.
- [13] E. McMullen and J. R. Saffran, "Music and language: a developmental comparison," *Music Perception*, vol. 21, no. 3, pp. 289–311, 2004.
- [14] W. T. Fitch, "The biology and evolution of music: a comparative perspective," *Cognition*, vol. 100, no. 1, pp. 173–215, 2006.
- [15] M. D. Hauser and J. McDermott, "The evolution of the music faculty: a comparative perspective," *Nature Neuroscience*, vol. 6, no. 7, pp. 663–668, 2003.
- [16] N. Steinbeis and S. Koelsch, "Shared neural resources between music and language indicate semantic processing of musical tension-resolution patterns," *Cerebral Cortex*, vol. 18, no. 5, pp. 1169–1178, 2008.
- [17] S. Jentschke and S. Koelsch, "Musical training modulates the development of syntax processing in children," *NeuroImage*, vol. 47, no. 2, pp. 735–744, 2009.
- [18] S. Koelsch, T. C. Gunter, M. Wittfoth, and D. Sammler, "Interaction between syntax processing in language and in music: an ERP Study," *Journal of Cognitive Neuroscience*, vol. 17, no. 10, pp. 1565–1577, 2005.
- [19] C. L. Magne, D. Schön, and M. Besson, "Musician children detect pitch violations in both music and language better than nonmusician children: behavioral and electrophysiological approaches," *Journal of Cognitive Neuroscience*, vol. 18, no. 2, pp. 199–211, 2006.
- [20] C. Neuhaus, T. R. Knösche, and A. D. Friederici, "Effects of musical expertise and boundary markers on phrase perception in music," *Journal of Cognitive Neuroscience*, vol. 18, no. 3, pp. 472–493, 2006.
- [21] A. Fox, *Prosodic Features and Prosodic Structure: The Phonology of Suprasegmentals*, Oxford University Press, New York, NY, USA, 2000.
- [22] A. D. Friederici, "Neurophysiological markers of early language acquisition: from syllables to sentences," *Trends in Cognitive Sciences*, vol. 9, no. 10, pp. 481–488, 2005.
- [23] P. K. Kuhl, "Early language acquisition: cracking the speech code," *Nature Reviews Neuroscience*, vol. 5, no. 11, pp. 831–843, 2004.
- [24] D. V. M. Bishop, C. V. Adams, and C. F. Norbury, "Distinct genetic influences on grammar and phonological short-term memory deficits: evidence from 6-year-old twins," *Genes, Brain and Behavior*, vol. 5, no. 2, pp. 158–169, 2006.
- [25] S. Weinert, "Deficits in acquiring language structure: the importance of using prosodic cues," *Applied Cognitive Psychology*, vol. 6, no. 6, pp. 545–571, 1992.
- [26] A. A. Benasich and P. Tallal, "Infant discrimination of rapid auditory cues predicts later language impairment," *Behavioural Brain Research*, vol. 136, no. 1, pp. 31–49, 2002.
- [27] P. Tallal, "Children with language impairment can be accurately identified using temporal processing measures: a response to Zhang and Tomblin," *Brain and Language*, vol. 65, pp. 395–403 (1998)," *Brain and Language*, vol. 69, no. 2, pp. 222–229, 1999.

- [28] P. Tallal and M. Piercy, "Defects of non-verbal auditory perception in children with developmental aphasia," *Nature*, vol. 241, no. 5390, pp. 468–469, 1973.
- [29] X. Zhang and J. B. Tomblin, "Can children with language impairment be accurately identified using temporal processing measures? A simulation study," *Brain and Language*, vol. 65, no. 3, pp. 395–403, 1998.
- [30] D. V. M. Bishop, C. V. Adams, K. Nation, and S. Rosen, "Perception of transient nonspeech stimuli is normal in specific language impairment: evidence from glide discrimination," *Applied Psycholinguistics*, vol. 26, no. 2, pp. 175–194, 2005.
- [31] D. V. M. Bishop and G. M. McArthur, "Individual differences in auditory processing in specific language impairment: a follow-up study using event-related potentials and behavioural thresholds," *Cortex*, vol. 41, no. 3, pp. 327–341, 2005.
- [32] G. M. McArthur and D. V. M. Bishop, "Which people with specific language impairment have auditory processing deficits?" *Cognitive Neuropsychology*, vol. 21, no. 1, pp. 79–94, 2004.
- [33] K. Corriveau, E. Pasquini, and U. Goswami, "Basic auditory processing skills and specific language impairment: a new look at an old hypothesis," *Journal of Speech, Language, and Hearing Research*, vol. 50, no. 3, pp. 647–666, 2007.
- [34] L. Przybylski, N. Bedoin, S. Krifi-Papoz et al., "Rhythmic auditory stimulation influences syntactic processing in children with developmental language disorders," *Neuropsychology*, vol. 27, no. 1, pp. 121–131, 2013.
- [35] R. Cumming, A. Wilson, and U. Goswami, "Basic auditory processing and sensitivity to prosodic structure in children with specific language impairments: a new look at a perceptual hypothesis," *Frontiers in Psychology*, vol. 6, article 972, 2015.
- [36] S. Jentschke, S. Koelsch, S. Sallat, and A. D. Friederici, "Children with specific language impairment also show impairment of music-syntactic processing," *Journal of Cognitive Neuroscience*, vol. 20, no. 11, pp. 1940–1951, 2008.
- [37] K. H. Corriveau and U. Goswami, "Rhythmic motor entrainment in children with speech and language impairments: tapping to the beat," *Cortex*, vol. 45, no. 1, pp. 119–130, 2009.
- [38] S. Clément, C. Planchou, R. Béland, J. Motte, and S. Samson, "Singing abilities in children with Specific Language Impairment (SLI)," *Frontiers in Psychology*, vol. 6, article 420, 2015.
- [39] International Labour Organization, *International Standard Classification of Occupations*, ISCO88, International Labour Office, Geneva, Switzerland, 1990.
- [40] H. B. G. Ganzeboom and D. J. Treiman, "Internationally comparable measures of occupational status for the 1988 International Standard Classification of Occupations," *Social Science Research*, vol. 25, no. 3, pp. 201–239, 1996.
- [41] H. Grimm, *SETK 3-5: Sprachentwicklungstest für Drei- bis Fünfjährige Kinder*, Hogrefe, Göttingen, Germany, 2001, (German).
- [42] A. S. Kaufman, N. L. Kaufman, and P. Melchers, *K-ABC: Kaufman Assessment Battery for Children*, Swets & Zeitlinger, Amsterdam, The Netherlands, 1994 (German).
- [43] E. E. Gordon, *Advanced Measures of Music Audiation*, GIA Publications, Chicago, Ill, USA, 1989.
- [44] C. E. Seashore, D. Lewis, and J. G. Saetveit, *Seashore Measures of Musical Talents. Manual*, The Psychological Corp, New York, NY, USA, 1960.
- [45] I. Peretz, N. Gosselin, Y. Nan, E. Caron-Caplette, S. E. Trehub, and R. Beland, "A novel tool for evaluating children's musical abilities across age and culture," *Frontiers in Systems Neuroscience*, vol. 7, article 30, 2013.
- [46] A. D. Baddeley, "Working memory and language: an overview," *Journal of Communication Disorders*, vol. 36, no. 3, pp. 189–208, 2003.

Research Article

The Influence of Music on Prefrontal Cortex during Episodic Encoding and Retrieval of Verbal Information: A Multichannel fNIRS Study

Laura Ferreri, Emmanuel Bigand, Patrick Bard, and Aurélia Bugaiska

UMR-CNRS 5022 “Laboratoire d’Etude de l’Apprentissage et du Développement”, Université de Bourgogne, 21000 Dijon, France

Correspondence should be addressed to Laura Ferreri; lf.ferreri@gmail.com

Received 23 January 2015; Revised 2 May 2015; Accepted 11 May 2015

Academic Editor: Kentaro Ono

Copyright © 2015 Laura Ferreri et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Music can be thought of as a complex stimulus able to enrich the encoding of an event thus boosting its subsequent retrieval. However, several findings suggest that music can also interfere with memory performance. A better understanding of the behavioral and neural processes involved can substantially improve knowledge and shed new light on the most efficient music-based interventions. Based on fNIRS studies on music, episodic encoding, and the dorsolateral prefrontal cortex (PFC), this work aims to extend previous findings by monitoring the entire lateral PFC during both encoding and retrieval of verbal material. Nineteen participants were asked to encode lists of words presented with either background music or silence and subsequently tested during a free recall task. Meanwhile, their PFC was monitored using a 48-channel fNIRS system. Behavioral results showed greater chunking of words under the music condition, suggesting the employment of associative strategies for items encoded with music. fNIRS results showed that music provided a less demanding way of modulating both episodic encoding and retrieval, with a general prefrontal decreased activity under the music versus silence condition. This suggests that music-related memory processes rely on specific neural mechanisms and that music can positively influence both episodic encoding and retrieval of verbal information.

1. Introduction

Episodic memory can be defined as a neurocognitive system, uniquely different from other memory systems, which enables human beings to remember past experiences [1]. Numerous studies have investigated the factors that can boost this system. According to the encoding specificity principle [2], the memory trace of an event and hence the properties of effective retrieval cues are determined by the specific encoding operations performed by the system on the input stimuli. Craik and Lockhart [3] first proposed that the durability of the trace depends on the “depth” of encoding processing, deeper semantic processing allowing better encoding of the target information. Furthermore, it has been demonstrated that the encoding context of an event plays a crucial role in successful memory performance. For instance, a rich context given by stimuli with a high (positive or negative) emotional valence can enhance the encoding of contextual information associated with an item [4]. In this scenario, music could offer a perfect example of an

enriched context. Indeed, given its complexity as a stimulus that evolves through time and has a strong emotional impact [5], music is likely to enrich the encoding context of an event, thereby improving subsequent memory performance. The evocative power of music is fascinating and undisputed: it can evoke both emotional states and personal events from the past [6]. Several studies have revealed a specific episodic memory for music, showing how it depends largely on emotion [7] and revealing the existence of specific related neural processes [8]. Nevertheless, the question of whether music as an encoding context can enhance episodic memory performance, especially concerning verbal material, remains debatable and controversial. Several studies have shown that music, presented either as background or as sung text, can enhance verbal learning and memory in both healthy and clinical populations [9–13]. However, several authors have recently claimed that music can also draw attention away from the to-be-remembered information, thus interfering in memory performance [14–16]. The key to solving this question seems to rely on a better understanding of the processes

involved: improving our knowledge of how music can boost memory performance at both behavioral and functional (i.e., neuronal) levels could shed new and essential light on the most efficient music-based paradigms and interventions.

In a series of functional near-infrared spectroscopy (fNIRS) studies, we previously showed that background music during the episodic encoding of verbal material can improve item and source memory performance and modulate prefrontal cortex (PFC) activity [10, 11]. More specifically, fNIRS studies have found that music leads to decreased activation (i.e., decrease in oxyhemoglobin- O_2Hb and deoxyhemoglobin- HHb increase) in the dorsolateral prefrontal cortex (DLPFC), known to be important for organizational, associative, and memory encoding [17]. In view of fNIRS studies showing decreased PFC activity during verbal learning in which subjects were helped during their performance [18, 19], we hypothesized that music could modulate episodic encoding by modifying the need of extra organizational and strategic encoding usually attributed to the DLPFC [20] and facilitating the creation of richer associative bindings crucial for subsequent retrieval [10, 11]. However, both methodological and theoretical caveats raise important issues. The present work therefore aims to increase our knowledge of music-related memory processes by extending investigations of background music and verbal memory through three main questions arising from these previous studies.

First, existing fNIRS data are limited to the encoding phase, raising the question of which mechanisms are involved during episodic retrieval. Research on episodic memory has clearly demonstrated that in order to understand how memories are formed, we need first to understand many cognitive and neurobiological processes involved in both encoding and retrieval, as well as the interactions among these phases [21]. Furthermore, in the light of the contrasting results in the literature, it is crucial to know whether the music facilitation reflected in decreased PFC activation is also found in the retrieval phase or whether by contrast it shows a more demanding pattern in line with the interference hypothesis. Therefore, in the present study, fNIRS prefrontal data were acquired during both encoding and retrieval of words in order to test the hypothesis that the PFC disengagement found during memory formation is also found during the retrieval phase.

Secondly, previous fNIRS acquisitions were limited to eight channels covering the bilateral DLPFC, thus hindering the possibility of ascertaining the music effect throughout the lateral prefrontal cortex, which is crucial during episodic memory processes [22–24]. Ventrolateral and dorsolateral regions of the PFC have been shown to implement different controls that provide complementary support for long-term memory encoding and retrieval. More specifically, during the encoding phase, ventrolateral prefrontal cortex (VLPFC) regions contribute to the ability to select goal-relevant item information and strengthen the representation of goal-relevant features, while DLPFC regions contribute to memory enhancement by allowing associations among items in long-term memory during encoding [17]. Concerning the retrieval phase, several studies on episodic memory retrieval have

found a fronto-parieto-cerebellar network, in which several bilateral frontal regions seem to mediate processes that act in the output of episodic retrieval operations (see [22] for a review). It is therefore important to understand whether the observed PFC deactivation is restricted to the dorsolateral region or whether it includes the whole lateral prefrontal cortex. While a delimited prefrontal deactivation would suggest that music specifically modulates certain cognitive processes, a decrease throughout the PFC during the music condition would indicate an overall PFC disengagement and suggest that music-related memory processes rely on music-specific and unusual neural mechanisms. Hence, in the present study, a multichannel (i.e., 48 measurement points) fNIRS system was used to monitor the whole PFC cortex during episodic encoding and retrieval.

The third important point concerns a behavioral issue. Our previous behavioral and functional results led us to explain the findings in terms of associative bindings. A musical context may afford efficient mnemonic strategies allowing the creation of interitem and item-source associations that can help subsequent retrieval. These mnemonic strategies would result in PFC deactivation [18, 19]. If confirmed, this would be an important contribution to the existing debate about the complex music-memory issue. However, previous studies used judgment memory tasks, whereby subjects were presented with a copy of old items and had to retrieve and judge whether or not each item had been presented previously (item memory) and, if so, in which context (source memory). Using this paradigm, it was not possible to investigate possible associative processes. Therefore, in the present study we used a free recall task in order to investigate if subjects adopted particular strategies during the retrieval phase.

To extend our knowledge of music-related memory processes and contribute to the current debate, the present study used multichannel fNIRS to test lateral prefrontal activations during music-related encoding and retrieval (i.e., free recall). We asked subjects to memorize lists of words, presented with a background of either music or silence, and to retrieve as many words as possible after an interference phase. We used a 48-channel fNIRS system to monitor their PFC activity bilaterally. Based on the hypothesis that a background of music would modulate PFC activity throughout the memory processing stages, we expected to find less cortical activation during both the encoding and the retrieval phases under the music condition, in line with our previous studies on verbal encoding with music [10, 11]. Furthermore, clustered retrieval of previously encoded words for the music condition when compared to the silence condition would suggest that music helps encoding through the implementation of associative strategies.

2. Materials and Methods

2.1. Participants. Nineteen young healthy students at the University of Burgundy (13 female, mean age 21.65 ± 3.2 years) took part in the experiment in exchange for course credits. All the participants were right-handed, nonmusicians, and native

French-speakers and reported having normal or corrected-to-normal vision and hearing. None were taking medication known to affect the central nervous system. Informed written consent was obtained from all participants prior to taking part in the experiment. The study was anonymous and fully complied with the Helsinki Declaration, Convention of the Council of Europe on Human Rights and Biomedicine.

2.2. Experimental Procedure. Subjects were seated in front of a computer in a quiet, dimly lit room. After the 48 fNIRS probe-set had been fitted on the forehead overlying the PFC (see fNIRS section below for detailed description), the in-ear headphones inserted and the sound recorder placed, subjects were informed that they would be presented with different lists of words with two auditory backgrounds: music or silence. They were asked to memorize the lists of words and were told that, after a brief backward counting task, they should mentally recall the previously seen words and then say as many as they could.

Verbal stimuli consisted of 90 taxonomically unrelated concrete nouns selected from the French “Lexique” database ([25]; <http://www.lexique.org/>). Words were randomly divided into six lists (15 words per list, 45 words for each condition), matched for word length and occurrence frequency. In the music condition, the background music used in all blocks was the instrumental jazz piece “Crab walk” (by Everything But The Girl, 1994). This excerpt was chosen after a pretest among a list of 8 pieces representing different musical genres (such as classical, jazz, new age) preselected by the authors. All the excerpts were instrumental in order to avoid possible interference between the lyrics and the verbal material to be encoded. The excerpts were evaluated by nonmusician participants in terms of arousal, emotional valence, and pleasantness using a 10-point scale. Participants were also asked to report if the music was familiar or not. The selected piece was chosen for its positive valence, medium arousal quality and for being rated as pleasant and unfamiliar.

Three blocks for each condition (music or silence) were presented to each subject, giving a total of 6 experimental blocks. Each block consisted of three phases, namely, encoding, interference, and retrieval. In the encoding phase, 15 words were displayed successively against a background of music or silence. The auditory stimulation started 15 s before the first word was displayed, continued during the sequential display of words, and ended 15 s after the last word. Words in each block were displayed for 3 s per word, amounting to 45 s for the sequential presentation of 15 words. Each encoding phase therefore lasted 75 s (15 s background, 45 s words, and 15 s background). Verbal stimuli were visually presented in white on black background in the middle of the screen. Auditory stimuli were presented using in-ear headphones, and the overall loudness of the excerpts was adjusted subjectively to ensure a constant loudness throughout the experiment.

Prior to the retrieval phase, subjects were asked to count down from a given number displayed on the screen until the word “stop” appeared. The interference phase lasted 30 seconds.

The retrieval phase was divided into 15 s of a “search for” phase, in which the previous encoding background

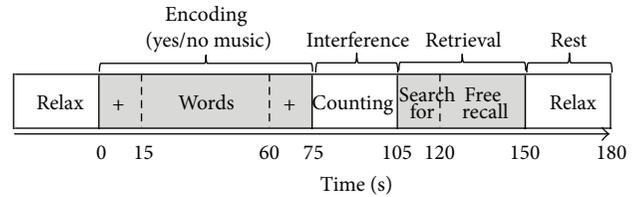
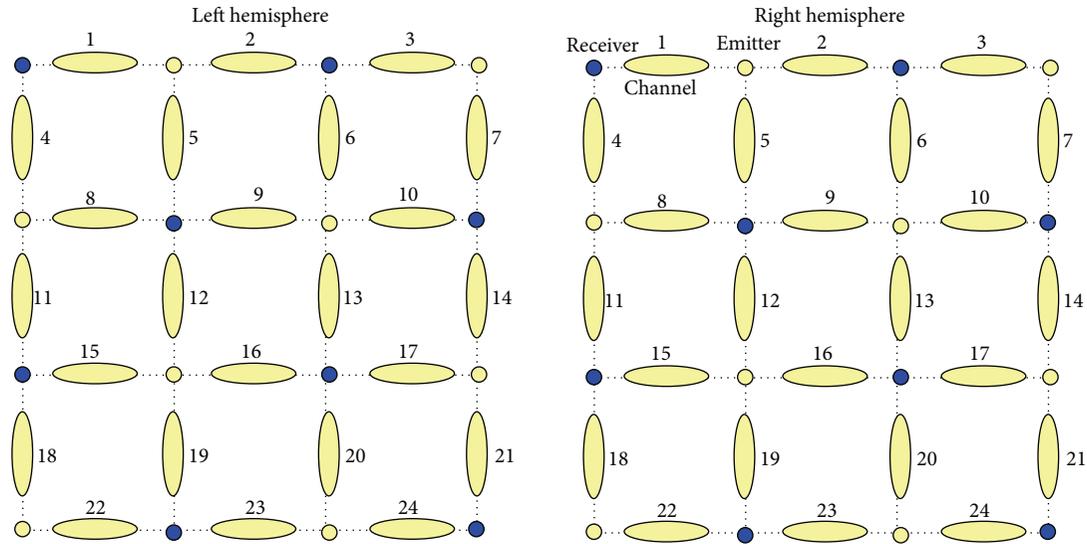


FIGURE 1: Representation of one encoding-interference-retrieval block between two 30 s rest blocks. Each block consisted of 15 s of context (+ on the screen) alone (music or silence in the earphones), then 45 s of context and word encoding (with either background music or silence), and then again 15 s of context (+) alone. After 30 seconds of interference phase (counting), subjects were asked to search for previously encoded words (search for, with either background music or silence, 15 s) and then to recall as many words as they could (free recall, 30 s).

was presented in the earphones and subjects were asked to mentally recall the previously seen words and 30 s of a “free recall” phase, in which subjects were asked to say aloud as many words of the previously encoded list as possible. There were two reasons for this procedure: first, to resituate the subjects in the source of encoding, enabling good memory performance (see, e.g., [26, 27]); secondly, to have a control condition for possible voice-movement artifacts. A sound recorder was used to record subjects’ free recall performance. The retrieval phase lasted 45 s. The total duration of each block was 3 minutes. Each block was followed by a 30 s rest (silent) (Figure 1).

The order of music/silence blocks was counterbalanced, as well as the order of word lists and the order of words in the lists. During the rest periods, subjects were instructed to try to relax and not to think about the task; in contrast, during the context-only phases of the blocks, participants were instructed to concentrate on a fixation cross on the screen and to focus on the task. Presentations of task instructions and stimuli were controlled by E-Prime software (Psychology Software Tools, Inc.) using a laptop with a 15” monitor. The entire experimental session, including fNIRS recording, lasted about 20 minutes.

2.3. fNIRS Measurements. A 48-channel fNIRS system (OxymonMkIII, Artinis Medical Systems B.V., The Netherlands) was used to measure the concentration changes of O₂Hb and HHb (expressed in μM) using an age-dependent constant differential path-length factor given by $4.99 + 0.0067 * (\text{age}^{0.814})$ [28]. Data were acquired at a sampling frequency of 10 Hz. The 48 fNIRS optodes (24 emitters and 24 detectors, Figure 2(a)) were placed symmetrically over the lateral PFC. The distance between each emitter and detector was fixed at 3 cm. For each hemisphere, fNIRS channels measured the hemoglobin concentration changes at 24 measurement points in a 12 cm² area, with the lowest optodes positioned along the Fp1-Fp2 line and the most central optodes 2 cm from the Cz line [29], in accordance with the international 10/20 system [30]. From top to bottom, these measurement points were labeled 1–24 (see Figure 2(a)).



(a)



(b)

FIGURE 2: Channels template (a) and localization (b). The 48-channel NIRS system consisted in 8 emitters (yellow circles) and 8 receivers (blue circles) for each hemisphere, resulting in 24-left and 24-right measurement points (yellow lengthened shapes).

To optimize signal-to-noise ratio during the fNIRS recording, the 48 optodes were masked from ambient light by a black plastic cap that was kept in contact with the scalp with elastic straps, and all cables were suspended from the ceiling to minimize movement artifacts [31] (Figure 2(b)). During data collection, O_2Hb and HHb concentration changes were displayed in real time, and the signal quality and the absence of movement artifacts were verified.

2.4. Data Analysis

2.4.1. Behavioral Data. Memory performance was calculated for each subject under both conditions by computing the total number of correctly retrieved words. Incorrectly retrieved items were considered as intrusions. Paired t -tests were used to compare the free recall memory and intrusion

scores in the silence and music conditions. Subjects' possible associative strategies at encoding were examined using cluster analysis, in which the chunks created at retrieval indicated the level of interitem associations at the encoding phase. We therefore calculated the number of items presented in a row (i.e., one following the other) during the encoding phase that were retrieved in chunks, identifying 2-, 3-, 4-, 5-, and 6-word chunks produced by each subject and under each condition (e.g., if the subject encoded “bottle,” “fork,” “match,” “coat,” and “pool” in the encoding phase and then subsequently retrieved “fork,” “match,” and “coat” during the free recall task, this constituted a 3-word chunk; if the subject retrieved “bottle,” “match,” and “coat,” this constituted a 2-word chunk). Paired t -tests successively compared the most consistent chunks (>2-words) over the total of chunk ratios between the silence and music conditions.

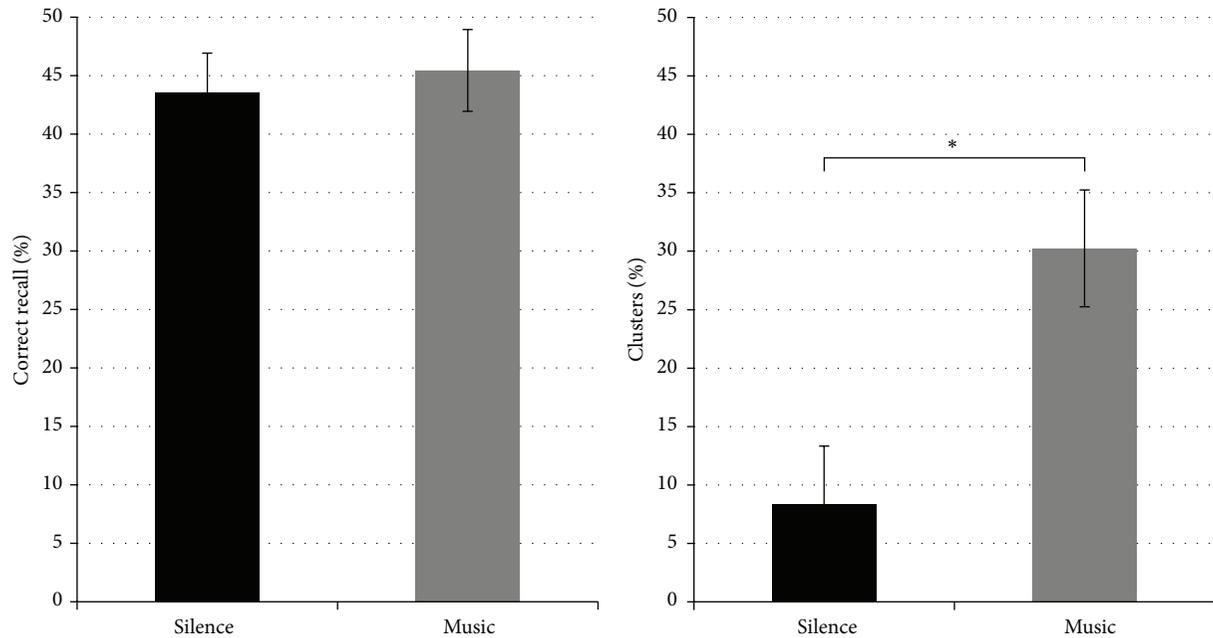


FIGURE 3: Behavioral results showing the total percentage of correctly retrieved words (left side) and of clusters (>2-word chunks) created in the free recall phase for the silence (black bars) and music (grey bars) conditions. * shows statistically significant differences ($P < .05$).

2.4.2. fNIRS Data. In order to eliminate task-irrelevant systemic physiological oscillations, the O_2Hb and HHb signals were first low-pass filtered (5th-order digital Butterworth filter with cut-off frequency 0.1 Hz) for each of the 48 fNIRS measurement points.

To determine the amount of activation during the encoding phase for the two conditions, data in each of the 6 experimental blocks was baseline corrected using the mean of the O_2Hb and HHb signals during the last 5 s of the rest phase. We then sample-to-sample averaged (i.e., 10 samples/s) the baseline-corrected signals over the 3 blocks of each condition, yielding one average music and silence O_2Hb and HHb signal per participant for both the encoding phase and the retrieval phase (both “search for” and “free recall” tasks). We then computed the maximum O_2Hb and the minimum HHb delta-from-baseline values over the 45 s (for the encoding), 15 s (for the “search for” retrieval), and 30 s (for the “free recall” retrieval) stimulus windows, for both the music and silence average block of each participant and for each channel (see Figure 4). Delta values were then statistically analyzed using a repeated measure ANOVA with 2 (music/silence condition) \times 2 (left/right hemisphere) \times 24 (optodes) repeated factors. Paired t -tests were also used to compare each channel as well as the means of left right activity for the silence and music condition and for each phase of the memory task [31] (see Figure 4).

Furthermore, in order to ascertain the PFC activation during the entire block of music/silence encoding and retrieval conditions, we ran a complete group time-series analysis in which we averaged O_2Hb , HHb , and total Hb (THb) concentrations over 5 s windows (i.e., one average

point for each 5 s) all over the blocks of encoding, interference, “search for,” and free recall phase, getting 35 successive measures of concentrations. Time-series data were then analyzed using a repeated measure ANOVA with 2 (music/silence condition) \times 2 (left/right hemisphere) \times 24 (optodes) \times 35 (points in time) within-subject factors.

3. Results

3.1. Behavioral Results. Paired t -tests on the free recall memory performance and intrusion scores revealed no differences in the total number of correctly retrieved words and false-alarm rates between the music and silence conditions ($t(18) = 1.17$, $P > .05$). However, cluster analysis revealed a significant difference between the two conditions concerning the number of chunks created at retrieval. While t -tests on the total number of words retrieved in chunks did not reveal a significant difference between the two conditions ($t(18) = -.165$, $P > .05$), a significant difference was found for cluster creation, data revealing that subjects created more consistent chunks (>2 words) in the music than in the silence condition, ($t(18) = 2.23$, $P = .02$) (Figure 3).

3.2. fNIRS Results. Figure 4 shows a channel-level analysis on O_2Hb delta-to-baseline values for each phase of the memory task (encoding, “search for,” and free recall). The repeated-measures ANOVA on O_2Hb delta-to-baseline values during the encoding phase showed a main effect of condition, with the whole PFC significantly less activated during encoding with music than with silence, $F(1, 18) = 9.78$, $P = .006$. For the retrieval phase, statistical analysis showed similar

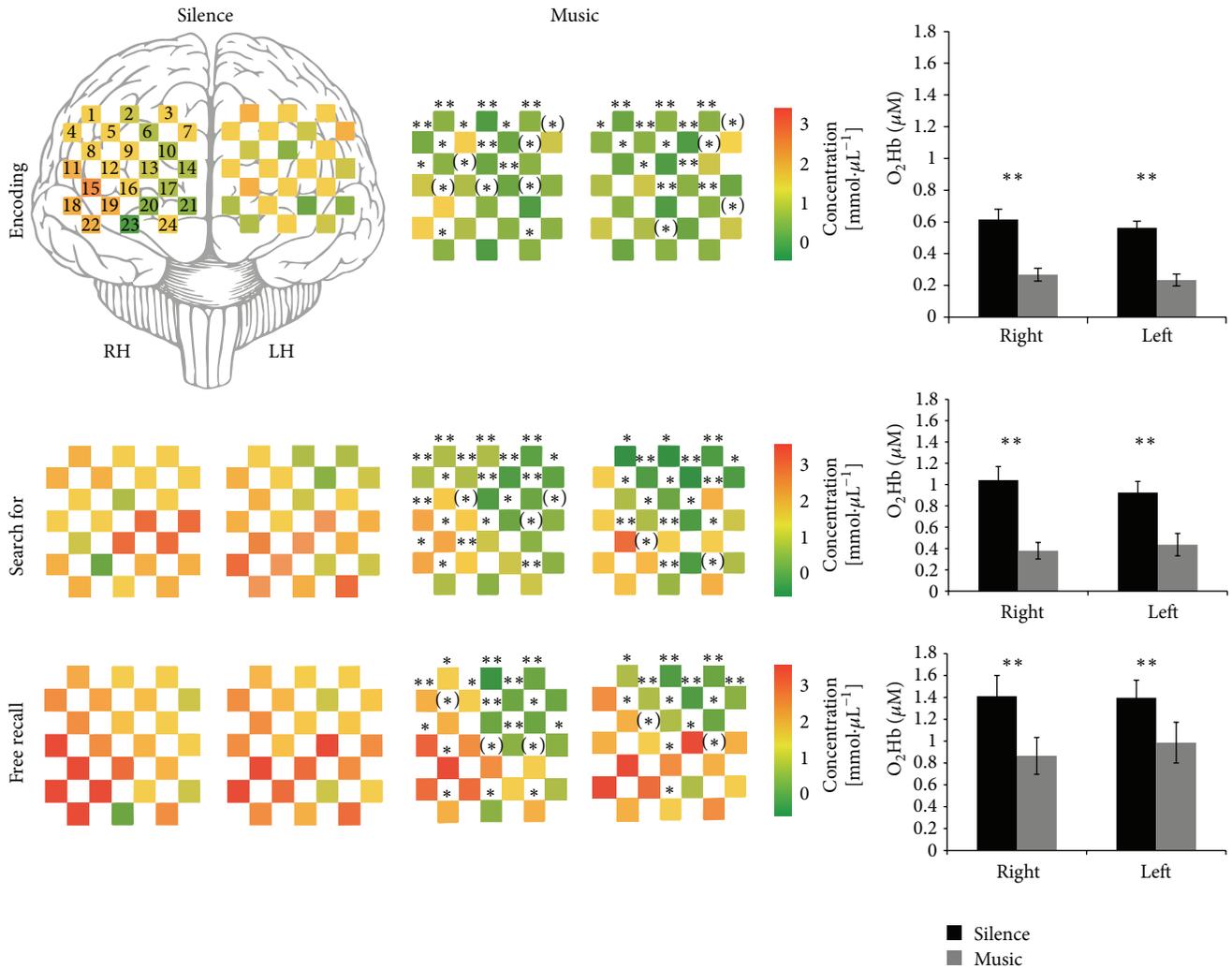


FIGURE 4: O₂Hb Δ-to baseline values expressed in micromoles for silence (left side) and music (right side) conditions, for all 48 channels (24 right -RH-, 24 left -LH-), during the encoding, “search for” and free recall phases. Red = more activated; green = less activated. The whole prefrontal cortex resulted significantly less activated in the music condition during the three phases. **, *, and (*) on the music channels show statistically significant differences (resp., $P < .01$, $P < .05$, and $.05 < P < .09$) between the two conditions for each channel. The difference between the two conditions is also showed in the left right of the figure, with black and grey bars representing O₂Hb Δ-to baseline mean values, respectively, for silence and music conditions, in the right and left hemisphere (** for $P < .01$ resulted from paired t -tests comparisons).

results for the “search for” and “free recall” tasks. Repeated-measures ANOVA on the “search for” phase revealed a main effect of condition, with higher O₂Hb concentrations for retrieval with silence than with music ($F(1, 18) = 9.62$, $P = .006$), which was also confirmed in the “free recall” phase ($F(1, 18) = 8.75$, $P = .008$). The decreased PFC activation under the music retrieval condition was also supported by higher HHb values (based on the balloon model, see, e.g., [32]) for the music condition ($F(1, 18) = 6.93$, $P = .017$ for the “search for” phase, $F(1, 18) = 3.56$, $P = .075$ for the “free recall” phase). These results were also confirmed by paired t -test comparing the mean values of left and right channel for the two conditions (Figure 4).

Time-series analysis on O₂Hb values confirmed a main effect of the condition ($F(1, 18) = 7.58$, $P = .013$), with less

PFC engagement for the music encoding and retrieval phases. This was supported by HHb time-series analyses ($F(1, 18) = 5.63$, $P = .008$) which showed higher values for the music condition. A condition \times laterality interaction was also found for O₂Hb concentrations ($F(1, 18) = 4.48$, $P = .048$), suggesting higher left and right hemisphere engagement, respectively, for silence and music condition. A main effect of time was also found for both O₂Hb ($F(612, 34) = 19.04$, $P < .001$) and HHb values ($F(612, 34) = 6.001$, $P < .001$), as shown in Figure 5.

4. Discussion

Extending previous studies of verbal memory encoding and music [10, 11], the present work investigated music-related

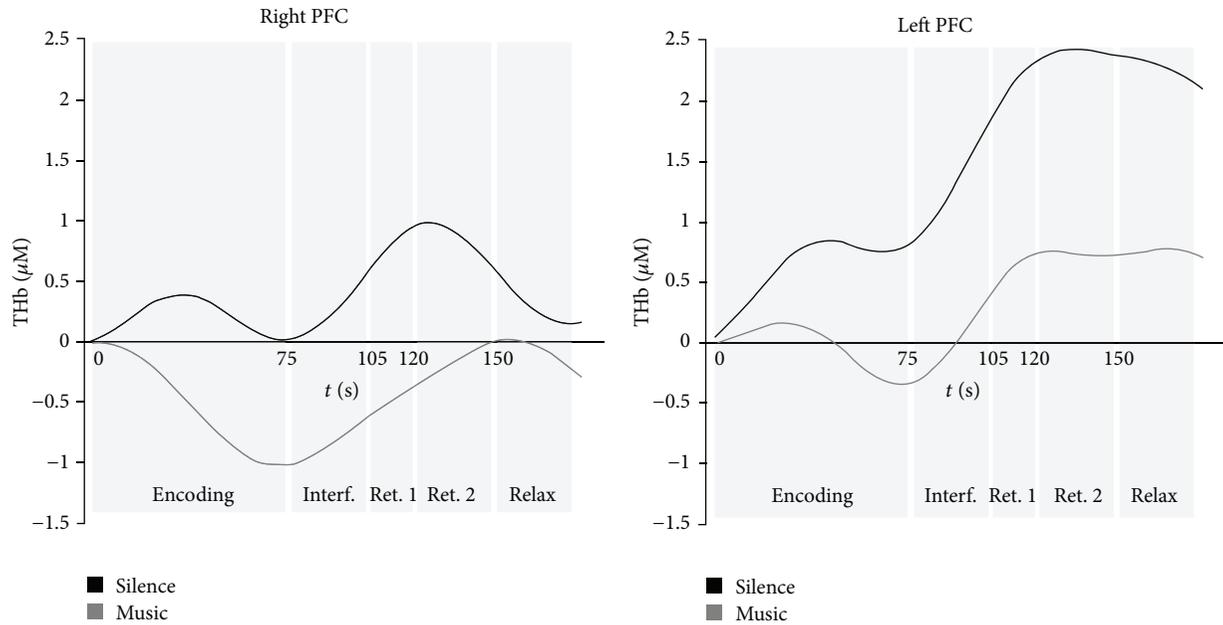


FIGURE 5: Representative single-subject time-series analysis showing THb values expressed in micromoles for silence (black lines) and music (grey lines) conditions, in right (left side) and left (right side) PFC during the encoding, interference, “search for” (ret. 1), free recall (ret. 2), and relax phases.

episodic encoding and retrieval processes using multichannel fNIRS to monitor cortical oxygenation changes over the lateral PFC during both episodic encoding and retrieval of verbal information.

One of the main findings of this study is that activity decreased under the music condition as compared to the silence condition. In line with our previous experiments, fNIRS results during word encoding revealed that the PFC was significantly more active under the silence condition than under the music condition [10, 11]. In the light of fNIRS studies showing PFC cortex deactivation when subjects’ memory performance was improved by given strategies or pharmacological stimulants [18, 19], we previously interpreted the decreased DLPFC activity during music encoding as a music-related facilitation process. More specifically, we postulated that background music, unlike silence, required less involvement of the DLPFC for organizational [17] and relational interitem processing [33] during verbal episodic encoding. The first new finding of the present study is that the decreased activity under the music condition extended across the entire lateral PFC. As shown by Figure 4, analysis on musical encoding and retrieval revealed lower O_2Hb values in almost all channels. As mentioned in the introduction, the DLPFC and VLPFC are jointly recruited to guide the processing of interitem relational information in working memory, which promotes long-term memory for this relational information [20, 34]. In particular, VLPFC is involved in both relational and item-specific memory formation, and it seems to select goal-relevant features during episodic encoding, thus contributing to subjects’ ability to select relevant item information to remember [17, 34, 35]. Although fNIRS limitations in channel localizations make it

hard to specifically identify which lateral prefrontal regions are specifically involved during all over the memory task, these results suggest that the facilitator effect of a musical background also relies on its capacity to disengage the most ventral part of the PFC from its goal-relevant selective functions. In other words, music may affect the encoding state, not only by disengaging the PFC during specific interitem relational strategies (related to DLPFC activity), but also and more generally by affecting episodic prefrontal functions, namely, the capacity to select the relevant information to remember and strategically organize it for successful memory formation.

Another crucial finding of the present study is that such PFC less activation continued during the retrieval phase. Figure 5 shows an example of time-course of the THb signal all over the block of encoding and retrieval: although the retrieval phase showed an increased activation in both conditions, especially in most ventral channels (Figure 4), this was always less pronounced for the music condition. In our opinion, this is important for two main reasons. First, the fact that the music-related PFC decrease was observed during both the “search for” phase (with background music) and the “free recall” phase (without background music) excludes the possibility that the observed PFC modulation was due to the presence of auditory stimulation rather than to a specific music effect. Secondly, music provides a less demanding way of modulating the recruitment of PFC areas crucial for long-term manipulation of information and active strategic retrieval [36–38], indicating a long-lasting effect. This is particularly important in view of the divergent results in the literature. Indeed, if music constitutes a dual-task interference [14, 15], we should have observed highest increase in neural

activity for the music condition in at least one of the memory phases, as previously observed in fNIRS studies investigating dual-task situations [39, 40]. On the contrary, our results suggest that music-related memory processes rely on specific neural mechanisms underlying a less demanding prefrontal engagement throughout the stages of memory formation and retrieval.

In the light of previous fNIRS studies on memory [18, 41, 42], our results should also be viewed in terms of the contribution of fNIRS to understanding the role of PFC in long-term memory processes. Unlike our previous studies, we did not find a main effect of lateralization during word encoding. However, a more thorough time-series analysis revealed a condition \times laterality interaction, suggesting higher left and right hemisphere engagement, respectively, for silence and music condition. Furthermore, a specific lateralization became evident at the retrieval “search for” phase, where we found a left and right lateralization for the silence and music condition, respectively. This condition by laterality interaction related to the presence of music when subjects tried to retrieve previously encoded words can be interpreted in the light of studies showing that the lateralization of PFC activity during retrieval depends on the availability of verbal codes, with left hemispheric involvement for verbally coded information and right hemispheric activation for nonverbally coded information [43].

Major criticism of PFC fNIRS data concerns the task-evoked changes occurring in forehead skin perfusion [44–47]; PFC activity interpretations must therefore be taken with caution. Nevertheless, our findings not only confirm that fNIRS is a good tool for noninvasive investigation of long-term memory [41, 48, 49], but can also help shed new light on music-related prefrontal episodic memory processes. In particular, we suggest that music is able to modulate all stages of memory processing in a state-dependent manner, enabling the creation of relational links that may constitute efficient mnemonic strategies, as well as the successful retrieval of relevant information. Accordingly, less PFC activity is required to put these strategies to use during either encoding or retrieval. Importantly, this explanation is supported by our behavioral results. Indeed, cluster analysis revealed that participants created significantly more chunks (i.e., formed by >2 words) during the free recall of words previously encoded with music [50]. This would indicate that subjects found it easier to adopt relational-associative strategies to create interitem (and possibly item-source) links during memory formation, which were then used as mnemonic strategies for successful retrieval. However, although we previously found that a musical background can boost item [10] and source [11] memory in recognition tasks, this was not the case for the free recall task, where no difference between music and silence was found in the number of correctly retrieved words. This suggests that behavioral paradigms often fail to characterize a reliable effect of music on memory performance, even when imaging methods are able to detect a music-related effect.

Considering many authors claiming that music hampers encoding and leads to negative results, as well as the different positive behavioral outcomes, it remains important to discuss when and how music can help memory performance. In our

opinion, it is crucial to note that many kinds of paradigms using many kinds of music stimuli exist in literature and hence can lead to contrasting results. In the present study, we used a pleasant musical background with a positive emotional valence and medium arousal quality with the specific idea that music can constitute a helpful encoding context. The results can therefore be discussed in the frame of an enriched context (see, e.g., [51, 52]) given by the presence of the music, in which many mechanisms (arousal-mood modulation, emotions, and reward) intervene to orchestrate the final music-related positive effect. In this perspective, the music-dependent prefrontal modulation observed opens new questions about the interpretation of such specific PFC decreased oxygenation pattern and the related facilitation. The mechanisms underlying PFC deactivation are matter of debate and can reflect several neural processes. Some explanations can be found in regard to BOLD signal decrease, which usually corresponds to an O_2Hb decrease and Hb increase in fNIRS signal. A BOLD decrease is usually interpreted as a deactivation that reflects a focal suppression of neural activity [53, 54] and several explanations have been proposed to clarify such deactivation. For instance, Harel et al. [55] claimed that BOLD decrease can be due to stealing of blood from less active regions into the most cerebral blood flow demanding areas. Therefore, the observed fNIRS prefrontal pattern could reflect a higher activation in other brain regions. The present multichannel fNIRS paradigm in part elucidated this question by investigating not only the DLPFC [10, 11], but also the entire PFC activity. Considering the different tasks attributed to the different regions of PFC for the episodic encoding and retrieval [17], it was reasonable to think that music could be more demanding for regions surrounding the DLPFC. Results revealed a prefrontal decrease in almost all the fNIRS channels, suggesting a huge and coherent prefrontal disengagement. However, such disengagement could be related to a greater activation in other (i.e., nonprefrontal) regions [55] that need therefore to be further investigated. Raichle and colleagues [54] proposed that such reduction of neuronal activities might be mediated through a reduction in thalamic inputs to the cortex during attention-demanding cognitive tasks or through the action of diffuse projecting systems like dopamine (see also [56]). fNIRS studies showing deactivation in nonverbal tasks (e.g., video games) have tried to interpret it in terms of attention-demanding tasks [57]. Nevertheless, this hypothesis seems in conflict with other fNIRS studies investigating prefrontal responses to attention tasks. Indeed, several authors have shown how alertness or attention states significantly increase rather than decrease PFC activation [58, 59]. Also the dopamine system can be responsible for PFC deactivation [54]. Dopamine is a neurotransmitter strongly associated with the reward system: it is released in regions such as the ventro-tegmental area (VTA), nucleus accumbens or PFC as a result of rewarding experiences such as sex, food, but also music [60, 61]. Therefore, if prefrontal less activation can be related to the action of diffuse dopamine systems and the positive effect of music may be related to reward-emotional responses as well, it is possible that music-related reward mechanisms play a crucial role in helping subjects

in engaging successful verbal memory processes reflected in PFC disengagement.

Another crucial point to consider concerns the strong relationship between music and language, which has been clearly shown on both behavioral and neurophysiological level (see, e.g., [62]). It is therefore possible that, among possible general mechanisms discussed above, more language-specific processes may directly intervene during the encoding of verbal material with music. More specifically, our findings suggest that semantic-associative mechanisms may be activated more easily in presence of a musical background, thus resulting in greater clustering during the free recall task. A good example is represented by what participants reported in informal posttask metacognitive follow-up: indeed, when asked how difficult they found the task, many of the subjects suggested that music helped them in creating stories (i.e., bindings) among items and between items and music. For example, if the words “pool” and “glass” were subsequently presented and music was present, participants reported these words were easier to remember because of the creation of a little story in their mind (e.g. “*I imagined myself drinking a glass of wine while playing the pool in a bar with a jazzy atmosphere*”). In this case, the musical context may help in creating new connections between the items and the source itself, namely, new episodes that participants can then retrieve during their subjective mental “time travel” [1], as reflected by behavioral findings. Further neurophysiological investigations (e.g., investigating gamma and theta oscillations, shown to bind and temporally order perceptual and contextual representations in cortex and hippocampus [63]) could in this case elucidate possible item-source bindings processes and further research is therefore needed in this domain.

Taken together, our results overall can be seen in the general framework of the music and memory literature, supporting the idea that music can help verbal memory processes and that associative strategies facilitated by the presence of a musical background may explain memory enhancement. Given the increasing need to understand when and through which mechanisms music is able to stimulate cognitive functions, these results offer in our opinion an important contribution to the existent literature and open interesting perspectives on music-based rehabilitation programs for memory deficits.

5. Conclusions

The aim of this study was to focus on the prefrontal processes involved in music-related episodic memory. More specifically, we wanted to extend previous findings of prefrontal deactivation in the encoding phase of verbal material to the whole prefrontal cortex and also to the retrieval phase.

Overall, fNIRS findings show that music may specifically act and modify normal cortical activity; namely, it can entirely modulate the lateral PFC activity during both encoding and retrieval in a less demanding way. In particular, our results suggest that music-related strategic memory processes rely on specific neural mechanisms recruited throughout the stages of memory formation and retrieval. These findings are

supported by behavioral evidence indicating music-related associative strategies in the recall of verbal information and offer interesting perspectives for music-based rehabilitation programs for memory deficits.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This work was supported by the European Project EBRAMUS (European BRAin and MUSic) ITN Grant Agreement no. 218357, the Conseil Régional de Bourgogne, and the MAAMI ANR Project, TecSan Program, ANR-12-TECS-0014.

References

- [1] E. Tulving, “Episodic memory: from mind to brain,” *Annual Review of Psychology*, vol. 53, pp. 1–25, 2002.
- [2] E. Tulving and D. M. Thomson, “Encoding specificity and retrieval processes in episodic memory,” *Psychological Review*, vol. 80, no. 5, pp. 352–373, 1973.
- [3] F. I. M. Craik and R. S. Lockhart, “Levels of processing: a framework for memory research,” *Journal of Verbal Learning and Verbal Behavior*, vol. 11, no. 6, pp. 671–684, 1972.
- [4] S. Hamann, “Cognitive and neural mechanisms of emotional memory,” *Trends in Cognitive Sciences*, vol. 5, no. 9, pp. 394–400, 2001.
- [5] V. N. Salimpoor, I. Van Den Bosch, N. Kovacevic, A. R. McIntosh, A. Dagher, and R. J. Zatorre, “Interactions between the nucleus accumbens and auditory cortices predict music reward value,” *Science*, vol. 340, no. 6129, pp. 216–219, 2013.
- [6] O. Sacks, “The power of music,” *Brain*, vol. 129, no. 10, pp. 2528–2532, 2006.
- [7] S. Eschrich, T. F. Münte, and E. O. Altenmüller, “Unforgettable film music: the role of emotion in episodic long-term memory for music,” *BMC Neuroscience*, vol. 9, article 48, 2008.
- [8] H. Platel, J.-C. Baron, B. Desgranges, F. Bernard, and F. Eustache, “Semantic and episodic memory of music are subserved by distinct neural networks,” *NeuroImage*, vol. 20, no. 1, pp. 244–256, 2003.
- [9] S. Bottiroli, A. Rosi, R. Russo, T. Vecchi, and E. Cavallini, “The cognitive effects of listening to background music on older adults: processing speed improves with upbeat music, while memory seems to benefit from both upbeat and downbeat music,” *Frontiers in Aging Neuroscience*, vol. 6, 2014.
- [10] L. Ferreri, J.-J. Aucouturier, M. Muthalib, E. Bigand, and A. Bugaïska, “Music improves verbal memory encoding while decreasing prefrontal cortex activity: an fNIRS study,” *Frontiers in Human Neuroscience*, vol. 7, article 779, 2013.
- [11] L. Ferreri, E. Bigand, S. Perrey, M. Muthalib, P. Bard, and A. Bugaïska, “Less effort, better results: how does music act on prefrontal cortex in older adults during verbal encoding? An fNIRS study,” *Frontiers in Human Neuroscience*, vol. 8, article 301, 2014.
- [12] N. R. Simmons-Stern, A. E. Budson, and B. A. Ally, “Music as a memory enhancer in patients with Alzheimer’s disease,” *Neuropsychologia*, vol. 48, no. 10, pp. 3164–3167, 2010.

- [13] W. T. Wallace, "Memory for music: effect of melody on recall of text," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, vol. 20, no. 6, pp. 1471–1485, 1994.
- [14] L. Jäncke, E. Brügger, M. Brummer, S. Scherrer, and N. Alahmadi, "Verbal learning in the context of background music: no influence of vocals and instrumentals on verbal learning," *Behavioral and Brain Functions*, vol. 10, no. 1, article 10, 2014.
- [15] L. Jäncke and P. Sandmann, "Music listening while you learn: no influence of background music on verbal learning," *Behavioral and Brain Functions*, vol. 6, article 3, pp. 1–14, 2010.
- [16] A. Racette and I. Peretz, "Learning lyrics: to sing or not to sing?" *Memory & Cognition*, vol. 35, no. 2, pp. 242–253, 2007.
- [17] R. S. Blumenfeld and C. Ranganath, "Prefrontal cortex and long-term memory encoding: an integrative review of findings from neuropsychology and neuroimaging," *The Neuroscientist*, vol. 13, no. 3, pp. 280–291, 2007.
- [18] M. Matsui, K. Tanaka, M. Yonezawa, and M. Kurachi, "Activation of the prefrontal cortex during memory learning: near-infrared spectroscopy study," *Psychiatry & Clinical Neurosciences*, vol. 61, no. 1, pp. 31–38, 2007.
- [19] R. Ramasubbu, H. Singh, H. Zhu, and J. F. Dunn, "Methylphenidate-mediated reduction in prefrontal hemodynamic responses to working memory task: a functional near-infrared spectroscopy study," *Human Psychopharmacology*, vol. 27, no. 6, pp. 615–621, 2012.
- [20] D. Bor, J. Duncan, R. J. Wiseman, and A. M. Owen, "Encoding strategies dissociate prefrontal activity from working memory demand," *Neuron*, vol. 37, no. 2, pp. 361–367, 2003.
- [21] S. C. Brown and F. I. Craik, "Encoding and retrieval of information," in *The Oxford Handbook of Memory*, E. Tulving and F. Craik, Eds., pp. 93–107, Oxford University Press, New York, NY, USA, 2000.
- [22] R. Cabeza, F. Dolcos, R. Graham, and L. Nyberg, "Similarities and differences in the neural correlates of episodic memory retrieval and working memory," *NeuroImage*, vol. 16, no. 2, pp. 317–330, 2002.
- [23] S. E. Prince, S. M. Daselaar, and R. Cabeza, "Neural correlates of relational memory: successful encoding and retrieval of semantic and perceptual associations," *The Journal of Neuroscience*, vol. 25, no. 5, pp. 1203–1210, 2005.
- [24] C. Ranganath, "Binding items and contexts: the cognitive neuroscience of episodic memory," *Current Directions in Psychological Science*, vol. 19, no. 3, pp. 131–137, 2010.
- [25] B. New, C. Pallier, M. Brysbaert, and L. Ferrand, "Lexique 2: a new French lexical database," *Behavior Research Methods, Instruments, and Computers*, vol. 36, no. 3, pp. 516–524, 2004.
- [26] W. R. Balch, K. Bowman, and L. A. Mohler, "Music-dependent memory in immediate and delayed word recall," *Memory and Cognition*, vol. 20, no. 1, pp. 21–28, 1992.
- [27] W. R. Balch and B. S. Lewis, "Music-dependent memory: the roles of tempo change and mood mediation," *Journal of Experimental Psychology: Learning Memory and Cognition*, vol. 22, no. 6, pp. 1354–1363, 1996.
- [28] A. Duncan, J. H. Meek, M. Clemence et al., "Measurement of cranial optical path length as a function of age using phase resolved near infrared spectroscopy," *Pediatric Research*, vol. 39, no. 5, pp. 889–894, 1996.
- [29] M. Kameyama, M. Fukuda, Y. Yamagishi et al., "Frontal lobe function in bipolar disorder: a multichannel near-infrared spectroscopy study," *NeuroImage*, vol. 29, no. 1, pp. 172–184, 2006.
- [30] M. Okamoto, H. Dan, K. Sakamoto et al., "Three-dimensional probabilistic anatomical cranio-cerebral correlation via the international 10–20 system oriented for transcranial functional brain mapping," *NeuroImage*, vol. 21, no. 1, pp. 99–111, 2004.
- [31] J. Gervain, F. Macagno, S. Cogoi, M. Peña, and J. Mehler, "The neonate brain detects speech structure," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 105, no. 37, pp. 14222–14227, 2008.
- [32] X. Cui, S. Bray, D. M. Bryant, G. H. Glover, and A. L. Reiss, "A quantitative comparison of NIRS and fMRI across multiple cognitive tasks," *NeuroImage*, vol. 54, no. 4, pp. 2808–2821, 2011.
- [33] L. J. Murray and C. Ranganath, "The dorsolateral prefrontal cortex contributes to successful relational memory encoding," *Journal of Neuroscience*, vol. 27, no. 20, pp. 5515–5522, 2007.
- [34] R. S. Blumenfeld, C. M. Parks, A. P. Yonelinas, and C. Ranganath, "Putting the pieces together: the role of dorsolateral prefrontal cortex in relational memory encoding," *Journal of Cognitive Neuroscience*, vol. 23, no. 1, pp. 257–265, 2011.
- [35] I. Kahn, A. Pascual-Leone, H. Theoret, F. Fregni, D. Clark, and A. D. Wagner, "Transient disruption of ventrolateral prefrontal cortex during verbal encoding affects subsequent memory performance," *Journal of Neurophysiology*, vol. 94, no. 1, pp. 688–698, 2005.
- [36] D. Badre and A. D. Wagner, "Left ventrolateral prefrontal cortex and the cognitive control of memory," *Neuropsychologia*, vol. 45, no. 13, pp. 2883–2901, 2007.
- [37] P. C. Fletcher and R. N. A. Henson, "Frontal lobes and human memory insights from functional neuroimaging," *Brain*, vol. 124, no. 5, pp. 849–881, 2001.
- [38] M. Petrides, B. Alivisatos, and A. C. Evans, "Functional activation of the human ventrolateral frontal cortex during mnemonic retrieval of verbal information," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 92, no. 13, pp. 5803–5807, 1995.
- [39] R. McKendrick, H. Ayaz, R. Olmstead, and R. Parasuraman, "Enhancing dual-task performance with verbal and spatial working memory training: continuous monitoring of cerebral hemodynamics with NIRS," *NeuroImage*, vol. 85, pp. 1014–1026, 2014.
- [40] H. Ohsugi, S. Ohgi, K. Shigemori, and E. B. Schneider, "Differences in dual-task performance and prefrontal cortex activation between younger and older adults," *BMC Neuroscience*, vol. 14, article 10, 2013.
- [41] M. Okamoto, Y. Wada, Y. Yamaguchi et al., "Process-specific prefrontal contributions to episodic encoding and retrieval of tastes: a functional NIRS study," *NeuroImage*, vol. 54, no. 2, pp. 1578–1588, 2011.
- [42] S. Basso Moro, S. Cutini, M. L. Ursini, M. Ferrari, and V. Quaresima, "Prefrontal cortex activation during story encoding/retrieval: a multi-channel functional near-infrared spectroscopy study," *Frontiers in Human Neuroscience*, vol. 7, 2013.
- [43] B. Opitz, A. Mecklinger, and A. D. Friederici, "Functional asymmetry of human prefrontal cortex: encoding and retrieval of verbally and nonverbally coded information," *Learning & Memory*, vol. 7, no. 2, pp. 85–96, 2000.
- [44] L. Gagnon, M. A. Yücel, M. Dehaes et al., "Quantification of the cortical contribution to the NIRS signal over the motor cortex using concurrent NIRS-fMRI measurements," *NeuroImage*, vol. 59, no. 4, pp. 3933–3940, 2012.
- [45] E. Kirilina, A. Jelzow, A. Heine et al., "The physiological origin of task-evoked systemic artefacts in functional near infrared spectroscopy," *NeuroImage*, vol. 61, no. 1, pp. 70–81, 2012.

- [46] T. Takahashi, Y. Takikawa, R. Kawagoe, S. Shibuya, T. Iwano, and S. Kitazawa, "Influence of skin blood flow on near-infrared spectroscopy signals measured on the forehead during a verbal fluency task," *NeuroImage*, vol. 57, no. 3, pp. 991–1002, 2011.
- [47] Y. Kubota, M. Toichi, M. Shimizu et al., "Prefrontal hemodynamic activity predicts false memory—a near-infrared spectroscopy study," *NeuroImage*, vol. 31, no. 4, pp. 1783–1789, 2006.
- [48] S. Cutini, S. B. Moro, and S. Bisconti, "Review: functional near infrared optical imaging in cognitive neuroscience: an introductory review," *Journal of Near Infrared Spectroscopy*, vol. 20, no. 1, pp. 75–92, 2012.
- [49] L. Ferreri, E. Bigand, S. Perrey, and A. Bugajska, "The promise of Near-Infrared Spectroscopy (NIRS) for psychological research: a brief review," *L'Année Psychologique*, vol. 114, no. 3, pp. 537–569, 2014.
- [50] M. McElhinney and J. M. Annett, "Pattern of efficacy of a musical mnemonic on recall of familiar words over several presentations," *Perceptual and Motor Skills*, vol. 82, no. 2, pp. 395–400, 1996.
- [51] S. Hamann, "Cognitive and neural mechanisms of emotional memory," *Trends in Cognitive Sciences*, vol. 5, no. 9, pp. 394–400, 2001.
- [52] M. Lövdén, M. Rönnlund, and L. G. Nilsson, "Remembering and knowing in adulthood: effects of enacted encoding and relations to processing speed," *Aging, Neuropsychology, and Cognition*, vol. 9, no. 3, pp. 184–200, 2002.
- [53] P. Fransson, G. Krüger, K.-D. Merboldt, and J. Frahm, "MRI of functional deactivation: temporal and spatial characteristics of oxygenation-sensitive responses in human visual cortex," *NeuroImage*, vol. 9, no. 6, pp. 611–618, 1999.
- [54] M. E. Raichle, A. M. MacLeod, A. Z. Snyder, W. J. Powers, D. A. Gusnard, and G. L. Shulman, "A default mode of brain function," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 98, no. 2, pp. 676–682, 2001.
- [55] N. Harel, S.-P. Lee, T. Nagaoka, D.-S. Kim, and S.-G. Kim, "Origin of negative blood oxygenation level-dependent fMRI signals," *Journal of Cerebral Blood Flow and Metabolism*, vol. 22, no. 8, pp. 908–917, 2002.
- [56] K. Sakatani, D. Yamashita, T. Yamanaka et al., "Changes of cerebral blood oxygenation and optical pathlength during activation and deactivation in the prefrontal cortex measured by time-resolved near infrared spectroscopy," *Life Sciences*, vol. 78, no. 23, pp. 2734–2741, 2006.
- [57] G. Matsuda and K. Hiraki, "Sustained decrease in oxygenated hemoglobin during video games in the dorsal prefrontal cortex: a NIRS study of children," *NeuroImage*, vol. 29, no. 3, pp. 706–711, 2006.
- [58] A.-C. Ehlis, C. G. Bähne, C. P. Jacob, M. J. Herrmann, and A. J. Fallgatter, "Reduced lateral prefrontal activation in adult patients with attention-deficit/hyperactivity disorder (ADHD) during a working memory task: a functional near-infrared spectroscopy (fNIRS) study," *Journal of Psychiatric Research*, vol. 42, no. 13, pp. 1060–1067, 2008.
- [59] M. J. Herrmann, E. Woidich, T. Schreppel, P. Pauli, and A. J. Fallgatter, "Brain activation for alertness measured with functional near infrared spectroscopy (fNIRS)," *Psychophysiology*, vol. 45, no. 3, pp. 480–486, 2008.
- [60] V. N. Salimpoor, I. van den Bosch, N. Kovacevic, A. R. McIntosh, A. Dagher, and R. J. Zatorre, "Interactions between the nucleus accumbens and auditory cortices predict music reward value," *Science*, vol. 340, no. 6129, pp. 216–219, 2013.
- [61] R. J. Zatorre, "Musical pleasure and reward: mechanisms and dysfunction," *Annals of the New York Academy of Sciences*, vol. 1337, no. 1, pp. 202–211, 2015.
- [62] B. Tillmann, "Music and language perception: expectations, structural integration, and cognitive sequencing," *Topics in Cognitive Science*, vol. 4, no. 4, pp. 568–584, 2012.
- [63] E. Nyhus and T. Curran, "Functional role of gamma and theta oscillations in episodic memory," *Neuroscience & Biobehavioral Reviews*, vol. 34, no. 7, pp. 1023–1035, 2010.

Research Article

Interaction of Musicianship and Aging: A Comparison of Cortical Auditory Evoked Potentials

Jennifer L. O'Brien,¹ Dee A. Nikjeh,² and Jennifer J. Lister²

¹Department of Psychology, University of South Florida St. Petersburg, 140 7th Street S, DAV 116, St. Petersburg, FL 33701, USA

²Department of Communication Sciences and Disorders, University of South Florida Tampa, 4202 E. Fowler Avenue, Tampa, FL 33621, USA

Correspondence should be addressed to Jennifer L. O'Brien; jenobrien@usf.edu

Received 15 February 2015; Revised 11 April 2015; Accepted 14 April 2015

Academic Editor: Stefan Evers

Copyright © 2015 Jennifer L. O'Brien et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Objective. The goal of this study was to begin to explore whether the beneficial auditory neural effects of early music training persist throughout life and influence age-related changes in neurophysiological processing of sound. **Design.** Cortical auditory evoked potentials (CAEPs) elicited by harmonic tone complexes were examined, including P1-N1-P2, mismatch negativity (MMN), and P3a. **Study Sample.** Data from older adult musicians ($n = 8$) and nonmusicians ($n = 8$) (ages 55–70 years) were compared to previous data from young adult musicians ($n = 40$) and nonmusicians ($n = 20$) (ages 18–33 years). **Results.** P1-N1-P2 amplitudes and latencies did not differ between older adult musicians and nonmusicians; however, MMN and P3a latencies for harmonic tone deviances were earlier for older musicians than older nonmusicians. Comparisons of P1-N1-P2, MMN, and P3a components between older and young adult musicians and nonmusicians suggest that P1 and P2 latencies are significantly affected by age, but not musicianship, while MMN and P3a appear to be more sensitive to effects of musicianship than aging. **Conclusions.** Findings support beneficial influences of musicianship on central auditory function and suggest a positive interaction between aging and musicianship on the auditory neural system.

1. Introduction

The perception of music infuses the human brain with a rich auditory sensory experience and is a cognitively complex task requiring the integration of multiple cortical levels and neural systems. Because of the intense training and skill acquisition a musician receives from an early age, a musician's brain provides unique opportunities to explore the impact of music perception and music training on neural structural and functional adaptation and development.

Ground-breaking anatomic studies using functional magnetic resonance imaging (fMRI) [1, 2] revealed that musicians compared to nonmusicians have 5% greater cerebellar volume, significantly larger anterior corpus callosum, and increased gray matter volume in the left Heschl's gyrus and left inferior frontal gyrus. In addition to altered cortical structure, electroencephalographic (EEG) and magnetoencephalographic (MEG) recordings of cortical auditory

evoked potentials (CAEPs) parallel anatomic studies showing cortical enlargement of auditory areas important for music perception [3, 4] and reveal superior preattentive auditory sensory memory representations for musicians compared to nonmusicians [5–14]. Specifically, findings suggest that musicians relative to nonmusicians (a) discriminate auditory differences that are undetectable for nonmusicians at a preattentive processing level implying that sensory memory traces containing auditory information may be enhanced by musical expertise [8], (b) discriminate embedded pitch shifts faster (shorter latencies) within familiar and unfamiliar interval patterns [5], (c) have larger amplitude responses to changes in relative pitch structure, such as melodic contour and interval [6], rhythmic deviation [15], and timbre of expertise instrument [16], (d) respond faster (shorter latencies) to pitch deviances of harmonic tone complexes [9, 12], and (e) enhanced sensitivity to acoustic changes of harmonic complexes and speech syllables [10, 13, 14]. Interestingly,

effects of musicianship on the plasticity of CAEPs can be seen even after short-term (e.g., two weeks) musical training (e.g., [17]).

These remarkable effects of music training on auditory neural structure and function have been observed from childhood through young adulthood; however, research is limited when it comes to exploring whether these auditory neural effects of music training extend throughout the life of the musician. Given their history of extraordinary exposure to and experience with spectrally and temporally rich auditory stimuli at an early age, it is questioned whether older adult musicians might experience age-related changes in central auditory processing differently than older nonmusicians.

It is a well-known fact that older adults often complain that although they can hear speech, they cannot understand it. Age-related physiological changes typically alter the way in which spectral and temporal information are encoded [18]. Older adults often exhibit neural activation patterns that are qualitatively different and more frontally oriented than those of younger adults [19–21]. These patterns suggest relatively slower neural transit time and altered auditory inhibition/arousal by irrelevant stimuli for older adults and characterize an inefficient aging auditory system that contributes to poor speech understanding in noisy, real-world listening environments [21]. Given that young adult musicians have altered auditory neural structures and superior auditory neural processing for music and speech, is it then possible that older adult musicians may have a more efficient auditory system than their nonmusician cohorts?

Investigations of age-related changes affecting auditory neural processing have used both behavioral and electrophysiological measures (e.g., EEG, MEG). Psychoacoustic measures provide information about the listener's perception of sound and reflect a conscious attentional process. Recent behavioral studies exploring the rate of age-related decline on peripheral and central auditory processing between older adult musicians and age-comparable nonmusicians (ages 45–65) reported that older musicians demonstrate increased auditory working memory [22] and less age-related decline for gap-detection and speech-in-noise thresholds than nonmusicians [20]. Findings suggest that musical training may diminish the impact of age-related auditory decline [22] and, further, that musicians may experience less age-related decline in central auditory processing than nonmusicians [23]. We question what impact, if any, the interaction between musicianship and aging has on preattentive neurophysiological processing of sound for older adults.

Preattentive cortical auditory evoked responses (e.g., P1-N1-P2, MMN, and P3a), the focus of the current study, are particularly suitable when investigating responses of older adults because they reflect the automatic detection and discrimination of acoustic sensory memories prior to attentional focus without being contaminated by attention, motivation, or cognitive demands of the task [24, 25]. CAEPs provide excellent temporal resolution and allow for a noninvasive evaluation of the various stages of auditory processing from preattentive sensory perception to later cognitive levels. To date, CAEP studies comparing older musicians and nonmusicians are unknown.

CAEP studies investigating age-related changes of central auditory processes in older nonmusician listeners have reported similar aging effects on P1-N1-P2 elicited by pure tones [19] and gap-detection [21]. The P1-N1-P2 complex reflects the physiological detection of audible stimulus energy [28]. In both studies comparing older and younger adults with normal hearing, there was a pattern of larger P1 amplitudes and slower P2 latencies for the older adults. For gap-detection, P1 latency was earlier for older adults than younger; however, there was no P1 latency difference between older and younger adults for detection of pure tones. In neither study did N1 or P2 amplitudes or the N1 latency differ by age group. It may be speculated that larger P1 amplitude and earlier P1 latency as well as longer P2 latencies may represent an adaptive neural response to aging. That is to say, compared to younger adults, older adults may activate or recruit more or different neural resources to increase the efficiency of stimulus detection and inhibition and may have slower processing or extended neural conduction time once the auditory stimulus is detected.

Mismatch negativity (MMN) is a passively elicited negative potential typically considered to be independent of attention and higher cognitive processing [29]. In general, the MMN is thought to be a preperceptual or preattentive central auditory response to an acoustic deviation based on the detection of an auditory regularity in a preceding sound sequence and an automatic sensory memory-based comparison process [30, 31]. Indexing in sensory memory as reflected by the MMN is considered to be a preattentive because it has been shown to occur in the absence of attention and even conscious awareness [30]. Note that, while the MMN can be elicited outside of attention, it can also be influenced by attentional modulation (see [47] for modeling evidence of temporal attention enhancing the MMN).

While the MMN can occur preattentively, the P3a, which often follows the MMN, is associated with an involuntary and automatic shift of attention to and conscious perception of a deviant or new stimulus [30, 32, 33]. Typically, the P3a is elicited in response to infrequent task-irrelevant stimuli that are unexpected (e.g., pitch “C”) in a sequence of frequently presented standard stimuli (e.g., pitch “A”) within which the listener is actively attending to the detection of infrequent target stimuli (pitch “B”). The infrequent presence of pitch “C” automatically shifts the listener's attention to that tone, resulting in a P3a.

Age-related studies using CAEPs to elicit MMN and P3a are limited; however, results are consistent. Comparisons of older and younger adults suggest that MMN is affected more by stimulus contrast and presentation rate than age alone [19, 26, 34]. Age alone did not affect the MMN latency or amplitude for stimulus-change detection of frequency [19] or stimulus duration with a short interstimulus interval (ISI = 0.5 sec); rather, as the time between the stimuli increased, the MMN attenuated more for older listeners than younger suggesting that MMN may reflect the gradual decay of the stimulus trace in the auditory system and that this trace decays faster for older than younger adults [26, 34]. We are aware of only two studies of P3a and aging for auditory stimuli. Those studies indicate that P3a is absent [27]

TABLE 1: Summary of older musician musical history.

Musician	Age (years)	Age training initiated (years)	Instrument/voice	Total years training	Total years music making	Average hours per week
OM1	61	8	I	10	52	4
OM2	67	8	I	7	59	2
OM3	61	8	I/V	20	29	2
OM4	58	10	I/V	8	20	0
OM5	65	13	I/V	17	52	13
OM6	64	8	I	8	50	2
OM7	59	16	I/V	8	43	3
OM8	65	8	I/V	14	57	8
Mean	62.5 yrs	9.87 yrs		11.5 yrs	45.25 yrs	4 h 15 mins/wk
Range	58–67 yrs	8–13 yrs		7–20 yrs	20–59 yrs	0–13 hrs

or delayed [35] for older as compared to younger adults, indicating that older adults are slower to switch attention to and evaluate distracting auditory stimulus outside of their focused attention.

Our current research sought to examine cortical auditory evoked responses of older musicians and is based on the framework of our earlier research with young adult musicians [9, 10]. It was an exploratory study to investigate the interaction between the aging process and possible effects of early music training and the possible impact of this interaction on the neurophysiological processing of sound. Preattentive CAEPs elicited by harmonic tones are compared between normal hearing older musicians and their nonmusician cohorts. In addition, CAEP data from older musicians and nonmusicians is compared to that of their younger counterparts [9] to further elucidate effects of age-related changes on central auditory processing. Specifically, the P1-N1-P2 complex was elicited by a standard harmonic tone stimulus to establish the physiological detection of sound at the level of the auditory cortex. MMN and P3a were elicited by small deviant frequencies (i.e., 1.5% and 6% lower in frequency than a standard) of a harmonic tone complex. Based on previous research of the P1-N1-P2 potentials [10, 19, 21] and the assumption that larger P1 amplitude reflects increased neural recruitment for regulating incoming auditory stimuli in older adults, it was predicted that, given their musical expertise and training, older musicians may have smaller P1 amplitudes for harmonic tones than older nonmusicians and similar group latencies. We also predicted a shorter P2 latency for older musicians compared to older nonmusicians, given the general increase in P2 latency with age and reduction effects of musicianship on aging.

Based on previous electrophysiological research among young adult musicians and nonmusicians, it was predicted that older musicians may have shorter MMN and P3a latencies and larger MMN amplitudes than age-comparable nonmusicians for harmonic complexes; however, no significant effects of musicianship were predicted for P3a amplitudes. MMN and P3a amplitude were predicted to increase as the size of frequency deviance increased. Based on the literature described above, the following aging effects were predicted

for P1-N1-P2, MMN, and P3a: (1) P1 amplitude would be larger; (2) P1 latency is earlier and P2 latency later for older adults compared to younger; (3) neither P2, N1, or MMN amplitudes nor MMN latency would be affected by age; and (4) P3a amplitude would be smaller and latency later for older compared to younger adults.

2. Method

2.1. Participants. Eight older adult musicians (mean age = 62.5 yrs, range = 58–67 yrs) and eight older nonmusicians (mean = 61.4 yrs, range = 55–70 yrs) were recruited from local communities, community music organizations, and the University of South Florida (USF). Musicians averaged 11.5 years of formal music training that began between the ages of 8 and 16 (see Table 1 for musical history). Nonmusicians had fewer than 12 months of music training. For the purpose of this investigation, formal music training refers to a minimum of 6 years of professionally directed and implemented music instruction and technical exercises provided by a professional musician and/or music educator [36]. All participants were right-handed and native speakers of English and had normal pure-tone thresholds (less than or equal to 25 dB HL bilaterally) for the frequencies of interest in the present study (≤ 3000 Hz). Performance on the Words-in-Noise (WIN) [37, 38] test, in terms of the dB signal-to-noise ratio for 50% correct word recognition in noise, was age-appropriate for all participants. A summary of hearing thresholds and WIN scores, shown in Table 2, indicates that the hearing sensitivity of the two older adult groups was highly similar. None had a history of neurological impairment, absolute pitch ability, exposure to tone languages, or previous participation in psychoacoustic experiments. The study was approved by the USF Institutional Review Board and documented informed consent was obtained from all subjects. Participants were reimbursed \$15 per hour of participation.

The identical inclusion and exclusion criteria were used for our earlier recruitment of 40 young adult musicians and 20 nonmusicians [9]. Young adult musicians were between the ages of 18 and 33 (mean age = 22 years) and had an average of 9.8 years (median = 9.18 yrs, range = 5 to 17 yrs) of formal

TABLE 2: Mean pure-tone hearing thresholds (dB HL) and 50% signal-to-noise ratios (dB) for the Words-in-Noise (WIN; [38, 39]) test plus standard deviations for each older participant group.

	WIN	Ear	Frequency (Hz)							
			250	500	1000	2000	3000	4000	6000	8000
Older nonmusicians ($n = 8$)	5.3 (2.4)	Right	15 (5)	14 (5)	15 (6)	16 (7)	16 (3)	26 (17)	26 (9)	30 (22)
		Left	13 (3)	16 (5)	13 (7)	16 (6)	22 (6)	26 (15)	29 (12)	26 (13)
Older musicians ($n = 8$)	5.6 (2.7)	Right	18 (6)	14 (5)	14 (8)	14 (13)	16 (12)	16 (8)	28 (11)	31 (23)
		Left	13 (3)	16 (6)	15 (8)	14 (11)	13 (9)	19 (9)	26 (11)	31 (17)

music training. Nonmusicians were between the ages of 20 and 34 (mean age = 23) and had fewer than 12 months of music training.

2.2. Stimuli. Harmonic complexes were digitally generated (sampling rate = 50 kHz), controlled, and presented using a Tucker-Davis Technologies (TDT) RP2 Real-Time Processor with model HB 7 headphone buffer. Duration of all stimuli was 200 ms and shaped with a \cos^2 window to create a 10 ms rise/fall time. Interstimulus interval (ISI) was 500 ms. All stimuli were presented bilaterally via Etymotic Research (ER2) insert earphones at 75 dB SPL.

The present study is a continuation of a series of studies investigating the relationship between pitch perception and vocal production by types of musicians (i.e., vocalists, string instrumentalists, and wind instrumentalists) and nonmusicians [10]. Thus, auditory stimuli consisted of harmonic tone complexes containing fundamental frequencies (F_0) that occur within the mid-frequency range of the female vocal register, $F_0 = 261.63$ Hz to 392 Hz (C4 to G4) [39]. The standard tone was G4, $F_0 = 392$ Hz, and was chosen from a behavioral task in a previous study because this tone elicited the best overall difference limen for frequency (DLF) across groups [10]. The two deviant harmonic complexes, $F_0 = 386$ Hz (HMD1.5), 1.5% difference from the standard tone F_0 (an eighth tone difference) and $F_0 = 370$ Hz (HMD6), 6% difference from the standard tone F_0 (a semitone difference), were selected based on DLFs measured behaviorally in a companion study and represented a range of behavioral performance. In physical terms, the interval between adjacent whole tones is a 12% difference between the fundamental frequencies of each tone. The interval between a reference tone and its semitones is a fundamental frequency difference of roughly 6% (e.g., G4 to F4# equals 392 Hz to 370 Hz). Each harmonic tone stimulus contained a F_0 and three harmonics. The amplitude of each harmonic was divided by its harmonic number to create a natural amplitude contour in the frequency spectrum.

2.3. Procedure

2.3.1. Electroencephalographic Recording. CAEPs were recorded and analyzed using a Compumedics Neuroscan

EEG system with a SynAmps 2 amplifier and Neuroscan Scan 4.3 acquisition software. A cap of 64 sintered electrodes was placed on the subject's head with additional electrodes above and below the left eye and at the outer canthus of each eye to monitor eyeblink activity. The nose served as reference and the forehead served as ground. To minimize any auditory attentive behavior, the participant was comfortably seated in a sound-attenuated booth and instructed to watch a closed-caption movie of choice and to ignore the auditory stimuli. All impedances were kept below 30 k Ω . The acquisition of EEG data was by continuous sampling, recorded at an AD (analog to digital) sampling rate of 1000 Hz and stored on the computer for offline averaging. The raw signal was amplified within a frequency band of 0.05–100 Hz.

The harmonic tone condition was designed as a multideviant oddball paradigm to elicit MMN and P3a. The protocol consisted of one standard tone ($F_0 = 392$ Hz) and two infrequently occurring deviant tones (HMD6 and HMD1.5). The standard tone occurred on a minimum of 75% (1080 minimum) of the trials and the two deviant frequencies occurred on 25% (180 per deviant, 360 total deviants) of the trials. Stimuli were presented in a pseudorandom sequence with at least three standard stimuli separating presentations of deviant stimuli; thus, two deviant stimuli did not occur in succession. The response to the standard stimulus was analyzed to establish physiological detection of the auditory stimuli (P1-N1-P2).

2.3.2. Data Analysis. Offline analysis of the continuous EEG waveforms was conducted using Neuroscan Scan 4.3 Edit software and began with manual artifact rejection. As a precaution for data analysis, the first 10 CAEP responses were omitted from the averaging process to exclude the variation of the N1 amplitude (i.e., the refractoriness) associated with the start of the stimulation sequence [40, 41]. EEG epochs of 700 ms (–100 to 600 ms) were obtained, baseline corrected (–100–0 ms), and averaged separately for the standard and deviant stimuli. To eliminate ocular movement contamination, epochs containing artifacts exceeding ± 80 μ V in the horizontal and vertical eye channels were rejected from averaging. CAEP waves were digitally band-pass filtered at 0.1–30 Hz with a squared Butterworth zero-phase filter (12 dB/octave roll-off). In order to maximize signal-to-noise

ratio at Fz, all of the processed average files for MMN/P3a analysis were individually rereferenced to the mastoids.

P1-N1-P2 response had the largest amplitudes measured at Fz; thus, reported measures and statistical analysis are based on CAEP responses measured from the Fz electrode for the standard stimulus. P1 was identified as the first positive peak occurring between 25 and 90 ms in the group average CAEP waveforms [17]. The N1 was defined as the largest negativity occurring between 70 and 140 ms. The P2 was defined as the largest positivity occurring between 140 and 255 ms. Latency windows of ± 25 ms around each amplitude peak for each group were determined and individual peak amplitudes and latencies were quantified within these preselected windows using scripts within the Neuroscan Scan 4.3 Edit software; the script selected the time point with the highest (for P1 and P2) or lowest (for N1) amplitude value within the designated window. Visual inspection of individual waveforms was conducted to ensure that clear peaks did not fall outside of the designated windows for any participant.

The MMN is illustrated by a difference wave obtained by subtracting the averaged CAEP elicited by the standard stimulus from the averaged CAEP elicited by a deviant stimulus [42]. Difference waveforms to illustrate MMN and P3a were calculated for each deviant stimulus condition. The MMN response was largest at electrodes Fz and Cz with the largest amplitudes measured at Fz. Thus, reported measures and statistical analysis are based on CAEP responses measured from Fz. The MMN was verified by polarity inversion at the mastoids prior to rereferencing all individual files to the mastoids. MMN has been shown to invert in polarity at electrodes below the level of the Sylvian fissure [42]. Polarity inversion at the mastoids is an accepted method to verify the MMN response to tonal changes [41]. CAEP amplitudes were quantified by first determining the peak latencies from the grand-average difference waves separately for each deviant as the largest peak between 100 and 300 ms at Fz for MMN [43]. The P3a was chosen as the first positive peak following the MMN.

Grand average group difference waveforms for each harmonic tone deviant stimulus were derived from the Fz electrode for purposes of illustrating MMN and P3a and for selecting latency windows. Latency windows of ± 25 ms around each amplitude peak for each group and deviant condition were determined [44]. MMN and P3a peak amplitudes and latencies for each participant were quantified within these preselected windows for each harmonic deviant condition using scripts within the Neuroscan Scan 4.3 Edit software to select the time point with the highest (for P3a) or lowest (for MMN) amplitude value within the designated window. Visual inspection of individual waveforms was conducted to ensure that clear peaks did not fall outside of the designated windows for any participant.

3. Results

Peak amplitude and latency data were analyzed using IBM SPSS Statistics software (version 21.0). To determine the effect of age and musicianship on P1-N1-P2, MMN, and P3a,

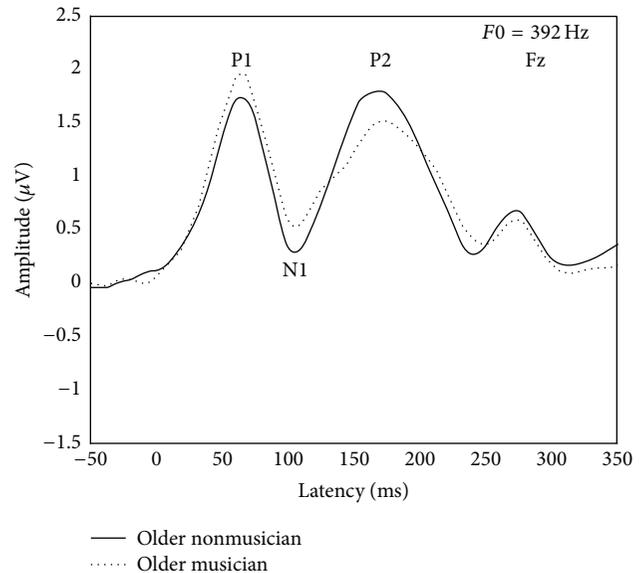


FIGURE 1: Group averaged CAEP waveforms illustrating P1, N1, and P2 elicited from older nonmusicians (solid line) and older musicians (dotted line) by the standard harmonic tone ($F_0 = 392$ Hz).

univariate analyses of variance (ANOVAs) were conducted on individual mean component latencies and amplitudes. An alpha level of 0.05 was used for all statistical tests.

3.1. Effects of Musicianship within Older Adults

3.1.1. P1-N1-P2 Components. Analysis of P1-N1-P2 was conducted at Fz for the standard harmonic tone stimulus ($F_0 = 392$ Hz). The peaks are designated in Figure 1 for visual inspection and the latency and amplitude means and standard deviations are reported in Table 3. There were no significant group differences of P1, N1, and P2 latencies and amplitudes between older adult musicians and nonmusicians, $F(1,14) = 1.19$, $p = 0.293$ for P1 latency; $F(1,14) = 0.18$, $p = 0.676$ for P1 amplitude; $F(1,14) = 0.71$, $p = 0.412$ for N1 latency; $F(1,14) = 0.34$, $p = 0.567$ for N1 amplitude; $F(1,14) = 0.11$, $p = 0.742$ for P2 latency; and $F(1,14) = 0.23$, $p = 0.638$ for P2 amplitude.

3.1.2. Mismatch Negativity (MMN). The MMN and P3a were elicited from older musicians and nonmusicians by two harmonic tone pitch deviants and are illustrated by comparing CAEPs to the standard stimulus and CAEPs to the deviant stimuli elicited from the oddball paradigm (Figure 2). The MMN and P3a latency and amplitude means and standard deviations are reported in Table 4. The grand average group CAEP difference waveforms reflect the MMN and P3a for the two frequency deviances (Figure 3). Parallel two-way repeated-measures ANOVAs were conducted on individual mean MMN and P3a latencies and amplitudes for each stimulus condition with the within-subject factor Stimulus (HMD1.5 [$F_0 = 386$ Hz], HMD6 [$F_0 = 370$ Hz]) and the between-subject factor Musicianship (older nonmusicians,

TABLE 3: Mean and standard deviation values for P1-N1-P2 cortical auditory evoked potential latencies and amplitudes elicited by standard harmonic tone ($F0 = 392$ Hz) from older and younger musicians and nonmusicians.

	P1		N1		P2	
	Latency (ms)	Amplitude (microvolts)	Latency (ms)	Amplitude (microvolts)	Latency (ms)	Amplitude (microvolts)
Older nonmusicians	65 (8)	2.02 (1.1)	104 (10)	-0.02 (0.95)	171 (15)	2.0 (1.4)
Older musicians	61 (8)	2.24 (0.8)	109 (12)	0.27 (1.1)	174 (15)	1.69 (1.0)
Younger musicians	74 (10)	1.84 (0.75)	105 (11)	0.21 (0.88)	147 (10)	2.21 (1.1)
Younger nonmusicians	74 (10)	2.00 (0.76)	103 (11)	0.199 (1.2)	146 (11)	2.43 (0.81)

TABLE 4: Mean and standard deviation values for MMN and P3a cortical auditory evoked potential latencies and amplitudes elicited by deviant harmonic tones (HMD 1.5 and HMD 6) from older and younger musicians and nonmusicians.

		MMN		P3a	
		Latency (ms)	Amplitude (microvolts)	Latency (ms)	Amplitude (microvolts)
Older nonmusicians	HMD 1.5	298 (15)	-1.1 (1.3)	330 (20)	-0.06 (1.3)
	HMD 6	200 (19)	-1.62 (0.9)	283 (13)	0.81 (1.1)
Older musicians	HMD 1.5	238 (12)	-1.26 (0.5)	314 (18)	-0.18 (0.8)
	HMD 6	172 (15)	-2.6 (1.7)	251 (16)	0.19 (1.5)
Younger nonmusicians	HMD 1.5	219 (18)	-1.90 (1.1)	297 (14)	0.43 (1.5)
	HMD 6	183 (12)	-3.44 (1.4)	265 (16)	1.47 (1.4)
Younger musicians	HMD 1.5	213 (14)	-2.21 (1.3)	284 (14)	0.20 (1.3)
	HMD 6	163 (15)	-2.63 (1.5)	259 (15)	2.04 (1.6)

older musicians). MMN latencies to harmonic tone pitch deviants were significantly shorter for older adult musicians than older nonmusicians, $F(1,14) = 7.96$, $p < 0.001$. Further, there was a significant effect of deviant magnitude, $F(1,14) = 200.68$, $p < 0.001$ for MMN latency. Across groups, as the magnitude of the frequency deviance increased, response latency decreased (Figure 3; see Table 4 for mean group latency values). The interaction between Musicianship and Stimulus was also significant, $F(1,14) = 7.57$, $p = 0.016$ for MMN latency. A post hoc analysis of the significant interaction revealed that the effect of Stimulus on MMN latency was significant for both older nonmusicians ($p < 0.001$) and older musicians ($p < 0.001$). The effect of Musicianship on MMN latency was also significant for both HMD1.5 ($p < 0.001$) and HMD6 ($p = 0.005$).

In contrast, MMN amplitude did not differ significantly by Musicianship, $F(1,14) = 1.66$, $p = 0.218$; however, there was a positive effect of deviant magnitude, $F(1,14) = 5.35$, $p = 0.036$. For all, as the magnitude of the frequency deviance increased, response amplitude increased. Interaction between Musicianship and Stimulus was not significant, $F(1,14) = 1.04$, $p = 0.326$ for MMN amplitude.

3.1.3. P3a Component. Overall, P3a latency occurred significantly earlier for older adult musicians than older nonmusicians, $F(1,14) = 12.65$, $p = 0.003$ (Figure 3). For all participants, the larger frequency deviance elicited a shorter latency response, $F(1,14) = 114.90$, $p < 0.001$ (Table 4). There was no interaction between group membership and the magnitude of the frequency deviant, $F(1,14) = 2.64$, $p = 0.126$. P3a amplitude was not dependent on Musicianship,

$F(1,14) = 0.59$, $p = 0.454$, nor magnitude of deviance, $F(1,14) = 3.07$, $p = 0.102$. The interaction was not significant, $F(1,14) = 0.50$, $p = 0.492$.

3.2. Effects of Age and Musicianship. Amplitude and latency of CAEPs P1-N1-P2, MMN, and P3a, elicited from older adult musicians and nonmusicians by harmonic tone complexes, were compared to CAEPs previously elicited from young adult musicians and nonmusicians [9] in univariate ANOVAs with Age (younger, older) and Musicianship (musicians, nonmusicians) as the between-subject factors. CAEP data from older and younger musicians is compared to further examine and elucidate the effects of age-related changes on central auditory processing. All of the previously collected data from young adult musicians and nonmusicians was reanalyzed using methods described in this paper.

3.2.1. P1-N1-P2 Components. P1-N1-P2 latency and amplitude means and standard deviations for all four groups are reported in Table 3. P1 latencies were significantly affected by Age, $F(1,72) = 14.93$, $p < 0.0011$, with earlier latencies for older adults than younger listeners, but was not affected by Musicianship, $F < 1$. There were no significant effects of Age or Musicianship or interaction between the two for P1 amplitude, N1 latency, or N1 amplitude, $F_s < 1$. P2 latencies were significantly affected by Age, $F(1,72) = 64.75$, $p < 0.001$, with significantly earlier latencies for younger listeners compared to older listeners, but was not affected by Musicianship or the interaction of Age and Musicianship, $F_s < 1$. P2 amplitude was not affected by Age, $F(1,72) = 2.44$, $p = 0.112$, or Musicianship or the interaction of the two,

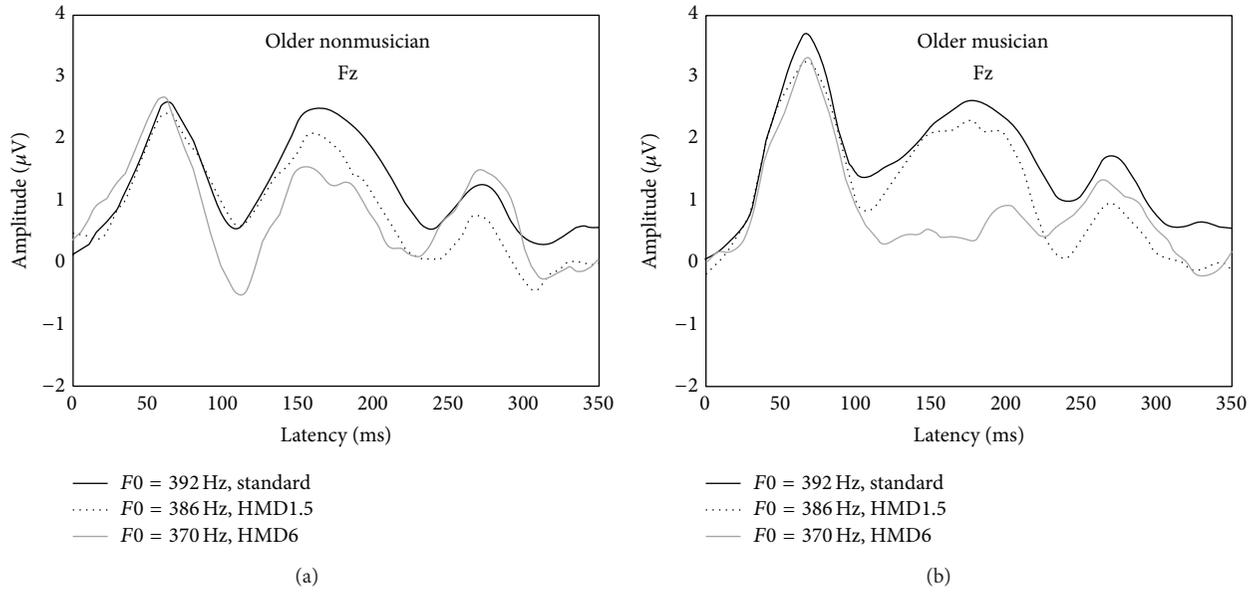


FIGURE 2: Grand average group CAEP waveforms at Fz for older nonmusicians (a) and older musicians (b). A solid black line represents the CAEP to the standard stimulus while the dotted and gray lines represent CAEPs to deviants presented in the oddball paradigm (i.e., HMD1.5 and HMD6).

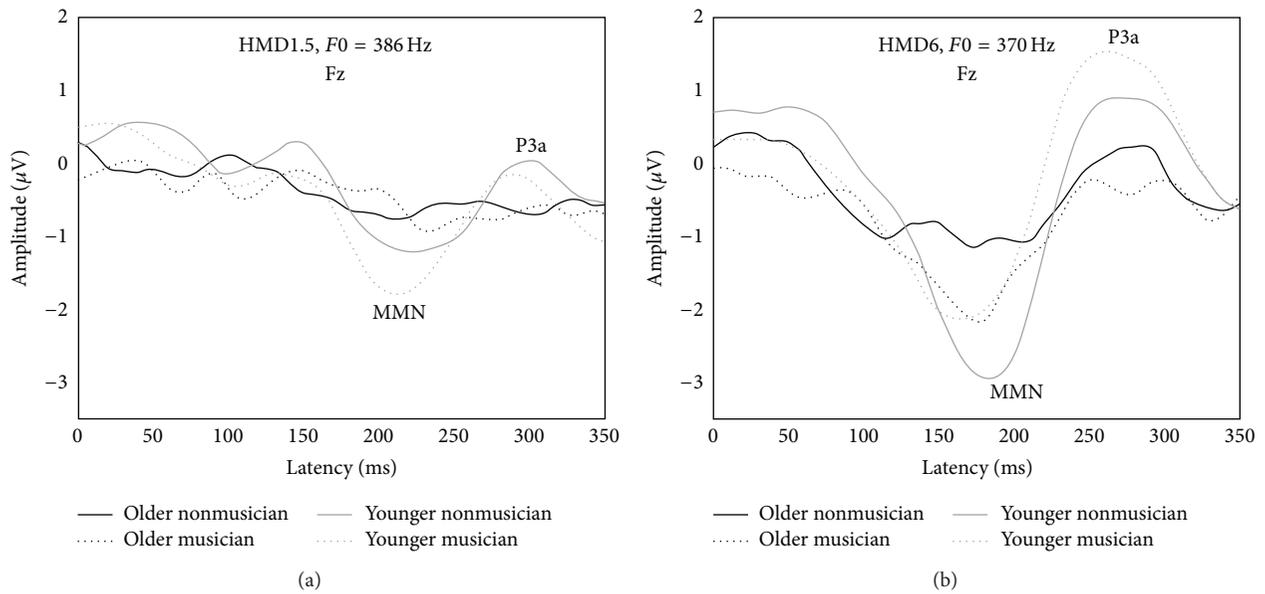


FIGURE 3: Grand average group difference waveforms (deviant minus standard) at Fz from older adult musicians (OM), older nonmusicians (ONM), young musicians (YM), and young nonmusicians (YNM) in response to a 1.5% frequency deviance (HMD1.5, (a)) and a 6% frequency deviance (HMD6, (b)).

$F_s < 1$. Averaged group waveforms for P1-N1-P2 components are illustrated in Figure 4.

3.2.2. *Mismatch Negativity (MMN)*. MMN latency and amplitude means and standard deviations are reported in Table 4. There were significant effects of Age for MMN latency: younger participants had shorter latencies than older participants for HMD1.5, $F(1,72) = 155.81, p < 0.001$, and

HMD6, $F(1,72) = 10.81, p = 0.002$. Younger participants also had significantly larger MMN amplitudes than older participants for HMD1.5, $F(1,72) = 6.21, p = 0.015$, and for HMD6 $F(1,72) = 5.43, p = 0.023$. There were also significant effects of Musicianship for MMN latency: musicians had shorter latencies than nonmusicians for HMD1.5, $F(1,72) = 60.75, p < 0.001$, and HMD6, $F(1,72) = 33.74, p < 0.001$. Effects of Musicianship on MMN amplitude for both stimuli were not significant, $F_s < 1$.

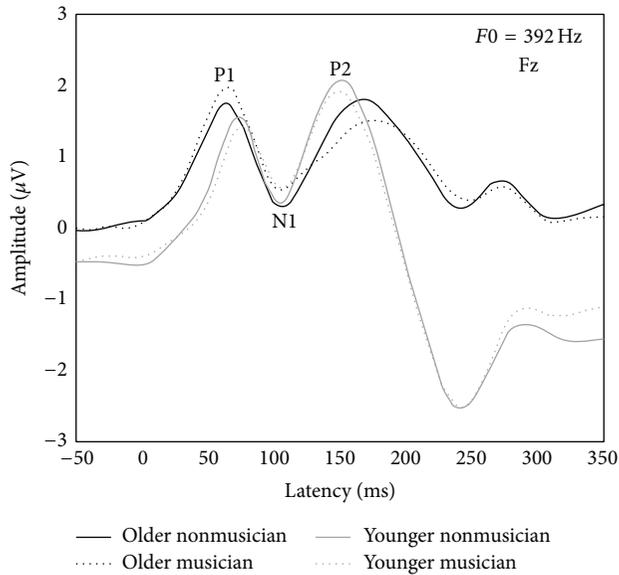


FIGURE 4: Group averaged CAEP waveforms from older (black lines) and young (gray lines) adult nonmusicians (solid lines) and musicians (dotted lines) illustrating P1-N1-P2 at Fz elicited by the standard harmonic tone ($F_0 = 392$ Hz).

MMN latency effects for HMD1.5 were qualified by an Age \times Musicianship interaction, $F(1,72) = 42.44$, $p < 0.001$. The two young adult groups (YM, YNM) had very similar, short latencies; the OMs had slightly longer latencies and the ONMs had the longest latencies of all. The latencies of the YM and YNM groups were not significantly different, $F(1,72) = 1.77$, $p = 0.187$, but all other groups showed latencies significantly different from each other, $ps < 0.001$. MMN amplitude effects for HMD6 were qualified by an Age \times Musicianship interaction, $F(1,72) = 4.45$, $p = 0.038$. The largest amplitude overall was found for the YNM group while the amplitudes of the two musician groups (YM, OM) were very similar and smaller than that of the YNM. The ONM group showed the smallest overall amplitude. The difference between the YMs and OMs was not significant, $F < 1$, but YNMs had significantly larger amplitudes than ONMs, $F(1,72) = 9.15$, $p = 0.003$.

3.2.3. P3a Component. P3a latency and amplitude means and standard deviations are reported in Table 4. There were significant effects of Age on P3a latency for HMD1.5, $F(1,72) = 55.00$, $p < 0.001$, and on P3a amplitude for HMD6, $F(1,72) = 8.90$, $p = 0.004$. There were also significant effects of Musicianship for P3a latency: musicians had shorter latencies than nonmusicians for HMD1.5, $F(1,72) = 11.52$, $p = 0.001$, and HMD6, $F(1,72) = 21.06$, $p < 0.001$. Effects of Musicianship on P3a amplitude for both stimuli were not significant, $F_s < 1$. P3a latency effects for HMD6 were qualified by an Age \times Musicianship interaction, $F(1,72) = 9.21$, $p = 0.003$. P3a latency for HMD6 was shortest for the OM group, followed closely by the YM group, and longest latencies were observed for the YNM and ONM groups, respectively. P3a latency for the YM group was not significantly different from that of the

OM or YNM groups but all other group comparisons were significant.

4. Discussion

Musicians from early childhood through young adulthood have been shown to have enhanced auditory sensory-memory and sensitivity to acoustic changes for music and speech; however, there is a paucity of research exploring whether older musicians retain this acoustic processing advantage throughout life. This was an exploratory study using preattentive CAEPs, specifically the P1-N1-P2, MMN, and P3a components, to compare responses between formally trained older musicians and an age-comparable group of nonmusicians. Moreover, to investigate the interaction between aging and musicianship, CAEPs from the older musicians and nonmusicians in the current study were compared to CAEPs of young adult musicians and nonmusicians from an earlier study [9]. As an exploratory study, the sample size for the older adult groups was small (n 's = 8) compared to the younger adult group previously published (musicians $n = 40$; nonmusicians $n = 20$) [9, 10]; however, the data are useful in showing neurophysiological responses of musicianship that are age-agnostic as well as responses that are age-specific.

Consistent with our hypotheses and similar to our findings among young adults, preattentive neurophysiological responses to music stimuli do distinguish older adult musicians from older nonmusicians. Unique to this investigation is the added comparison of preattentive CAEPs between groups of young and older musicians and nonmusicians and findings that further differentiate the obligatory evoked potentials from the change-detection potentials. It appears that the early obligatory P1-N1-P2 complex may be more affected by aging than musicianship, whereas the inverse is true of the preattentive responses representing the change-detection paradigm. That is, MMN and P3a derived components appear to be influenced by musicianship rather than aging.

4.1. P1-N1-P2 Complex. As predicted, older adults regardless of music background detected the presence of acoustic energy earlier than younger adults (i.e., shorter P1 latency); yet once arousal was triggered, neural conduction was longer from P1 to P2 (i.e., later P2 latency) for older adults than the younger. Findings are consistent with others [19, 21] and suggest perhaps an age-related adaptive neural response that increases neural recruitment to initiate auditory arousal to the presence of acoustic energy. Once aroused, the delayed time-course may be explained by age-related refractory-time differences between younger and older slowing of synchronous neural firing in the central auditory pathways. That is, once the neurons are activated, the older system takes longer to recover before the neurons can fire again. This increase in refractory time results in slower neural travel time.

We expected an effect of both aging and musicianship on P1 amplitude and found neither. We also expected an effect of musicianship on P2 latency in older adults, which was not present in this data set. As for the effect of aging on P1

amplitude, it is well documented that older listeners typically demonstrate an increase in the strength of the obligatory response (P1 amplitude) as a result of an adaptive increase in neural recruitment for the physiological detection of the auditory stimulus [18, 19, 21]. Thus, it was predicted that, as a group, the older listeners would have larger P1 amplitude than the young listeners. As for the influence of musicianship, smaller P1 amplitude has been shown to reflect reduced processing demands [45]. Therefore, it was speculated that highly trained listeners, such as musicians, would be more efficient at inhibiting irrelevant stimuli and regulating familiar stimuli (i.e., harmonic tones) than untrained listeners and, thus, musicians were predicted to have smaller P1 amplitude than nonmusicians. One plausible explanation for the absence of significant differences may be the choice of stimulus. The P1-N1-P2 complex was elicited by a single repetitive harmonic tone. A review of aging studies indicates that both the latency and amplitude patterns of P1, N1, and P2 may be altered depending on the level and spectrum of the stimulus as well as the presentation paradigm (for review see [46]). It is plausible that an adaptive increase in neural activation (i.e., larger P1 amplitude) may have been unnecessary to perceive this less complex auditory stimulus; whereas, previously reported P1 amplitude findings have been linked to more complex stimuli such as neural detection of gaps in noise [21], pure tones with and without noise [19], and detection of consonant-vowel speech syllables [25]. Another possible explanation may be due to the small number of older participants compared to the larger group of young participants. Future research will have larger numbers of older participants more compatible in size to the younger groups.

4.2. MMN and P3a Components. MMN and P3a derived components appear to be more sensitive to neural effects of musicianship than aging. Consistent with our hypotheses, MMN and P3a elicited by small deviant frequencies in harmonic tones distinguished older adult musicians from older nonmusicians. Prior to attentional focus, older musicians discriminated small changes of pitch in harmonic tone complexes faster (i.e., shorter MMN latencies) than older untrained listeners indicating superior sensory memory-based comparison processes and supporting the premise that early and intensive music training may affect central auditory processing throughout life. Consistent with the literature, for both groups of older listeners, as the magnitude of stimulus deviance increased, MMN latency decreased and amplitude increased [5, 9, 24]. As predicted, MMN amplitude did not significantly differ between the two older groups suggesting that older musicians discriminated changes in pitch faster than nonmusicians without an increase in neural activation.

P3a latency and amplitude patterns for the older adult groups were similar to the MMN, but not identical. Older musicians switched attention faster (i.e., shorter P3a latency) to acoustic changes in music without an increase in neural activation. This difference between trained and untrained listeners speaks to the musician's extensive auditory experience and exposure to multiple, concurrent, and sequential music stimuli. Consequently, even a minor change in pitch

elicited a swift shift of attention toward the distraction. Like the MMN latency, all older listeners responded faster to the larger frequency deviance (earlier P3a latency).

To further elucidate the interaction between aging and musicianship, MMN and P3a data were compared among four groups: young musicians (YM), young nonmusicians (YNM), older musicians (OM), and older nonmusicians (ONM). For the larger frequency deviance (i.e., 6% or 22 Hz), musicianship appears to have the advantage; that is, all the musicians (YM and OM) reacted to the pitch change faster with comparable neural effort than the nonmusicians (YNM and ONM). Since attentional modulation of the MMN is possible [47], this effect of musical expertise on the MMN could in part be due to a difference in the attentional demand of the stimuli since tone stimuli might be inherently more "interesting" for musicians, or musicians might be more used to directing their attention to the sounds.

The YNM had the strongest neural activation (largest MMN amplitude), while the ONM had the smallest. Findings support previous literature suggesting that preattentive change-detection (MMN) may be affected more by stimulus contrast and presentation rate than age alone [19, 26, 34]. However, despite the fact that the older adults had normal pure-tone thresholds up to 3 kHz and that low-frequency stimulus contrasts were used to minimize the confounding effects of age-related high-frequency hearing loss, the possibility remains that physiological discrimination of the 1.5% pitch deviance (6 Hz difference) may have been compromised by aging effects.

P3a latency patterns were similar to those of the MMN. Like the MMN response to the larger harmonic tone deviant, musicianship appears to have prevailed. All musicians (YM and OM) compared to the untrained listeners (YNM and ONM) demonstrated faster preattentive registration of a deviant harmonic stimulus outside attentional focus and shifted their attention toward the distraction. This reflexive auditory-neural response may be a residual adaptation related to music training and experience during childhood. Consistent with the MMN response to the smaller 1.5% pitch deviance, older adults were slower to register this distraction outside of their attentional focus than the younger adults and neural effort was not a distinguishing factor. In summary, MMN and P3a components appear to be more influenced by musicianship rather than aging. Further, it appears that early and intensive music training may alter aging effects on central auditory processing throughout life.

5. Conclusions

Within the neuroscience of music, this study was an initial foray into the investigation of the effects of aging and musicianship on the auditory neural system. Youth and musicianship appear to be an advantageous combination for efficient and enhanced preattentive auditory neural processing. In terms of physiological detection of sound, P1-N1-P2 appears to be more sensitive to effects of aging rather than musicianship. All older listeners, regardless of musicianship, demonstrated faster auditory arousal to the presence of

acoustic energy than all younger listeners suggesting an adaptive increase in neural recruitment; yet once arousal was initiated, neural conduction of acoustic energy was slower for the older adults than younger. In terms of automatic acoustic discrimination, MMN and P3a appear to be more affected by musicianship than aging. All musicians, regardless of age group, demonstrated a pattern of enhanced auditory sensory-memory-based comparison processes for harmonic tone stimuli and exceptional sensitivity to and involuntary distraction by an acoustic change of music. Findings suggest that musicianship has beneficial neurophysiological consequences on central auditory processing throughout life, and, further, some of these neurophysiological effects may be independent of age-related changes.

The lifelong ability to adapt to environmental demands and sensory stimulation is based on the dynamic capacity of the human brain to modify and alter its structure and function. Formal music training has been shown to facilitate and enhance encoding of the acoustic signal, shape subcortical, and early cortical stages of acoustic perception and may retard age-related neural changes and facilitate adaptive neural function. The benefits of music training are socially, clinically, and educationally relevant. Music training and practice may be used as an educational tool or a rehabilitative strategy to facilitate neurophysiological processing of sound. The possibilities for off-setting age-related physiological changes of the auditory neural system through music training have far-reaching effects not only for the field of neuroscience and music education, but also for gerontology, speech-language pathology, and hearing science. Regardless of the remarkable technological advances in hearing devices such as programmable digital hearing aids and cochlear implants, none can duplicate our original auditory system. Consequently, it is essential to investigate other means by which our auditory abilities may be enhanced and protected.

This was an introductory study to explore possibilities that early music training may influence auditory processing in later life and to assess the methodological challenges faced when working with an older population. Implications from this study have been interpreted with caution and generalizations were kept to a minimum due to the small sample of older musicians and the intrasubject variability that occurs when working with older adults. Future recommendations for studies of older adults include recruiting larger numbers to offset individual variability and allowing for frequent breaks during the experiment to offset participant fatigue. To further elucidate the effects of aging and musicianship, it would be beneficial to include higher-level listening tasks using more complex auditory stimuli such as speech stimuli and to also include psychoacoustic measures of auditory perception. Finally, while the information presented suggests that formal music training and extensive auditory sensory exposure facilitate neurophysiological responses throughout life, the influence of genetic factors and inherent musical ability cannot be dismissed. It is yet unknown whether musical abilities and cortical structural differences are due to learning or whether these differences reflect innate abilities and capacities that are advanced by early exposure to music.

Abbreviations

ANOVA:	Analysis of variance
CAEPs:	Cortical auditory evoked potentials
DLF:	Difference limen for frequency
EEG:	Electroencephalographic
F0:	Fundamental frequency
fMRI:	Functional magnetic resonance imaging
HMD1.5:	Stimulus with 1.5% difference from the standard tone F0
HMD6:	Stimulus with 6% difference from the standard tone F0
ISI:	Interstimulus interval
MEG:	Magnetoencephalographic
MMN:	Mismatch negativity
ONM:	Older nonmusicians
OM:	Older musicians
TDT:	Tucker-Davis Technologies
USF:	University of South Florida
WIN:	Words-in-Noise
YNM:	Younger nonmusicians
YM:	Younger musician.

Disclosure

Portions of the work described here were presented as follows: (1) D. Nikjeh and J. Lister (2012): Preattentive CAEPs: Exploring Effects of Musicianship and Aging in American Speech-Language-Hearing Association conference, Atlanta, GA; (2) J. Lister, J. Edwards, J. O'Brien, D. Nikjeh, and V. Williams (2011): Effects of Auditory Training and Music Background: A CAEP Study of Older Adults in Aging and Speech Communication—An International and Interdisciplinary Conference, Bloomington, IN; and (3) D. Nikjeh, J. Lister, and K. Weinbaum (2010): Aging Musicians: Influence of Early Music Training on Cortical Auditory Evoked Potentials in AHEAD III (Assessment of Hearing in the Elderly: Aging and Degeneration-Integration through Immediate Intervention) Adult Hearing Screening conference, Cernobbio, Italy.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

References

- [1] G. Schlaug, "The brain of musicians: a model for functional and structural adaptation," *Annals of the New York Academy of Sciences*, vol. 930, pp. 281–299, 2001.
- [2] C. Gaser and G. Schlaug, "Brain structures differ between musicians and non-musicians," *The Journal of Neuroscience*, vol. 23, no. 27, pp. 9240–9245, 2003.
- [3] T. Elbert, V. Candia, E. Altenmüller et al., "Alteration of digital representations in somatosensory cortex in focal hand dystonia," *NeuroReport*, vol. 9, no. 16, pp. 3571–3575, 1998.
- [4] C. Pantev, A. Engelien, V. Candia, and T. Elbert, "Representational cortex in musicians: plastic alterations in response to

- musical practice,” *Annals of the New York Academy of Sciences*, vol. 930, pp. 300–314, 2001.
- [5] E. Brattico, R. Näätänen, and M. Tervaniemi, “Context effects on pitch perception in musicians and nonmusicians: evidence from event-related-potential recordings,” *Music Perception*, vol. 19, no. 2, pp. 199–222, 2001.
 - [6] T. Fujioka, L. J. Trainor, B. Ross, R. Kakigi, and C. Pantev, “Musical training enhances automatic encoding of melodic contour and interval structure,” *Journal of Cognitive Neuroscience*, vol. 16, no. 6, pp. 1010–1021, 2004.
 - [7] S. Koelsch, B.-H. Schmidt, and J. Kansok, “Effects of musical expertise on the early right anterior negativity: an event-related brain potential study,” *Psychophysiology*, vol. 39, no. 5, pp. 657–663, 2002.
 - [8] S. Koelsch, E. Schröger, and M. Tervaniemi, “Superior pre-attentive auditory processing in musicians,” *NeuroReport*, vol. 10, no. 6, pp. 1309–1313, 1999.
 - [9] D. A. Nikjeh, J. J. Lister, and S. A. Frisch, “Hearing of note: an electrophysiologic and psychoacoustic comparison of pitch discrimination between vocal and instrumental musicians,” *Psychophysiology*, vol. 45, no. 6, pp. 994–1007, 2008.
 - [10] D. A. Nikjeh, J. J. Lister, and S. A. Frisch, “Preattentive cortical-evoked responses to pure tones, harmonic tones, and speech: influence of music training,” *Ear & Hearing*, vol. 30, no. 4, pp. 432–446, 2009.
 - [11] D. A. Nikjeh, J. J. Lister, and S. A. Frisch, “The relationship between pitch discrimination and vocal production: comparison of vocal and instrumental musicians,” *Journal of the Acoustical Society of America*, vol. 125, no. 1, pp. 328–338, 2009.
 - [12] M. Tervaniemi, V. Just, S. Koelsch, A. Widmann, and E. Schröger, “Pitch discrimination accuracy in musicians vs non-musicians: an event-related potential and behavioral study,” *Experimental Brain Research*, vol. 161, no. 1, pp. 1–10, 2005.
 - [13] D. Schön, C. Magne, and M. Besson, “The music of speech: music training facilitates pitch processing in both music and language,” *Psychophysiology*, vol. 41, no. 3, pp. 341–349, 2004.
 - [14] P. C. M. Wong, E. Skoe, N. M. Russo, T. Dees, and N. Kraus, “Musical experience shapes human brainstem encoding of linguistic pitch patterns,” *Nature Neuroscience*, vol. 10, no. 4, pp. 420–422, 2007.
 - [15] C. Lappe, O. Steinsträter, and C. Pantev, “Rhythmic and melodic deviations in musical sequences recruit different cortical areas for mismatch detection,” *Frontiers in Human Neuroscience*, vol. 7, article 260, 2013.
 - [16] C. Pantev, L. E. Roberts, M. Schulz, A. Engelen, and B. Ross, “Timbre-specific enhancement of auditory cortical representations in musicians,” *NeuroReport*, vol. 12, no. 1, pp. 169–174, 2001.
 - [17] C. Lappe, S. C. Herholz, L. J. Trainor, and C. Pantev, “Cortical plasticity induced by short-term unimodal and multimodal musical training,” *The Journal of Neuroscience*, vol. 28, no. 39, pp. 9632–9639, 2008.
 - [18] K. Tremblay and B. Ross, “Effects of age and age-related hearing loss on the brain,” *Journal of Communication Disorders*, vol. 40, no. 4, pp. 305–312, 2007.
 - [19] S. Bertoli, J. Smurzynski, and R. Probst, “Effects of age, age-related hearing loss, and contralateral cafeteria noise on the discrimination of small frequency changes: psychoacoustic and electrophysiological measures,” *Journal of the Association for Research in Otolaryngology*, vol. 6, no. 3, pp. 207–222, 2005.
 - [20] J. J. Lister, N. D. Maxfield, and G. J. Pitt, “Cortical evoked response to gaps in noise: within-channel and across-channel conditions,” *Ear and Hearing*, vol. 28, no. 6, pp. 862–878, 2007.
 - [21] J. J. Lister, N. D. Maxfield, G. J. Pitt, and V. B. Gonzalez, “Auditory evoked response to gaps in noise: older adults,” *International Journal of Audiology*, vol. 50, no. 4, pp. 211–225, 2011.
 - [22] A. Parbery-Clark, D. L. Strait, S. Anderson, E. Hittner, and N. Kraus, “Musical experience and the aging auditory system: implications for cognitive abilities and hearing speech in noise,” *PLoS ONE*, vol. 6, no. 5, Article ID e18082, 2011.
 - [23] B. R. Zendel and C. Alain, “Musicians experience less age-related decline in central auditory processing,” *Psychology and Aging*, vol. 27, no. 2, pp. 410–417, 2012.
 - [24] N. Novitski, M. Tervaniemi, M. Huottilainen, and R. Näätänen, “Frequency discrimination at different frequency levels as indexed by electrophysiological and behavioral measures,” *Cognitive Brain Research*, vol. 20, no. 1, pp. 26–36, 2004.
 - [25] K. L. Tremblay, M. Piskosz, and P. Souza, “Aging alters the neural representation of speech cues,” *NeuroReport*, vol. 13, no. 15, pp. 1865–1870, 2002.
 - [26] E. Pekkonen, T. Rinne, K. Reinikainen, T. Kujala, K. Alho, and R. Näätänen, “Aging effects on auditory processing: an event-related potential study,” *Experimental Aging Research*, vol. 22, no. 2, pp. 171–184, 1996.
 - [27] J. Rimmele, E. Sussman, C. Keitel, T. Jacobsen, and E. Schröger, “Electrophysiological evidence for age effects on sensory memory processing of tonal patterns,” *Psychology and Aging*, vol. 27, no. 2, pp. 384–398, 2012.
 - [28] C. Alain and K. Tremblay, “The role of event-related brain potentials in assessing central auditory processing,” *Journal of the American Academy of Audiology*, vol. 18, no. 7, pp. 573–589, 2007.
 - [29] H. Menning, L. Roberts, and C. Pantev, “Plastic changes in the auditory cortex induced by intensive frequency discrimination training,” *NeuroReport*, vol. 11, no. 4, pp. 817–822, 2000.
 - [30] R. Näätänen, T. Jacobsen, and I. Winkler, “Memory-based or afferent processes in mismatch negativity (MMN): a review of the evidence,” *Psychophysiology*, vol. 42, no. 1, pp. 25–32, 2005.
 - [31] J. M. K. Nousak, D. Deacon, W. Ritter, and H. G. Vaughan Jr., “Storage of information in transient auditory memory,” *Cognitive Brain Research*, vol. 4, no. 4, pp. 305–317, 1996.
 - [32] C. Escera, K. Alho, I. Winkler, and R. Näätänen, “Neural mechanisms of involuntary attention to acoustic novelty and change,” *Journal of Cognitive Neuroscience*, vol. 10, no. 5, pp. 590–604, 1998.
 - [33] E. Yago, M. J. Corral, and C. Escera, “Activation of brain mechanisms of attention switching as a function of auditory frequency change,” *NeuroReport*, vol. 12, no. 18, pp. 4093–4097, 2001.
 - [34] H. Gaeta, D. Friedman, W. Ritter, and J. Cheng, “An event-related potential study of age-related changes in sensitivity to stimulus deviance,” *Neurobiology of Aging*, vol. 19, no. 5, pp. 447–459, 1998.
 - [35] Z. A. Gaál, R. Csuhaj, and M. Molnár, “Age-dependent changes of auditory evoked potentials—effect of task difficulty,” *Biological Psychology*, vol. 76, no. 3, pp. 196–208, 2007.
 - [36] C. McNamara, *Ways to Look at Training and Development Processes: Informal/Formal and Self-Directed/Other-Directed*, Free Management Library, 2005, <http://www.managementhelp.org/trng.dev/ways/ways.htm>.
 - [37] R. H. Wilson and C. A. Burks, “Use of 35 words for evaluation of hearing loss in signal-to-babble ratio: a clinic protocol,” *Journal of Rehabilitation Research and Development*, vol. 42, no. 6, pp. 839–852, 2005.

- [38] R. H. Wilson, "Development of a speech-in-multitalker-babble paradigm to assess word-recognition performance," *Journal of the American Academy of Audiology*, vol. 14, no. 9, pp. 453–470, 2003.
- [39] M. Hirano, *Clinical Examination of Voice*, Springer, New York, NY, USA, 1981.
- [40] E. Pekkonen, T. Rinne, and R. Näätänen, "Variability and replicability of the mismatch negativity," *Electroencephalography and Clinical Neurophysiology*, vol. 96, no. 6, pp. 546–554, 1995.
- [41] J. Sinkkonen and M. Tervaniemi, "Towards optimal recording and analysis of the mismatch negativity," *Audiology and Neuro-Otology*, vol. 5, no. 3-4, pp. 235–246, 2000.
- [42] R. Näätänen, "The mismatch negativity: a powerful tool for cognitive neuroscience," *Ear & Hearing*, vol. 16, no. 1, pp. 6–18, 1995.
- [43] M. Tervaniemi, S. Kruck, W. de Baene, E. Schröger, K. Alter, and A. D. Friederici, "Top-down modulation of auditory processing: effects of sound context, musical expertise and attentional focus," *European Journal of Neuroscience*, vol. 30, no. 8, pp. 1636–1642, 2009.
- [44] T. W. Picton, S. Bentin, P. Berg et al., "Guidelines for using human event-related potentials to study cognition: recording standards and publication criteria," *Psychophysiology*, vol. 37, no. 2, pp. 127–152, 2000.
- [45] M. H. Giard and F. Peronnet, "Auditory-visual integration during multimodal object recognition in humans: a behavioral and electrophysiological study," *Journal of Cognitive Neuroscience*, vol. 11, no. 5, pp. 473–490, 1999.
- [46] R. Čeponiene, M. Westerfield, M. Torki, and J. Townsend, "Modality-specificity of sensory aging in vision and audition: evidence from event-related potentials," *Brain Research*, vol. 1215, pp. 53–68, 2008.
- [47] R. Auksztulewicz and K. Friston, "Attentional enhancement of auditory mismatch responses: a DCM/MEG study," *Cerebral Cortex*, 2015.

Research Article

Inferior Frontal Gyrus Activation Underlies the Perception of Emotions, While Precuneus Activation Underlies the Feeling of Emotions during Music Listening

Ken-ichi Tabei^{1,2}

¹Department of Dementia Prevention and Therapeutics, Graduate School of Medicine, Mie University, Tsu 514-8507, Japan

²Department of Neurology, Graduate School of Medicine, Mie University, Tsu 514-8507, Japan

Correspondence should be addressed to Ken-ichi Tabei; kenichi.tabei@gmail.com

Received 21 April 2015; Revised 15 June 2015; Accepted 16 June 2015

Academic Editor: Enzo Emanuele

Copyright © 2015 Ken-ichi Tabei. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

While music triggers many physiological and psychological reactions, the underlying neural basis of perceived and experienced emotions during music listening remains poorly understood. Therefore, using functional magnetic resonance imaging (fMRI), I conducted a comparative study of the different brain areas involved in perceiving and feeling emotions during music listening. I measured fMRI signals while participants assessed the emotional expression of music (perceived emotion) and their emotional responses to music (felt emotion). I found that cortical areas including the prefrontal, auditory, cingulate, and posterior parietal cortices were consistently activated by the perceived and felt emotional tasks. Moreover, activity in the inferior frontal gyrus increased more during the perceived emotion task than during a passive listening task. In addition, the precuneus showed greater activity during the felt emotion task than during a passive listening task. The findings reveal that the bilateral inferior frontal gyri and the precuneus are important areas for the perception of the emotional content of music as well as for the emotional response evoked in the listener. Furthermore, I propose that the precuneus, a brain region associated with self-representation, might be involved in assessing emotional responses.

1. Introduction

Music has strong emotional effects on listeners. When investigating the effect of music on emotions, it is necessary to consider the felt emotion separately from the perceived emotion [1, 2]. When we listen to music, various emotions are induced in our mind. On the other hand, we can objectively appreciate the emotions expressed in the music independently of our own emotions. The former is referred to as felt emotion, which is an emotional response evoked by music in listeners. The latter is referred to as perceived emotion, which is an emotional quality detected by the listeners of music [1, 2]. Hunter and Schellenberg [3] argued about perceived emotion and felt emotion; when a listener is in a happy mood, it is possible for them to maintain that happy feeling even when they are listening to sad music, which they perceive as being sad. However, if the music is associated with an unpleasant past experience, a disagreeable feeling could be evoked in the

listener even when they recognize the music as expressing a happy feeling. Together with previous findings, it seems that feeling and perception ratings are correlated, although perceived emotion tends to be stronger than felt emotion [4–7]. Therefore, they advocated that it was necessary to separate these two types of loci.

As for the relationship between these two types of perception, past studies using psychological measurements have reported a correlation between the perceived emotion and felt emotion [4–7]. However, compared to the perceived emotion, the felt emotion includes various factors, such as recalling memories associated with the music [7]. Therefore, when the mean values were calculated from the listeners' assessments of the emotional expression of music and their assessments of their own emotional response, the absolute value of the rating scales of the listeners' emotional response tended to be lower than that of the emotional expression of the music. This finding suggests that while listeners can easily perceive

the emotional expression intended by the composer, listeners may not necessarily experience the emotional response that the composer intended to convey [3, 7].

Past neuropsychological studies of brain-damaged patients have suggested that the perceived emotion and felt emotion while listening to music may differ. Satoh et al. [8] reported a patient with damage to the parietal lobe in his right cerebral hemisphere who had selective impairment of felt emotion while listening to music; however, the patient's intellectual function, memory, constructional ability, and perception of the emotional expression of the music remained normal. Similar impairments have also been reported for a patient with damage to the temporal and parietal lobes in the right hemisphere [9] and a patient with a lesion centering on the insular cortex in the left hemisphere, which widely included the frontal lobe and amygdala [10]. On the other hand, there are also patients who show the opposite symptoms [11, 12]. Matthews et al. [11] reported a case of neurodegenerative disease with auditory agnosia in which the auditory agnosia was accompanied by impaired perception of the emotional expression of music and musical components such as pitch, while the emotional response when listening to music was maintained.

Thus, past psychological and brain injury studies have demonstrated that the perceived emotion and felt emotion while listening to music could be separable loci, which suggests that they may involve different localizations in the brain.

Recently, there have been brain function imaging studies that have attempted to visualize human brain activity while listening to music. Several studies pertaining to the perceived emotion [13–15] and felt emotion [16–20] while listening to music have been conducted. These studies have reported that when participants are assessing the emotional expression of music, activity occurs in the medial frontal gyrus, superior frontal gyrus, middle frontal gyrus, and cingulate gyrus; and when participants are assessing their emotion felt to music, activity occurs in the ventral striatum, midbrain, amygdala, orbitofrontal cortex, and ventral prefrontal area.

The purpose of these studies was to determine the areas of the brain that are involved in cognitive processing by assessing perceived emotion alone or felt emotion alone while listening to music. However, as seen in the results of psychological and brain injury studies, it is obvious that there is some interaction between the cognitive processes of emotion felt to and emotional expression of music, which suggests that there are both common and independent areas in the brain related to the processing of these two emotional loci. Therefore, in order to clarify the whole process, it is essential to compare directly both the emotion felt and emotional expression and to investigate the brain localizations related to these two emotional loci. The objectives of this study were as follows: (1) to use functional magnetic resonance imaging (fMRI) to measure brain activity while a participant is assessing their emotion felt and the emotional expression evoked by the music and (2) to compare the areas of brain activity involved in these respective assessments. Many common brain mechanisms exist for the assessment of emotional response and of emotional expression while listening to

music. However, the locus differs between the recognition of one's own emotion (felt emotion) and the understanding of what the music is expressing (perceived emotion). Therefore, it is predicted that the information processing of the two respective loci will involve independent brain localizations.

2. Materials and Methods

2.1. Participants. Seventeen healthy right-handed undergraduate students majoring in music (10 women, 7 men; mean age 21.4 years, SD = 2.0) participated in this study. Participants provided written informed consent before the experiment in accordance with the Declaration of Helsinki. The study was approved by the Ethics Committee of Nihon University.

2.2. Stimulus Selection. Four hundred and fifty-two healthy volunteers (311 men and 141 women), between 18 and 29 years of age (mean age 19.3 years, SD = 1.4), performed one of two rating tasks. None of the volunteers participated in the subsequent fMRI study. A set of 56 musical stimuli (16 happy, 16 sad, 16 scary, and 16 peaceful stimuli from Viellard et al. [21]) were copied onto two CDs in a random order, with 6 s of silence between excerpts. Two groups of listeners were defined based on the type of ratings they were requested to provide. Each excerpt was presented only once. In the felt emotion task, 232 listeners judged to what extent they experienced each of the four emotions labeled as happy, sad, scary, and peaceful. In the perceived emotion task, 220 listeners judged to what extent they perceived each of the four emotions labeled as happy, sad, scary, and peaceful in each excerpt. For each labeled emotion, listeners gave a rating on a 10-point scale from 0 (*absent*) to 9 (*present*). The listeners had been previously informed that they had to provide a rating for the four emotion labels. I calculated the average scale of each stimulus. I used only happy and sad stimuli for the fMRI experiment because it was easy for participants to differentiate happy and sad stimuli. As a result, 12 happy (g01, g03, g04, g07, g10, and g13) and sad (t04, t06, t08, t10, t13, and t14) [21, APPENDIX 2] musical stimuli were selected for the fMRI experiment from the six top average stimuli from the perceived and felt emotion tasks. The happy stimuli were written in a major mode at an average tempo of 137 metronome markings (MM; range: 92–196) [21]. The sad stimuli were written in a minor mode at an average slow tempo of 46 MM (range: 40–60) [21]. The average stimuli duration was 12.4 s (range: 9.2–16.4 s) [21].

2.3. Tasks. The stimuli were presented through stereo headphones and the instructions were presented using a projector in the fMRI scanner. The participants performed three tasks. In the felt emotion task, they judged whether they felt happy, sad, or neutral emotions and pressed the button assigned to each emotional response. I asked participants "Please indicate which of the three emotions you experience in each excerpt." In the perceived emotion task, they judged whether the stimulus expressed happy, sad, or neutral feelings and pressed the button assigned to each emotional quality. I asked participants "Please indicate which of the three

emotions you recognize in each excerpt.” The participants performed the three-forced-choice judgment in each task. In the passive listening task, the participants listened to the stimulus and then pressed a button. A block design was used. Each task condition lasted for 60 s with instruction slides followed by a baseline period (32 s). In each task block, three musical stimuli were presented in a random order every 20 s. Participants performed each task six times in a recording session.

2.4. fMRI Measurements. The MRI scans were obtained using a 1.5-Tesla MR scanner (Siemens Symphony). Functional images were obtained using a T2*-weighted gradient-echo planar imaging sequence (40 horizontal slices, repetition time (TR) = 4,000 ms, echo time (TE) = 50 ms, slice thickness = 3 mm, gapless, and field of view (FOV) = 192 mm, 64 × 64 matrix). Additionally, a T1-weighted anatomical image was obtained for each participant (TR = 2,200 ms, TE = 3.93 ms, FA = 15°, TI = 1,100 ms, 1 mm³ voxel, and FOV = 256 mm).

2.5. fMRI Data Analysis. Data analysis was performed using SPM5 (Wellcome Department of Imaging Neuroscience, London, UK). The functional images were realigned to the first image to correct for movement-related effects, coregistered to the anatomical image, normalized to the Montreal Neurological Institute brain template, and spatially smoothed with an isotropic Gaussian kernel (full width at half maximum = 8 mm). I conducted voxelwise statistical analyses based on the general linear model. For the statistical model, a block design was modeled using the canonical hemodynamic response function and the temporal derivative, and low-frequency drifts were removed using a high-pass filter (128 s). For each participant, I computed contrasts for “perceived emotion task > baseline,” “felt emotion task > baseline,” and “passive listening task > baseline.” In addition, I also computed contrasts for “perceived emotion task > passive listening task,” “perceived emotion task > felt emotion task,” “felt emotion task > passive listening task,” and “felt emotion task > perceived emotion task.” In order to exclude the possibility of negative activations, I applied an inclusive mask using the contrast image of “perceived emotion task > baseline” (in the former two contrasts) and “felt emotion task > baseline” (in the latter two contrasts). A random-effects model was used for the group analysis (a voxel level after correction for false discovery rate $p < 0.05$ or cluster threshold $p < 0.05$ (corrected) and a voxel threshold $p < 0.001$ (uncorrected)).

3. Results

3.1. Perceived Emotion Is Stronger Than Felt Emotion across Participants. Each participant’s judgment in the perceived emotion and felt emotion tasks was calculated as the congruency ratio of selected adjective to music category (i.e., happy to happy-stimuli and sad to sad-stimuli; Figure 1). In the perceived emotion task, the mean of the congruency ratio was 93.5% and 94.1% for happy to happy-stimuli and sad to sad-stimuli, respectively. In the felt emotion task, the mean of the



FIGURE 1: Congruency ratio of selected adjective to music categories. The bars represent standard deviation.

congruency ratio was 78.4% and 72.5% for happy to happy-stimuli and sad to sad-stimuli, respectively. The effect of tasks on perceived versus felt emotion was tested by performing two-way analysis of variance with the congruency ratio of selected adjective to musical category for the task condition (perceived versus felt) and musical category (happy versus sad) as within-participants factors. A significant main effect of task ($F(1, 16) = 7.64, p < 0.05$) was observed, with a higher congruency ratio in the perceived emotion condition than in the felt emotion condition. No significant main effect of musical category or interaction between task condition and musical category was found.

3.2. Perceived Emotions Activate the Bilateral Inferior Frontal Gyri, While Felt Emotions Activate the Precuneus. Using baseline subtraction analysis, I found that the perceived emotion task was associated with increased blood-oxygen-level dependent (BOLD) signals in areas of the frontal cortex (the inferior, middle, and superior frontal gyri and medial wall), auditory areas, the posterior parietal cortex, precuneus, parahippocampal gyrus, cingulate gyrus, right lentiform nucleus, left thalamus, and cerebellum (Figure 2(a)). The felt emotion task was associated with increased BOLD signals in areas of the frontal cortex (the inferior, middle, and superior frontal gyri and medial wall), auditory areas, the posterior parietal cortex, precuneus, cingulate gyrus, right lentiform nucleus, left thalamus, and cerebellum (Figure 2(b)). Meanwhile, the passive listening task was associated with increased BOLD signals in areas of the frontal cortex (inferior and middle frontal gyri), auditory areas, the posterior parietal cortex, and the cerebellum (Figure 2(c)).

When I performed subtraction analysis with the passive listening task, the perceived emotion task was associated with increased BOLD signals in the bilateral inferior frontal gyri (Table 1(a); Figure 3(a)) and the felt emotion task was associated with increased BOLD signals in the precuneus (Table 1(b); Figure 3(b)). On the other hand, a direct comparison of the passive listening and the perceived emotion tasks and of the passive listening and the felt emotion tasks did not show significant activation. A direct comparison between the

TABLE 1: Cortical areas where the blood-oxygen-level dependent signals increased with perceived emotion > passive listening (a) and felt emotion > passive listening (b); cluster threshold $p < 0.05$ (corrected).

(a)							
Contrast	L/R	Area	Brodmann area (BA)	Z value	Talairach coordinates (mm)		
					X	Y	Z
Perceived > passive	Right	Inferior frontal gyrus	BA 45	3.63	53	28	8
				3.45	55	26	17
				3.78	-53	27	-5
	Left	Inferior frontal gyrus	BA 47	3.49	-55	17	-3
				3.39	-46	31	0
(b)							
Contrast	L/R	Area	Brodmann area (BA)	Z value	Talairach coordinates (mm)		
					X	Y	Z
Felt > passive	Right	Precuneus	BA 7	3.69	2	-75	46
	Left	Precuneus	BA 7	3.87	-8	-66	35
				3.32	-4	-70	44

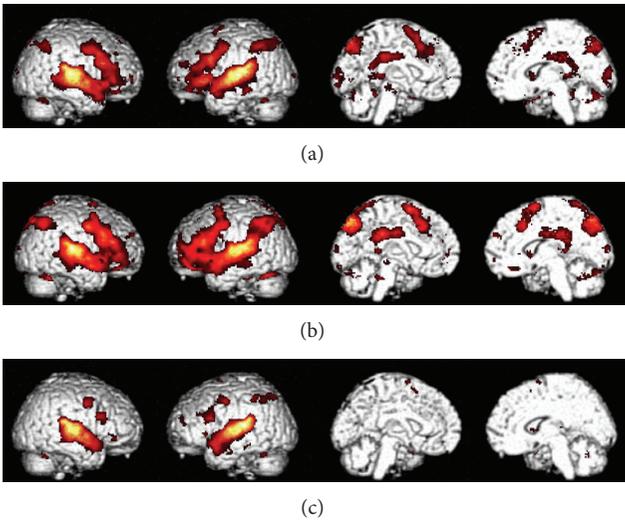


FIGURE 2: Cortical areas where the blood-oxygen-level dependent signals increased with the perceived emotion (a), felt emotion (b), and passive listening (c) tasks; voxel level after correction for false discovery rate ($p < 0.05$).

perceived and felt emotion tasks also did not show significant activation.

4. Discussion

In this study, brain activity was measured using fMRI while the participants were assessing the emotional expression of music and while they were assessing the emotional response they felt while listening. I compared the active brain areas that were involved in each of the assessments. I found that the study participants assessed sad stimuli as more “sad” and

happy stimuli as more “happy” in the perceived emotion task than in the felt emotion task. Similar to previous studies [4–7], I confirmed that perceived emotion was assessed more reliably by the participants than emotional responses.

A comparison of brain activity in each of the tasks showed that the precuneus region was activated more strongly in the felt emotion task than in the passive listening task. In addition, greater activity in the bilateral inferior frontal gyri was observed during the perceived emotion task than during the passive listening task. The precuneus region is known to be the area responsible for “self-representation” and is activated during tasks involving self-referential judgments and tasks consisting of judgments from a first-person perspective [22]. In addition, previous work where only the listeners’ emotional responses were assessed found that activity in the precuneus region was correlated with the emotional response [16]. Therefore, this study suggests that the precuneus region is concerned with evaluating one’s own cognitive emotional changes during music listening.

Activity in the inferior frontal gyri increased during the perceived emotion task. Although there have been no previous reports on inferior frontal gyri activation during emotional assessment, previous studies have reported activity in the dorsolateral prefrontal cortex (Brodmann area 9) during the assessment of emotion expressions [23]. One reason for the paucity of reports may stem from the fact that previous studies have focused on the brain lateralization of the perception of emotional expressions of happy and sad music. However, the dorsolateral prefrontal cortex has been reported to be active during the judgment of tonality [24] and the detection of deviations from harmonic progressions [25], suggesting that this area might be involved in the processing of tonality in musical structures. The musical stimuli used in the current study also consisted of music involving major and minor tonalities and may account for the inferior frontal gyri activation.

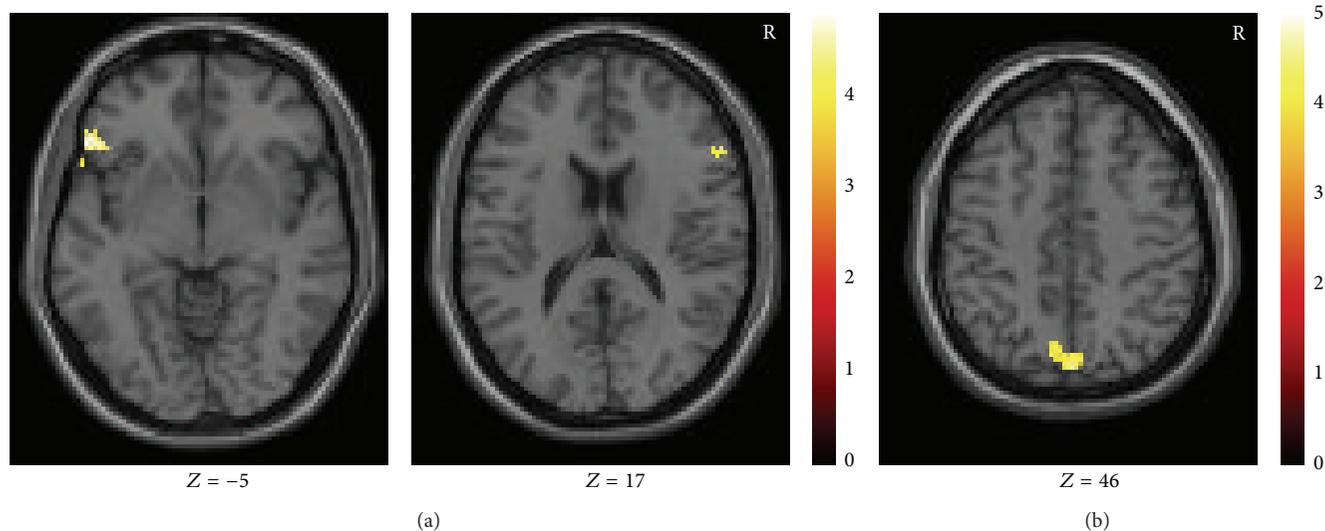


FIGURE 3: Cortical areas where the blood-oxygen-level dependent signals increased with perceived emotion > passive listening (a) and felt emotion > passive listening (b); cluster threshold $p < 0.05$ (corrected).

5. Conclusion

In summary, the findings suggest that the bilateral inferior frontal gyri might be involved in assessing the emotional expression of music, while the precuneus processes the emotional responses evoked in the listener. Moreover, the fact that the precuneus region is likely responsible for self-representation implicates this brain area in the assessment of the emotional response to music.

Conflict of Interests

The author declares that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

The author would like to thank Nobuo Usui, Izuru Nose, and Masato Taira, whose comments and suggestions were innumerable valuable. Finally, the author would like to express their gratitude to JSPS for their financial support.

References

- [1] A. Gabrielsson, "Emotion perceived and emotion felt: same or different?" *Musicae Scientiae*, vol. 5, pp. 123–147, 2002.
- [2] E. Schubert, "Emotion felt by the listener and expressed by the music: literature review and theoretical perspectives," *Frontiers in Psychology*, vol. 4, article 837, 2013.
- [3] P. G. Hunter and E. G. Schellenberg, "Music and emotion," in *Music Perception*, M. Riess Jones, R. R. Fay, and A. Popper, Eds., pp. 129–164, Springer, New York, NY, USA, 2010.
- [4] P. Evans and E. Schubert, "Relationships between expressed and felt emotions in music," *Musicae Scientiae*, vol. 12, no. 1, pp. 75–99, 2008.
- [5] K. Kallinen and N. Ravaja, "Emotion perceived and emotion felt: same and different," *Musicae Scientiae*, vol. 10, no. 2, pp. 191–213, 2006.
- [6] E. Schubert, "Locus of emotion: the effect of task order and age on emotion perceived and emotion felt in response to music," *Journal of Music Therapy*, vol. 44, no. 4, pp. 344–368, 2007.
- [7] E. Schubert, "The influence of emotion, locus of emotion and familiarity upon preference in music," *Psychology of Music*, vol. 35, no. 3, pp. 499–515, 2007.
- [8] M. Satoh, T. Nakase, K. Nagata, and H. Tomimoto, "Musical anhedonia: selective loss of emotional experience in listening to music," *Neurocase*, vol. 17, no. 5, pp. 410–417, 2011.
- [9] M. Mazzoni, P. Moretti, L. Pardossi, M. Vista, A. Muratorio, and M. Puglioli, "A case of music imperception," *Journal of Neurology Neurosurgery and Psychiatry*, vol. 56, no. 3, article 322, 1993.
- [10] T. D. Griffiths, J. D. Warren, J. L. Dean, and D. Howard, "When the feeling's gone: a selective loss of musical emotion," *Journal of Neurology, Neurosurgery and Psychiatry*, vol. 75, no. 2, pp. 344–345, 2004.
- [11] B. R. Matthews, C.-C. Chang, M. de May, J. Engstrom, and B. L. Miller, "Pleasurable emotional response to music: a case of neurodegenerative generalized auditory agnosia," *Neurocase*, vol. 15, no. 3, pp. 248–259, 2009.
- [12] I. Peretz and L. Gagnon, "Dissociation between recognition and emotional judgements for melodies," *Neurocase*, vol. 5, no. 1, pp. 21–30, 1999.
- [13] A. J. Blood, R. J. Zatorre, P. Bermudez, and A. C. Evans, "Emotional responses to pleasant and unpleasant music correlate with activity in paralimbic brain regions," *Nature Neuroscience*, vol. 2, no. 4, pp. 382–387, 1999.
- [14] S. Brown, M. J. Martinez, and L. M. Parsons, "Passive music listening spontaneously engages limbic and paralimbic systems," *NeuroReport*, vol. 15, no. 13, pp. 2033–2037, 2004.
- [15] V. Menon and D. J. Levitin, "The rewards of music listening: response and physiological connectivity of the mesolimbic system," *NeuroImage*, vol. 28, no. 1, pp. 175–184, 2005.

- [16] A. J. Blood and R. J. Zatorre, "Intensely pleasurable responses to music correlate with activity in brain regions implicated in reward and emotion," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 98, no. 20, pp. 11818–11823, 2001.
- [17] S. Koelsch, T. Fritz, D. Y. V. Cramon, K. Müller, and A. D. Friederici, "Investigating emotion with music: an fMRI study," *Human Brain Mapping*, vol. 27, no. 3, pp. 239–250, 2006.
- [18] T. Baumgartner, M. Esslen, and L. Jäncke, "From emotion perception to emotion experience: emotions evoked by pictures and classical music," *International Journal of Psychophysiology*, vol. 60, no. 1, pp. 34–43, 2006.
- [19] T. Baumgartner, K. Lutz, C. F. Schmidt, and L. Jäncke, "The emotional power of music: how music enhances the feeling of affective pictures," *Brain Research*, vol. 1075, no. 1, pp. 151–164, 2006.
- [20] M. T. Mitterschiffthaler, C. H. Y. Fu, J. A. Dalton, C. M. Andrew, and S. C. R. Williams, "A functional MRI study of happy and sad affective states induced by classical music," *Human Brain Mapping*, vol. 28, no. 11, pp. 1150–1162, 2007.
- [21] S. Viellard, I. Peretz, N. Gosselin, S. Khalfa, L. Gagnon, and B. Bouchard, "Happy, sad, scary and peaceful musical excerpts for research on emotions," *Cognition and Emotion*, vol. 22, no. 4, pp. 720–752, 2008.
- [22] A. E. Cavanna and M. R. Trimble, "The precuneus: a review of its functional anatomy and behavioural correlates," *Brain*, vol. 129, no. 3, pp. 564–583, 2006.
- [23] S. Khalfa, D. Schon, J.-L. Anton, and C. Liégeois-Chauvel, "Brain regions involved in the recognition of happiness and sadness in music," *NeuroReport*, vol. 16, no. 18, pp. 1981–1984, 2005.
- [24] T. Mizuno and M. Sugishita, "Neural correlates underlying perception of tonality-related emotional contents," *NeuroReport*, vol. 18, no. 16, pp. 1651–1655, 2007.
- [25] N. Steinbeis and S. Koelsch, "Shared neural resources between music and language indicate semantic processing of musical tension-resolution patterns," *Cerebral Cortex*, vol. 18, no. 5, pp. 1169–1178, 2008.

Review Article

The Mismatch Negativity: An Indicator of Perception of Regularities in Music

Xide Yu, Tao Liu, and Dingguo Gao

Department of Psychology, Sun Yat-Sen University, No. 135 Xingang Xi Road, Guangzhou 510275, China

Correspondence should be addressed to Dingguo Gao; edsgao@mail.sysu.edu.cn

Received 24 April 2015; Revised 23 July 2015; Accepted 26 July 2015

Academic Editor: Stefan Evers

Copyright © 2015 Xide Yu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper reviews music research using Mismatch Negativity (MMN). MMN is a deviation-specific component of auditory event-related potential (ERP), which detects a deviation between a sound and an internal representation (e.g., *memory trace*). Recent studies have expanded the notion and the paradigms of MMN to higher-order music processing such as those involving short melodies, harmony chord, and music syntax. In this vein, we firstly reviewed the evolution of MMN from sound to music and then mainly compared the differences of MMN features between musicians and nonmusicians, followed by the discussion of the potential roles of the training effect and the natural exposure in MMN. Since MMN can serve as an index of neural plasticity, it thus can be widely used in clinical and other applied areas, such as detecting music preference in newborns or assessing wholeness of central auditory system of hearing illness. Finally, we pointed out some open questions and further directions. Current music perception research using MMN has mainly focused on relatively low hierarchical structure of music perception. To fully understand the neural substrates underlying processing of regularities in music, it is important and beneficial to combine MMN with other experimental paradigms such as early right-anterior negativity (ERAN).

1. Introduction

Music, as an aesthetic entity, is constituted of complex structures and regularities. The basic components of musical sound include pitch, duration, intensity, and timbre. While the high-level elements consist of beat, tempo, chord, rhythm, and harmony melody following certain rules [1, 2]. It has been acknowledged that music plays important roles in evolution and development of human cognitive functions such as intelligence [3], language [4], and memory [5] and in turn affects human social behaviors, for example, emotion management, self-identity, and interpersonal relationship [6]. More recently, growing neurology research has further reported a close relationship between music and rehabilitation/remission of some neural diseases, such as Alzheimer's disease [7–12], Parkinson's disease [13], autism [14], and traumatic brain injury [15]. Therefore, it is important and beneficial to systemically examine the neural substrates underlying music perception/appreciation.

Because of its fundamental importance on human cognition and behavior, music has been attracting increasing

attention in fields of neuroscience and neurology, opening a new hotspot of research termed as “music neuroscience” [16–18]. Previous literature on music neuroscience has primarily focused on Mismatch Negative (MMN)—a deviation-specific component of auditory event-related potential (ERP) recorded by electroencephalography (EEG)—in “music” perception. MMN is closely related to deviations between different sound features, such as pitch, timbre, location of sound source, intensity, rhythm, and abstract rule. Specifically, MMN peaks at around 100–200 ms after deviation onset with the amplitude and latency altering depending on deviation magnitude and perceptual discriminability [19].

Concerning the neural substrates of MMN, previous EEG studies [20–22] have revealed that the main source of MMN may localize at the supratemporal plane of the parietal lobes and is close to the primary and secondary auditory cortices, while the additional sources are in proximity to the inferior frontocentral lobes [23, 24]. After being found firstly in 1978 [25], MMN has been detected at a wide variety of studies from single sine tone [26, 27], alternating piano tones [28, 29], pitch discrimination [30], timbre processing [31, 32], rhythm

[33], and beat [34] to chord and melody [35]. And recently, MMN in music has greatly broken through the limitations of traditional MMN paradigms in sound perception, even though the nature of MMN in detecting deviations from standards will never change. In context-free experiments, for example, frequency deviation in sequences of pure tones, the standardized internal representation is commonly assumed as an auditory memory, induced by frequent repetition of a standard sound, and is hence termed as “*memory trace*.” In the same vein, in music this may interact with internal representation of rules, cultural preferences, and so forth. Thus, the MMN is involved in neural system that detects the violation of rule-based contexts in memory trace.

Due to the social nature of music, various music styles are different in their cultures [36, 37]. From a perspective of social culture, a large number of previous studies have demonstrated that human brain has a neural template that is derived from our musical culture, allowing for the detection of deviations of culture-based regularities by MMN [38, 39]. Besides MMN, on the other hand, ERP research has also revealed several components related with cultural familiarity in terms of music phrase boundary process, such as a P3-like component and music closure positive shift (CPS) [40, 41]. For instance, in a brain-imaging study, Nan et al. (2008) have reported that the culturally familiar, native (i.e., Western) condition elicited higher activation in the bilateral motor regions, the left planum temporale, and the right ventromedial prefrontal cortex than the nonnative condition, indicating better sensorimotor integration. In contrast, the unfamiliar, nonnative music style (i.e., Chinese) yielded higher activation in a right lateralized network, including the angular gyrus, the middle frontal gyrus, and the posterior insula, indicating increased demands on attentional control and auditory processing [42].

These ERP components and functional activities distinguish between music-related neural regularities. However, they touch little on deviation detection in memory trace (induced via frequently repeated standard musical sounds) that can be monitored by MMN. And there is an interesting question that whether or not culture-related neural regularities play a role in facilitation or interference to MMN. More studies are needed to address this issue.

Although MMN is widely detected in music research, it alone is not enough to fully understand the neural substrates underlying music perception, since MMN only reflects simple deviation from internal memory trace. The early right-anterior negativity (ERAN) is a specific response to harmonic violations, reflecting the processing of music-syntactic information. That is, ERAN involves acoustic information structured according to complex and abstract regularities [43]. Hence, combination of MMN and ERAN may provide us more rich and complete information about music perception.

Previous studies have demonstrated that both MMN and ERAN, having a feature of plasticity, are modulated by musical experience and culture surroundings [43], even in early childhood. Thus, they may facilitate cognitive treatments on individuals with amusia and other auditory-related diseases. To achieve this goal, many questions are needed to be resolved in future music neuroscience research, such as

how to use MMN and ERAN to deal with the debates of top-down (determined by the implicit or explicit knowledge of music sound) and down-top (context-driven, a human inborn nature) modulations in music and where is the “music processing sub-module” proposed by Peretz and Coltheart [44]?

Following this logical stream, we will firstly review MMN history from sound to music briefly and discuss its critical functions. Focusing on the feature of neuroplasticity in MMN, we will then primarily introduce different MMN responses between musicians and nonmusicians, its underlying mechanisms, and the practical applications. Finally, we will conclude with the role of MMN in music perception and point out further directions in music neuroscience research, for example, combining MMN and ERAN.

2. Brief Review on MMN

2.1. MMN in Sound. Research on MMN covers a wide range, from simple sound detection to complex music perception. Accordingly, our brief review on MMN begins with studies using sound stimuli followed by ones using music paradigms. Although MMN exists not only in ERP but also in other forms, such as MMN in MEG [45], MMN in PET [46], and MMN in fMRI [47, 48], the later ones are out of range of this review, and we will primarily focus on the MMN in ERP. MMN is originally defined as a *deviation-sensitive* component of auditory event-related potential (ERP). However, it is worthy to note that the MMN highly depends on the presence of a short-term *memory trace* in the auditory cortex representing the repetitive aspects of the preceding auditory events, which usually lasts for a few seconds. Thus, single sounds, even deviants, with no former sounds during the last few seconds elicit no MMN but enhance obligatory responses of P1, N1, and P2 [49].

By now, most of conclusions about MMN are induced by the research using an oddball paradigm that consists of two sounds with different auditory features [19]. And follow-up researchers only changed some auditory attributions based on the traditional oddball paradigm in a principle of keeping two variants of one feature (one is for standard, and the other is for deviant). The rationale of the traditional oddball paradigm is the detection of the rare deviant stimulus from the relatively high quantity of standard stimulus. For instance, a sequence of frequent, standard sounds (with a probability of occurrence of 0.8 to form a stable memory trace in the auditory cortex) and infrequent deviant sounds (with a probability of 0.2) is randomly presented.

Interestingly, the strength (i.e., amplitude) and the speed (i.e., latency) of the MMN signal are related with both the size of the deviance (i.e., how the infrequent stimulus deviates from the defined memory trace) and the probability of occurrence of the deviant but are independent of the demands of primary tasks. Increasing deviant probability (i.e., decreasing the occurrence probability of standard stimuli) leads to an attenuation of MMN amplitude, which is thought to be partially due to a weakened standard memory trace. More importantly, this attenuation may also result from conflict between “two memory traces.” That is, increased

probability of deviant stimuli may lead to development of their own memory trace, which in turn may inhibit the MMN generation of the standard stimuli.

Concerning effect of task demands, the MMN detects sound deviations even in preattentive condition, where the (in)frequent stimuli are presented while participants are engaging in other sound-irrelevant tasks, such as watching a silent movie with subtitles, playing a computer game, or reading a magazine [50, 51]. These results suggest that MMN is sensitive to auditory deviations implicitly with little effect from the explicit task demands.

Last but not least, previous studies have revealed that MMN is elicited not only by basic auditory features (e.g., sine tone and single instrumental tone) but also by different kinds of abstract changes in auditory stimulation such as language grammar and musical syntax violations (for a review, see [52]). That is, we can record the MMN effect even without the so-called “standard stimuli” [53].

2.2. Multifeature MMN: From Sound to Music. Almost before 2004, most of studies on MMN have primarily employed the traditional oddball paradigm or its variants. This paradigm, however, has two main questions: one is that the oddball paradigm requires long-time periods in experiments that is unsuitable to children participants and other participants with some cognitive diseases [38]. The other question relates to ecological validity. That is, the oddball paradigm is oversimplified, and in turn, only one MMN acoustic feature can be examined every time. Although the oddball paradigm is highly appropriate to assess sensory dissonance, understanding of realistic acoustic environments and music that have more complex structures and consist of rich components cannot be achieved by the oddball paradigm itself.

To address this issue, Näätänen et al. (2004) have investigated the MMN in a “multi-feature paradigm.” Compared to the traditional oddball paradigm, the multifeature paradigm enables fast recording of responses to several deviant types in one-stimulus sequence, such as intensity, frequency, duration, stimulus omission, timbre, and pitch contour [26]. In addition, the MMN acquired in the new multifeature paradigm is equal in amplitude to those in the traditional oddball paradigm. Thus, the multifeature paradigm is effective and more appropriate to investigation of the music-related MMN. However, there still exists a question. The multifeatures in Näätänen’s paradigm were presented without context and rules that might be different from real music.

2.3. MMN in Music. In the original multifeature paradigm by Näätänen et al. in 2004 [26], the deviant attributions consist of frequency, intensity, duration, location, and a silent gap, but, as discussed above, all these five features compose no real music due to lack of context and rules. Recently, Vuust et al. (2011) have demonstrated a new, fast, musical multifeature MMN paradigm, in which 6 types of acoustic changes, relevant to musical processing in different musical genres, are presented in the same sequence (termed as melodic multifeature paradigm) [28]. Specifically, 5 of the 6 musical features are aspects of the musical sound that elicit different MMNs depending on levels of musical expertise:

pitch mistuning, intensity, timbre, sound-source location, and rhythm [54–56]. And to create a paradigm that could be used to compare nonmusicians to musicians, as well as musicians with different musical genres, they included a pitch slide, which is typical for improvisational music instead of classical music [57].

In comparison with the traditional multifeature paradigm by Näätänen et al. (2004), the melodic multifeature paradigm has thus shown a greater similarity to real music. The traditional multifeature paradigm in music research is based on a simple musical figure, which is well-known in many genres of Western tonal music: the Alberti bass, an accompaniment that is originally encountered in classical music such as Mozart’s sonatas or Beethoven’s rondos, and is later adopted with variations in other contemporary musical genres. In contrast, the melodic multifeature paradigm is composed of short melodies and includes deviations of complex spectral and temporal regularities, such as melody, rhythm, key, timbre, tuning, and timing.

In sum, using multifeature MMN paradigms and the variants, researchers have demonstrated that the neural substrates underlying conscious musical experience may involve the following three stages: (1) the encoding and temporal integration of each sound characterized by its specific acoustic and perceptual features (e.g., pitch, duration, and timbre) into brief neural traces, (2) the simultaneous maintenance and integration of the neural traces for acoustic features leading possibly to the memorization of musical motifs, and (3) the modulation of sound perception by the memory of the previous sounds (for a review, see [58]).

2.4. MMN Features Specific to Music

2.4.1. From Pitch to Interval or Contour. Taking Western tonal music as an example, it is based on a small subset of 12 pitches in the chromatic equal-tempered scale, where the intervals between consecutive pitches are semitones. And this discrete selection of pitches and their relationships to each other form the most fundamental rules of Western tonal music, resulting in jazz, classical, pop musical genres, and so forth. Human beings can extract the pitch difference (interval) with a sensitivity of 10–30 cents [30]. In musical domain, pitch change (frequency change) within chromatic scale is generally named as “out of tune” or “mistune,” and pitch change within diatonic scale is named as “out of key.” Two different pitches form intervals, including melodic interval (tones played sequentially) and harmonic interval (tones played simultaneously). Sounds presented continuously are characterized not only by the pitch, but also by the contour, the perceived duration, the perceived loudness (which does not necessarily coincide with the physical intensity), the timbre, and so forth. These rules imply that the way, in which those musical features are processed, is another important factor for disclosing the neural mechanisms of music perception, distinguished with the general sound perception.

(1) Pitch Change as out of Tune. Previous studies have induced a paradox about preattentive versus attentive processing in pitch deviations using MMN paradigm. For instance,

Koelsch et al. (1999) have used the ERP MMN to assess the accuracy of *preattentive* pitch discrimination in expert versus novice violin players [59]. In their study, the standard stimuli consisted of major chords of three sinusoidal tones with a perfect major third and fifth. The deviant stimuli were the same with the standard chords, except that the middle tone of the chords was marginally mistuned. The behavioral data revealed that nonmusicians detected 10% of the deviant chords, whereas violin players detected 80%. Consistent with the behavioral result, the ERP data showed that the deviants elicited the MMN in the musicians, but not in the nonmusicians. Specifically, in the discrimination condition, a significant MMN was followed by N2b and P3b deflections only in musicians. These results suggest that pitch discrimination may be of crucial importance for well-trained violin players, whereas such automatic detection of pitch differences is not available for nonmusicians.

Tervaniemi et al. (2005) have further argued that the differences between musicians and nonmusicians in pitch change processing result from the effects of parametric manipulations on the magnitudes of pitch deviance and the subjects' attentional focus [60]. To confirm the specificity of pitch discrimination accuracy in musicians, all the participating musicians in their study reached professional levels in different instruments (i.e., guitar, wind, piano, and string instruments). The authors utilized small pitch differences in relation to threshold to provide evidence that musicians have smaller, just noticeable differences in pitch perception than nonmusicians. Meanwhile, large pitch differences were employed to examine whether and how musicians' superiority is reflected in behavioral and ERP measures at suprathreshold levels. Partially consistent with the conclusion by Koelsch et al. (1999), the behavioral data in Tervaniemi et al. (2005) have revealed that the superiority of pitch discrimination accuracy in musicians over nonmusicians was observed not only with the 0.8% (small pitch differences) but also with the 2% (large pitch differences) frequency changes. Here, we need to explain briefly that there were 3 kinds of deviant sounds (0.8%, 2%, or 4% changes in frequency), and the so-called "preattentive" (0.8%) and "attentive" (2% or 4%) level were defined by the just noticeable differences in pitch perception. Furthermore, nonmusicians could also detect quite reliable and small pitch changes of 0.8%. However, the ERP data, that is, the MMN and the P3a, did not differentiate between nonmusicians and musicians in both two levels of pitch differences. These results suggest that musical expertise may exert its effects merely at attentive levels but is not necessary at the preattentive levels.

(2) *Interval Made by Pitches*. Besides pitch perception, MMN is also affected by some specific interval structure changes in music, that is, the distance between two serial pitches or several simultaneous pitches. In an ERP study by Brattico et al. (2003), they revealed that MMN showed a larger amplitude to an infrequent dissonant interval (the major seventh) in a context of repeated consonant intervals (the major sixth) than to an infrequent consonant interval in a consonant context (e.g., the perfect octave replacing the major sixth) [61]. Importantly, the similar result was obtained

even if the pitch deviation between consonant standards and dissonant deviants was smaller in pitch distance than that between consonant deviants and consonant standards. This result may reject the assumption that MMN increases its amplitude with acoustically larger pitch shifts (e.g., Tiitinen et al. [62]). Accordingly, Tervaniemi and Brattico (2005) have demonstrated that human brain responds more vigorously to the distance in dissonance between infrequent musical interval and the repeated context than to the distance in pitch between their tone components.

(3) *Contour: A Kind of Sound Context*. Several studies have suggested that contour may be a more fundamental attribute for melody recognition than interval size [63]. It is a defining feature of melodies and hence is a potent cue to the identity of musical pieces, such as short note sequences that form recognizable "hooks" and "themes." Ample evidence has proved that contour processing provides an essential basis for melody recognition [64–66]. For example, Dowling [67] has investigated the role of both interval and contour in melody recognition. In the interval condition, subjects were asked to detect subtle deviations in interval size between referential and target melodies that were played in either the same or different keys. The results showed that subjects yielded more errors when the target melodies were presented in different keys than when those were presented in the original key. In the contour condition, subjects were instructed to detect deviations in the contours of target melodies that were played in either the same or different keys, and subjects' performance was robust across transpositions between different keys

Contour also provides a kind of musical sound context in MMN paradigms. Brattico et al. (2001) have compared musicians with nonmusicians on their accuracy of processing pitch deviations of identical magnitude in temporally complex auditory contexts (Western versus non-Western scales) and in single sounds [68]. The results showed that the pitch deviation in the Western condition evoked larger MMN amplitude than that in the non-Western condition and in the single-tone condition. This suggests that cognitive processing of pitch deviation between the subsequent tones is facilitated in a complex sound context with familiar frequency ratios (Western condition) compared to unfamiliar frequency ratios (non-Western condition) and single tones. In addition, violations of music scale rules in unfamiliar melodies also elicit an MMN-like response with main generators localized in the nonprimary auditory cortex [38]. All these findings suggest that the music MMN reflects violations of musical regularities that are stored in long-term memory, rather than that developed during the short-term memory formation.

2.4.2. *Attentive or Preattentive of Beat Processing?* Beat is perceived regularly as salient events in time sequence by listeners. And this has been suggested to be domain-specific as well as species-specific to human beings [69]. Although some previous studies have reported that rhythmic entrainment may be a byproduct of the vocal learning mechanisms that are shared by some kinds of animals, including human beings [70, 71], Honing et al. (2012) have used the oddball paradigm to compare different abilities to detect the deviant musical

features between human beings and rhesus monkeys (*Macaca mulatta*) and revealed that human beings detected the beat in music, but the rhesus monkeys did not [72].

Then the beat perception in human beings is an attentive or a preattentive processing? Some researchers have suggested that focused attention is generally necessary for both beat perception [64, 65] and regularity detection [66]. In contrast, others have evidenced that beat processing is in fact preattentive [73], since it runs in sleeping newborns [74]. In the original beat-processing paradigm, researchers manipulated temporal structures of rhythm as metrical structures [64] and highly syncopated rhythms [65]. Ladinig et al. (2009) later controlled metrical stimuli more strictly by adding the variation in the timbre and intensity of tones to convey the metrical structure [73]. And thus, it is more acoustically rich and ecologically valid, allowing listeners to detect a beat preattentively.

However, a question arises here: whether or not the different responses to tones in different metrical positions are due to acoustic differences rather than beat processing? To address this issue, Bouwer et al. (2014) have used an MMN paradigm to examine the beat processing [75]. In their study, both musicians and nonmusicians were presented with a varying rhythm with a clear accent structure, in which sound was omitted occasionally. They compared MMN responses to the omissions of identical sounds in different metrical positions. The result showed that omissions in strong metrical positions (on the beat) elicited higher amplitude MMN responses than those in weak positions (not on the beat). This suggests that detection of a beat is preattentive when highly beat inducing stimuli are used. Importantly, they did not find any effect of musical expertise, suggesting that beat processing in metrically simple rhythms with clear accents does not need attention or musical expertise. Therefore, we indeed need to pay much attention to the fact that how acoustically varying stimuli may influence ERP results in further studies.

2.4.3. MMN in Single Chord Manipulation. Western music has two classifications that are familiar to all Western listeners: dichotomy of minor and major modalities and that of consonance and dissonance. By changing the interval structure of a chord, there exist some chord MMN paradigms, such as discrimination of slightly mistuned chords [54, 63]. However, in previous MMN studies that used chords as standard and deviant stimuli, the deviant chord included tones that were not present in the standard chord. Consequently, one possibility is that MMN could be simply elicited by the infrequently occurred tones, without necessarily reflecting any higher-order, musical differences (minor versus major) between the standard and deviant chords.

To resolve this problem, Virtala et al. (2011) have examined processing of chords in a new paradigm, where different standard chords were presented [76]. The deviant chords consisted of the same notes as in the standards but were different in qualitative combinations. They used all possible root major triads as standard stimuli. In order to assess

discrimination of both the major-minor and consonance-dissonance categories, there were 3 kinds of deviant stimuli: minor chords, dissonant chords, and inverted major chords (each was represented by three different, equiprobable chords from that category). In addition, to ensure elicitation of a classical physical-feature MMN, their sequences also included occasional standard major chords but are slightly softer in loudness. In the ignorance condition, the MMN was significantly elicited by all types of chords, but not by the inverted major chords. In the detection condition, the MMN was significantly elicited by the dissonant chords and the soft target chords. However, whether the classifications of major versus minor modalities and consonance versus dissonance are innate or based on implicit or explicit learning remain a question for future studies.

2.4.4. Timbre Difference in MMN. Many researchers do agree that, compared with chord processing, different instrumental tones (timbre) are processed by a different cortical area [77], and the stimulus complexity might influence the processing of musical sounds. Using a novel, fast multifeature MMN paradigm, Vuust et al. (2012) have investigated preattentive processing of auditory features in musicians with expertise in four distinct styles of music (classical, jazz, rock, or pop) and in nonmusicians [78]. Their results revealed that jazz musicians showed the greatest auditory expertise effect, especially when compared with rock musicians and nonmusicians. One possible interpretation is that jazz is characterized by complex harmonies and elaborate melodic and harmonic materials that present a challenge to many listeners. Furthermore, since jazz is often improvised, the variations of the harmonic progressions and rhythmic feelings are more frequently communicated across jazz musicians [79]. Thus, specific auditory skills, required for performing different musical tasks such as conducting an orchestra [80, 81], playing certain instruments or musical genres, lead to sensitivity to specific sound features, inducing different MMN responses in the amplitude and the latency (for a review, see [82]).

3. Neural Plasticity in Music MMN

3.1. Do Musicians Show More MMN Effects? Musicians have various pronounced skills in auditory perception that correlate with their expertise in music. For instance, musicians can detect smaller pitch differences, but nonmusicians cannot. In addition, musicians show better performance in discriminating pitch intervals [83] and in structuring rhythms [84] than nonmusicians. And these superior auditory processing skills of musicians are reflected in the MMN. For example, chords that deviated from repetitive standard chords by only 0.75% in pitch are enough to elicit MMN in violinists, whereas a much larger pitch deviation is required to elicit MMN in musical novices.

3.1.1. Different Time Grouping. Musicians also have a longer time window for integrating sounds as indicated by MMN elicitation [85]. van Zuijlen et al. (2005) have showed that the auditory system of musicians organize sounds differently

from that of nonmusicians [86]. Furthermore, professionally trained classical musicians showed auditory grouping of a tone sequence into four-tone groups according to a more complex Gestalt-grouping rule than nonmusicians. The authors found evidence on auditory grouping according to pitch-similarity in both musicians and nonmusicians. According to a good-continuation-of-pitch, however, the grouping was only observed in musicians. Grouping of the sequence was indicated by MMN response to an occasional fifth tone that extended the length of the standard-tone groups. The deviant could therefore have been detected by encoding a temporal and/or numerical regularity.

Furthermore, van Zuijlen et al. (2005) have examined whether auditory system could encode a temporal regularity or a numerical regularity and the differences between musicians and nonmusicians. To achieve this goal, van Zuijlen et al. presented tone sequences containing either a temporal regularity or a numerical regularity to participants. The sequence with the temporal regularity could be divided into segments with a constant duration but contained a varying number of tones. And the sequence with the numerical regularity could be divided into segments containing a constant number of tones but varied in duration. Auditory encoding of the regularity was then determined by examining whether MMN was elicited by an occasional segment lengthening varied in time (duration) or in number. The results revealed that, in both musicians and nonmusicians, an MMN was elicited when the temporal regularity was unexpected. In contrast, an MMN was elicited to violations of the numerical regularity only in musicians. These results suggest that temporal processing is important in audition, since a complex temporal regularity can be encoded at an involuntary auditory processing stage regardless of musical expertise. Furthermore, the auditory system of professional musicians can encode a numerical regularity without attention, reflecting the functional importance of beat tracking in the perceptual organization of music.

3.1.2. What Factors Shaped Musician's Expertise? Musicians' brains are shaped by their training types, musical styles or genres, and listening experiences. Accordingly, their neural activations depend highly on the instruments they played and the practice strategies, as well as the levels of their expertise. For instance, musicians who need to intone while playing such as violinists normally show a greater sensitivity to small deviations in pitch than musicians playing other instruments and nonmusicians [78]. And singers also yield a stronger MMN response than instrumentalists to small pitch deviations [87]. In addition, rhythmic deviations elicited a stronger, faster, and more left-lateralized MMN in musicians than in nonmusicians [88]. More recently, it is reported that musicians who perform music primarily without a score learn more easily to detect contour deviations in melodic patterns than those who read music score while playing, and this difference is detected preattentively by an enhanced MMN response [89]. Taken together, these studies indicate that musicians' brains process auditory information differently according to their practice strategies, and the MMN is sensitive to acoustic features that are specific to musical objects or practice.

3.1.3. Different Speed and Magnitude of Neuroplasticity. The MMN is modulated by musicianship as well [85, 86]. For instance, increased ERPs are measured in musicians especially when the sounds are complex [85]. Consistently, in Seppänen's study (2013), musicians also showed enhanced auditory processing compared with nonmusicians [90]. They examined the neural bases of the musical training effects on rapid plasticity in auditory processing. The learning-related plastic changes in ERP responses to pitch and duration deviants between passive blocks were compared between musicians and nonmusicians. Passive blocks were interleaved with an active discrimination task, in which participants were instructed to focus on an unrelated task but to ignore the testing task. In contrast, in the active blocks participants were asked to pay their attention on both tasks. The results showed that deviant-related ERP responses around the parietal areas decreased after finishing the active task in both musicians and nonmusicians. Compared with nonmusicians, musicians showed larger MMN responses, especially to deviations in musical sounds such as chords [54], melodies [51], and rhythms [56] as well as other complex auditory stimuli [78, 91].

3.1.4. Differences of Brain Structure. Previous studies have revealed increased gray matter volume and density in auditory cortices of musicians [92, 93]. On the basis of the structural findings, changes in the auditory ERPs in musicians may indicate expanded activation areas, increased number of neurons, greater synchronization, or faster connectivity. For instance, the results of James et al. (2013) revealed rapid plastic changes in the bilateral temporal and the left frontal areas only in musicians, but not in nonmusicians [94].

3.2. How Do the MMN Effects Change over Time by Natural Exposure? It seems that human beings are born for detection of sound discriminations. The MMN responses to auditory stimuli have been successfully recorded even in fetus and newborns or preterm infants [95]. Previous research has mainly focused on five developmental periods: fetus, inborn, preschool children, adults, and aging individuals ([96], for a review, see [97]), and suggests that MMN is developmentally stable and has a unique U-shaped curve. The MMN amplitude is only slightly smaller in infants than school-age children, and there is no MMN difference between school-age children and adults. With respect to the MMN latency, it is slightly longer in infants than in adults but reaches adult level quickly by early school-age years. Child MMN, however, does not seem to be analogous to adult MMN. For example, compared with adult, a prominent MMN response could be obtained in both waking and sleep states in infants. Furthermore, the distribution of MMN scalp is relatively broader and more central in children than in adults (for a review, see [98]).

Concerning music conception, previous behavioral studies have demonstrated that, at around the age of six months, infants are already equipped with certain perceptual and cognitive prerequisites for putative beneficial effects of immersing in a musical enriched environment. They not only show fairly accurate discrimination of basic musical features

such as pitch and duration but also are sensitive to some abstract aspects of musical sounds. For example, infants, encoding melodies and rhythms depending on features of relative pitch and duration, are able to group individual tones by pitch and show long-term memory for musical pieces. More recently, ERP studies have shown that music-relevant auditory abilities, such as discrimination of different intervals, sound grouping, perception of missing fundamental, auditory stream segregation, and detection of the beat of rhythmic sounds, are available before the age of six months or even at birth.

These early perceptual skills are also important in childhood: young children who typically receive ample musical exposure tend to find music both interesting and enjoyable. Thus, everyday musical activities are rich sources of experience that may shape the development of auditory skills. In an ERP study, Trainor et al. (2011) randomly assigned 4-month-old infants either to a guitar group, in which they were exposed to recordings of melodies in a guitar timbre, or to a marimba group, in which they were presented with the same melodies in a marimba timbre [99]. After a week-long, 20 min per day exposure to one of the two timbres, the guitar-exposed infants showed a larger obligatory response to guitar tones than to marimba tones, and the opposite response pattern was found for the marimba-exposed group. Furthermore, occasional pitch deviations in guitar tones elicited an MMN only in the guitar-exposed infants, whereas pitch deviations in the marimba tones elicited a significant MMN in both groups. These results suggested that a relatively short-term exposure to music in infants is enough to strengthen the neural representations of a given timbre, which is further reflected by MMN response to that timbre pitch.

3.3. Could MMN Effects Be Trained? Recently, growing studies have provided evidence on a modulation role of training (or experience) in MMN effects. For instance, François et al. (2012) have demonstrated that 8–10-year-old children who had been randomly assigned to (12 months) a music lesson group showed a larger increase in MMN amplitude than the control group children who were assigned to a painting lesson [100]. Consistently, Putkinen and Saarikivi (2013) have also revealed that the MMN elicited by minor chord deviants in a sequence of major chord standards showed more amplitude increase in music-trained children (7–13 years old) than in children without music training [101].

It is noteworthy that the children participants in music studies all received formal training of music. However, in our daily life only a few people have opportunities to obtain professional musical trainings. Most of normal people are nonmusicians but amateurs or just influenced by musical contexts in an implicit way. For instance, various computer and console games attract children into “musical play,” which could be considered as informal environments for music learning [101]. Then, an important question is that do these informal musical activities affect development of auditory discrimination and attention in preschool-aged children? This has been confirmed at least in adult participants. In previous behavioral studies, adults who have no experience of musical training showed (implicit) competence in processing

of some fairly nuanced aspects of music, as a similar manner of learning through mere incidental exposure.

Previous ERP studies have demonstrated that brains of nonmusicians process some aspects of Western tonality and harmony automatically [102]. One assumption is that these idiosyncrasies of Western tonal music are internalized in nonmusicians via everyday musical experiences. Recently, Putkinen and Saarikivi (2013) have investigated an issue of whether amount of informal musical activities is related to electrophysiological correlates of auditory deviation detection in 2–3-year-old children [101]. They used a multifeature MMN paradigm (Näätänen et al., 2004) to assess several auditory ERP responses and revealed that the MMN reflects different stages of auditory processing particularly in childhood. This result suggests that ambient exposure to musical contexts without special training is sufficient for learning of culture-specific musical knowledge. More importantly, one question rises as whether informal exposure to music may also modulate development of auditory processing beyond musical domain.

4. Conclusions and Further Directions

There is no doubt that MMN has opened a window for better understanding of music perception from a perspective of cognitive neuroscience. The MMN paradigms have been gradually evoking to get close to the real music. And consequently, it has become an effective index of perception of regularities in music. By now, almost all neuroscience studies on music had found the MMN features in basic levels of music, such as pitches, timbres, intensities, and short-term melodies. And the temporal processing in MMN is similar among these basic musical elements. That is, the MMN normally peaks at around 100–200 ms after deviation onset, and its amplitude and latency alter depending on deviation magnitude and related perceptual discriminability, which may be modulated by expertise effects and musical sound contexts. Concerning the neural basis of MMN, we still cannot reach a common idea about the exact position of the MMN's main source in brain due to nonuniform musical materials and paradigms. It is certain that, however, our brain is inborn to detect MMN features in music, and this ability changes over time in U-shaped route [103].

On the other hand, MMN has shown its feature of plasticity (i.e., neuroplasticity). For instance, different MMN responses have been consistently revealed between musicians and nonmusicians. And these differences mainly resulted from training efforts, which include natural exposure to music contexts and informal musical trainings. Because of its features of deviation-sensitivity and neuroplasticity, MMN has also functioned well in some clinical and practical areas, such as assessment of the wholeness of central auditory system in infants, detection of innate predisposition for musical preferences, and treatment of musical abilities of adult cochlear implant users. Although we have made great achievements in music MMN, there still exist many open questions needed to be resolved in future studies.

4.1. MMN and Music Syntax. The relationship between serial and hierarchical processing has been discussed in language literature [104]. Little study, however, has addressed this issue in music perception. Since music has complicated structures and regularities, perception of music involves various cognitive functions. In contrast, MMN is one indicator of brain responses to music and reflects primarily low level of music elements. The higher hierarchical structures of music, such as chord progressions (harmonic progressions), may be also a key component of music perception and are out of reach of MMN. Previous research has suggested that there is no single way to fully understand syntax in music, but combination of MMN and an early right-anterior negativity (ERAN) may create a new direction of this area. Both ERAN and MMN are event-related potentials elicited by detection of deviations of regularities and have similar peak amplitude and latency. MMN mainly detects abstract-feature deviants (frequency, intensity, location, gap, etc.), while ERAN is an ERP response to harmonically inappropriate chords when listening to music. That is, MMN representations of regularities of intersound relationships are extracted online from acoustic environments, whereas ERAN represents music-syntactic regularities that already exist in long-term memory, which have been shaped by different musical cultures (for a review, see [105]). Taken together, to fully understand the neural bases of music perception, in the current stage, it is necessary to combine ERAN and MMN in music neuroscience research. However, the relationship between ERAN and MMN needs to be examined further.

4.2. MMN from a Developmental Perspective. Previous studies have suggested that we process MMN features congenitally. And this has been evidenced by several research results, such as frequency and intensity discrimination in fetus and infants [95]. To confirm this assumption, however, it is essential to examine more acoustic features, music features in particular. The neuroscience studies on real music features enable us to better understand a series of important issues on music perception/apprehension; for instance, whether or not human beings have an innate ability to distinguish musical regularities by MMN, and which MMN features are flexible to be influenced via training? When do MMN differences between musicians and nonmusicians begin to develop? And whether or not significant brain structural changes induced by early training were sustained even if musical training ceased thereafter?

Research outside of music domain in healthy adults has reported that brain structural changes that resulted from a complex motor task (e.g., juggling) normally occur within 1 week of training onset but return to baseline without ongoing training. In field of music neuroscience, however, no study has focused on these points. And two main issues needed to be considered in future studies are that “is there any ‘optimal period’ that maximizes difference during music training” and “whether or not levels of musical expertise can slow down attenuation of MMN effects”?

4.3. Wide Applicability of Musical MMN. MMN is well accepted for its features of being noninvasive and having no need for explicit awareness (preattentive) and hence has a wide range of applications in research and clinical areas. For example, MMN recorded in an unconscious coma patient provides reliable predictor of recovery of consciousness. With respect to MMN in sleep, it has weak effect in adults, but in infants and newborns MMN is very helpful in assessing their central auditory functions. It is worthy to note that, besides healthy individuals, MMN paradigm is also important for individuals with specific diseases or disabilities, such as cochlear implant (CI) users and individuals with schizophrenia. Specifically, for the CI users, MMN paradigm is not only an evaluation criterion to the quality of implantation, but also a good investigation tool to assess residual neural functions of processing musical sound features or inborn musical abilities [98]. However, this still exists as an academic discussion now, and it is necessary to develop related products following rationale of MMN in near future.

4.4. Studying Music in a More Musical Context. Music is a fine-grained aesthetic entity. When presented with simple sinusoidal tones with greater mistuning instead of fine-grained differences, violinists were not superior to nonmusicians in discriminating pitch deviations. A similar lack of MMN differences was also obtained when comparing processing of isolated infrequent sinusoidal tones or infrequent minor chords within a sequence of major chords between musicians and nonmusicians [38]. Seppänen et al. (2007) have compared musicians who mainly employ auditory rehearsal and playing strategies to a nonaural group as determined by a questionnaire. They found that practice strategies modulate the speed of neural discrimination of interval and contour deviations presented in melody-like patterns [89], but not in simple sound features. Although so-called “perfect melodic multi-feature MMN paradigm” [28] has showed much better ecological effects than other paradigms, it is just music-like. Music has various styles, consisting of different genres and composing ways, and thus it will be hard for researchers to develop a unified, standard paradigm that represents all acoustic features. Therefore, it is necessary to classify music styles first and then design different music paradigms correspondingly.

4.5. Daily Interactions with Music. Previous music MMN studies have been primarily conducted in well-controlled laboratory environments and in turn lack ecological effects leading to distinctive results in different musical contexts. One possibility is that different participants may have different levels of interactions with music in their daily life [102]. Some researchers have investigated several indices of musical experiences of participating children from their parents, such as how often the children engaged in different types of musical activities at home (e.g., singing and dancing) and how often they played music with their children (e.g., how often they sang to their children or sang with their children together). And a composite score representing the amount of such musical activities in daily life was significantly correlated

with the ERP response amplitudes, such as enlarged P3a responses to duration and gap deviants and a diminished LDN (late discriminative negativity) across all five deviant types.

4.6. Enculture. From a perspective of society, few people will deny that music is a universal language across different cultures. However, music still does not reach a common consensus cross-culturally in cognitive neuroscience level. Previous studies have consistently demonstrated that differences in cultural backgrounds affect the way in which people perceive auditory signals [106]. For instance, subjects from tonal language cultures (e.g., Chinese) performed better than those from nontonal language cultures (e.g., French) in pitch discrimination [107]. Accordingly, various open questions, such as whether there is an early sensitive period for the culture effects, whether the culture effects are relatively stable or malleable by musical experience, and whether the children who experience different musical cultures have the same level of sensitivities to musical MMN, are all interesting topics.

On the other hand, previous research of music neuroscience has mainly focused on Western music system (i.e., diatonic scale or equal temperament). Little is known, however, about the MMN in non-Western music (i.e., pentatonic scale). Therefore, we also need to pay much attention to non-Western music in further studies, such as ancient Chinese music, Indic music, and Japanese music. That is, the similarities and differences of processing of music perception cross-cultures are needed to be examined via intercultural comparison for better understanding of neural mechanisms of music perception/apprehension and music per se.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Authors' Contribution

The authors Xide Yu and Tao Liu contributed equally.

Acknowledgments

Dingguo Gao was supported by the “National Natural Science Foundation of China” [31371028] and Tao Liu was supported by the “Fundamental Research Funds for the Central Universities” [26000-31610113].

References

- [1] C. L. Spohn, “Programming the basic materials of music for self-instructional development of aural skills,” *Journal of Research in Music Education*, vol. 11, no. 2, pp. 91–98, 1963.
- [2] S. F. Michalski Jr., “Development and evaluation of a visual-aural program in conceptual understanding of the basic elements of music,” *Journal of Research in Music Education*, vol. 19, no. 1, pp. 92–97, 1971.
- [3] E. G. Schellenberg, “Music and cognitive abilities,” *Current Directions in Psychological Science*, vol. 14, no. 6, pp. 317–320, 2005.
- [4] S. Feld and A. A. Fox, “Music and language,” *Annual Review of Anthropology*, vol. 23, no. 1, pp. 25–53, 1994.
- [5] W. J. Dowling and B. Tillmann, “Memory improvement while hearing music: effects of structural continuity on feature binding,” *Music Perception*, vol. 32, no. 1, pp. 11–32, 2014.
- [6] D. J. Hargreaves and A. C. North, “The functions of music in everyday life: redefining the social in music psychology,” *Psychology of Music*, vol. 27, no. 1, pp. 71–83, 1999.
- [7] M. Sambandham and V. Schirm, “Music as a nursing intervention for residents with Alzheimer’s disease in long-term care. Music may be a memory trigger for patients with Alzheimer’s and provide a means of communication,” *Geriatric Nursing*, vol. 16, no. 2, pp. 79–83, 1995.
- [8] M. Brotons and P. K. Pickett-Cooper, “The effects of music therapy intervention on agitation behaviors of Alzheimer’s Disease Patients,” *Journal of Music Therapy*, vol. 33, no. 1, pp. 2–18, 1996.
- [9] A. M. Kumar, F. Tims, D. G. Cruess et al., “Music therapy increases serum melatonin levels in patients with Alzheimer’s disease,” *Alternative Therapies in Health and Medicine*, vol. 5, no. 6, pp. 49–57, 1999.
- [10] L. L. Cuddy and J. Duffin, “Music, memory, and Alzheimer’s disease: is music recognition spared in dementia, and how can it be assessed?” *Medical Hypotheses*, vol. 64, no. 2, pp. 229–235, 2005.
- [11] A. Moussard, E. Bigand, S. Belleville, and I. Peretz, “Music as an aid to learn new verbal information in Alzheimer’s disease,” *Music Perception*, vol. 29, no. 5, pp. 521–531, 2012.
- [12] R. G. Deason, N. R. Simmons-Stern, B. S. Frustace, B. A. Ally, and A. E. Budson, “Music as a memory enhancer: differences between healthy older adults and patients with Alzheimer’s disease,” *Psychomusicology: Music, Mind, and Brain*, vol. 22, no. 2, pp. 175–179, 2012.
- [13] C. Pacchetti, F. Mancini, R. Aglieri, C. Fundaró, E. Martignoni, and G. Nappi, “Active music therapy in Parkinson’s disease: an integrative method for motor and emotional rehabilitation,” *Psychosomatic Medicine*, vol. 62, no. 3, pp. 386–393, 2000.
- [14] M. W. Hardy and A. B. LaGasse, “Rhythm, movement, and autism: using rhythmic rehabilitation research as a model for autism,” *Frontiers in Integrative Neuroscience*, vol. 19, no. 7, pp. 19–23, 2013.
- [15] S. Hegde, “Music-based cognitive remediation therapy for patients with traumatic brain injury,” *Frontiers in Neurology*, vol. 5, no. 34, pp. 1–7, 2014.
- [16] M. Besson, F. Faïta, and J. Requin, “Brain waves associated with musical incongruities differ for musicians and non-musicians,” *Neuroscience Letters*, vol. 168, no. 1-2, pp. 101–105, 1994.
- [17] A. R. Damasio and H. Damasio, “Musical faculty and cerebral dominance,” in *Music and the Brain*, M. Critchley and R. A. Henson, Eds., pp. 141–155, Camelot Press, Southampton, UK, 1978.
- [18] M. Tervaniemi, T. Ilvonen, K. Karma, K. Alho, and R. Näätänen, “The musical brain: brain waves reveal the neurophysiological basis of musicality in human subjects,” *Neuroscience Letters*, vol. 226, no. 1, pp. 1–4, 1997.
- [19] R. Näätänen, P. Paavilainen, K. Alho, K. Reinikainen, and M. Sams, “The mismatch negativity to intensity changes in an auditory stimulus sequence,” *Electroencephalography and Clinical Neurophysiology*, vol. 40, pp. 125–131, 1987.

- [20] R. Hari, M. Hämäläinen, R. Ilmoniemi et al., “Responses of the primary auditory cortex to pitch changes in a sequence of tone pips: neuromagnetic recordings in man,” *Neuroscience Letters*, vol. 50, no. 1–3, pp. 127–132, 1984.
- [21] K. Alho, M. Huutilainen, H. Tiitinen, R. J. Ilmoniemi, J. Knuutila, and R. Naatanen, “Memory-related processing of complex sound patterns in human auditory cortex: a MEG study,” *NeuroReport*, vol. 4, no. 4, pp. 391–394, 1993.
- [22] S. Shalgi and L. Y. Deouell, “Direct evidence for differential roles of temporal and frontal components of auditory change detection,” *Neuropsychologia*, vol. 45, no. 8, pp. 1878–1888, 2007.
- [23] M.-H. Giard, F. Perrin, J. Pernier, and P. Bouchet, “Brain generators implicated in the processing of auditory stimulus deviance: a topographic event-related potential study,” *Psychophysiology*, vol. 27, no. 6, pp. 627–640, 1990.
- [24] T. Rinne, K. Alho, R. J. Ilmoniemi, J. Virtanen, and R. Näätänen, “Separate time behaviors of the temporal and frontal mismatch negativity sources,” *NeuroImage*, vol. 12, no. 1, pp. 14–19, 2000.
- [25] R. Näätänen, A. W. K. Gaillard, and S. Mäntysalo, “Early selective-attention effect on evoked potential reinterpreted,” *Acta Psychologica*, vol. 42, no. 4, pp. 313–329, 1978.
- [26] R. Näätänen, S. Pakarinen, T. Rinne, and R. Takegata, “The mismatch negativity (MMN): towards the optimal paradigm,” *Clinical Neurophysiology*, vol. 115, no. 1, pp. 140–144, 2004.
- [27] P. Paavilainen, P. Arajärvi, and R. Takegata, “Preattentive detection of nonsalient contingencies between auditory features,” *NeuroReport*, vol. 18, no. 2, pp. 159–163, 2007.
- [28] P. Vuust, E. Brattico, E. Glerean et al., “New fast mismatch negativity paradigm for determining the neural prerequisites for musical ability,” *Cortex*, vol. 47, no. 9, pp. 1091–1098, 2011.
- [29] R. Torppa, E. Salo, T. Makkonen et al., “Cortical processing of musical sounds in children with Cochlear Implants,” *Clinical Neurophysiology*, vol. 123, no. 10, pp. 1966–1979, 2012.
- [30] C. L. Krumhansl, “Rhythm and pitch in music cognition,” *Psychological Bulletin*, vol. 126, no. 1, pp. 159–179, 2000.
- [31] C. A. Christmann, T. Lachmann, and S. Berti, “Earlier timbre processing of instrumental tones compared to equally complex spectrally rotated sounds as revealed by the mismatch negativity,” *Neuroscience Letters*, vol. 581, pp. 115–119, 2014.
- [32] K. N. Goydke, E. Altenmüller, J. Möller, and T. F. Münte, “Changes in emotional tone and instrumental timbre are reflected by the mismatch negativity,” *Cognitive Brain Research*, vol. 21, no. 3, pp. 351–359, 2004.
- [33] C. Lappe, O. Steinsträter, and C. Pantev, “Rhythmic and melodic deviations in musical sequences recruit different cortical areas for mismatch detection,” *Frontiers in Human Neuroscience*, vol. 7, article 260, 2013.
- [34] F. L. Bouwer, T. L. Van Zuijlen, and H. Honing, “Beat processing is pre-attentive for metrically simple rhythms with clear accents: an ERP study,” *PLoS ONE*, vol. 9, no. 5, Article ID e97467, 2014.
- [35] P. Virtala, M. Huutilainen, E. Partanen, and M. Tervaniemi, “Musicianship facilitates the processing of Western music chords—an ERP and behavioral study,” *Neuropsychologia*, vol. 61, no. 1, pp. 247–258, 2014.
- [36] C. L. Krumhansl and L. L. Cuddy, “A theory of tonal hierarchies in music,” in *Music Perception*, M. R. Jones, R. Fay, and A. N. Popper, Eds., vol. 36 of *Springer Handbook of Auditory Research*, pp. 51–87, Springer, New York, NY, USA, 2010.
- [37] I. Cross, “Music, cognition, culture, and evolution,” *Annals of the New York Academy of Sciences*, vol. 930, pp. 28–42, 2001.
- [38] E. Brattico, M. Tervaniemi, R. Näätänen, and I. Peretz, “Musical scale properties are automatically processed in the human auditory cortex,” *Brain Research*, vol. 1117, no. 1, pp. 162–174, 2006.
- [39] R. Matsunaga, K. Yokosawa, and J.-I. Abe, “Magnetoencephalography evidence for different brain subregions serving two musical cultures,” *Neuropsychologia*, vol. 50, no. 14, pp. 3218–3227, 2012.
- [40] Y. Nan, T. R. Knösche, and A. D. Friederici, “The perception of musical phrase structure: a cross-cultural ERP study,” *Brain Research*, vol. 1094, no. 1, pp. 179–191, 2006.
- [41] Y. Nan, T. R. Knösche, and A. D. Friederici, “Non-musicians’ perception of phrase boundaries in music: a cross-cultural ERP study,” *Biological Psychology*, vol. 82, no. 1, pp. 70–81, 2009.
- [42] Y. Nan, T. R. Knösche, S. Zysset, and A. D. Friedend, “Cross-cultural music phrase processing: an fMRI study,” *Human Brain Mapping*, vol. 29, no. 3, pp. 312–328, 2008.
- [43] S. Koelsch, T. C. Gunter, E. Schröger, M. Tervaniemi, D. Sammler, and A. D. Friederici, “Differentiating ERAN and MMN: an ERP study,” *NeuroReport*, vol. 12, no. 7, pp. 1385–1389, 2001.
- [44] I. Peretz and M. Coltheart, “Modularity of music processing,” *Nature Neuroscience*, vol. 6, no. 7, pp. 688–691, 2003.
- [45] H. Thönnessen, M. Zvyagintsev, K. C. Harke et al., “Optimized mismatch negativity paradigm reflects deficits in schizophrenia patients. A combined EEG and MEG study,” *Biological Psychology*, vol. 77, no. 2, pp. 205–216, 2008.
- [46] B. W. Müller, M. Jüptner, W. Jentzen, and S. P. Müller, “Cortical activation to auditory mismatch elicited by frequency deviant and complex novel sounds: a PET study,” *NeuroImage*, vol. 17, no. 1, pp. 231–239, 2002.
- [47] S. Molholm, A. Martinez, W. Ritter, D. C. Javitt, and J. J. Foxe, “The neural circuitry of pre-attentive auditory change-detection: an fMRI study of pitch and duration mismatch negativity generators,” *Cerebral Cortex*, vol. 15, no. 5, pp. 545–551, 2005.
- [48] T. Rosburg, P. Trautner, T. Dietl et al., “Subdural recordings of the mismatch negativity (MMN) in patients with focal epilepsy,” *Brain*, vol. 128, no. 4, pp. 819–828, 2005.
- [49] G. M. Bidelman, “Towards an optimal paradigm for simultaneously recording cortical and brainstem auditory evoked potentials,” *Journal of Neuroscience Methods*, vol. 241, pp. 94–100, 2015.
- [50] K. Alho, “Selective attention in auditory processing as reflected by event-related brain potentials,” *Psychophysiology*, vol. 29, no. 3, pp. 247–263, 1992.
- [51] T. Fujioka, L. J. Trainor, B. Ross, R. Kakigi, and C. Pantev, “Musical training enhances automatic encoding of melodic contour and interval structure,” *Journal of Cognitive Neuroscience*, vol. 16, no. 6, pp. 1010–1021, 2004.
- [52] R. Näätänen, P. Paavilainen, T. Rinne, and K. Alho, “The mismatch negativity (MMN) in basic research of central auditory processing: a review,” *Clinical Neurophysiology*, vol. 118, no. 12, pp. 2544–2590, 2007.
- [53] S. Pakarinen, M. Huutilainen, and R. Näätänen, “The mismatch negativity (MMN) with no standard stimulus,” *Clinical Neurophysiology*, vol. 121, no. 7, pp. 1043–1050, 2010.
- [54] E. Brattico, K. J. Pallesen, O. Varyagina et al., “Neural discrimination of nonprototypical chords in music experts and laymen: an MEG study,” *Journal of Cognitive Neuroscience*, vol. 21, no. 11, pp. 2230–2244, 2009.

- [55] C. Pantev, B. Ross, T. Fujioka, L. J. Trainor, M. Schulte, and M. Schulz, "Music and learning induced cortical plasticity," *Annals of the New York Academy of Sciences*, vol. 999, no. 1, pp. 438–450, 2003.
- [56] P. Vuust, L. Ostergaard, K. J. Pallesen, C. Bailey, and A. Roepstorff, "Predictive coding of music—brain responses to rhythmic incongruity," *Cortex*, vol. 45, no. 1, pp. 80–92, 2009.
- [57] P. Vuust, K. J. Pallesen, C. Bailey et al., "To musicians, the message is in the meter: pre-attentive neuronal responses to incongruent rhythm are left-lateralized in musicians," *NeuroImage*, vol. 24, no. 2, pp. 560–564, 2005.
- [58] M. Tervaniemi and E. Brattico, "From sounds to music towards understanding the neurocognition of musical sound perception," *Journal of Consciousness Studies*, vol. 11, no. 3–4, pp. 27–3, 2004.
- [59] S. Koelsch, E. Schröger, and M. Tervaniemi, "Superior pre-attentive auditory processing in musicians," *NeuroReport*, vol. 10, no. 6, pp. 1309–1313, 1999.
- [60] M. Tervaniemi, V. Just, S. Koelsch, A. Widmann, and E. Schröger, "Pitch discrimination accuracy in musicians vs non-musicians: an event-related potential and behavioral study," *Experimental Brain Research*, vol. 161, no. 1, pp. 1–10, 2005.
- [61] E. Brattico, M. Tervaniemi, V. Valimäki, T. van Zuijen, and I. Peretz, "Cortical correlates of acquired deafness to dissonance," *Annals of the New York Academy of Sciences*, vol. 999, pp. 158–160, 2003.
- [62] H. Tiitinen, P. May, K. Reinikainen, and R. Näätänen, "Attentive novelty detection in humans is governed by pre-attentive sensory memory," *Nature*, vol. 372, no. 6501, pp. 90–92, 1994.
- [63] Y.-S. Lee, P. Janata, C. Frost, M. Hanke, and R. Granger, "Investigation of melodic contour processing in the brain using multivariate pattern-based fMRI," *NeuroImage*, vol. 57, no. 1, pp. 293–300, 2011.
- [64] E. Geiser, E. Ziegler, L. Jancke, and M. Meyer, "Early electrophysiological correlates of meter and rhythm processing in music perception," *Cortex*, vol. 45, no. 1, pp. 93–102, 2009.
- [65] H. L. Chapin, T. Zanto, K. J. Jantzen, S. J. Kelso, F. Steinberg, and E. W. Large, "Neural responses to complex auditory rhythms: the role of attending," *Frontiers in Psychology*, vol. 1, article 224, 18 pages, 2010.
- [66] M. Schwartz, K. Rothermich, M. Schmidt-Kassow, and S. A. Kotz, "Temporal regularity effects on pre-attentive and attentive processing of deviance," *Biological Psychology*, vol. 87, no. 1, pp. 146–151, 2011.
- [67] W. J. Dowling, "Scale and contour: two components of a theory of memory for melodies," *Psychological Review*, vol. 85, no. 4, pp. 341–354, 1978.
- [68] E. Brattico, R. Näätänen, and M. Tervaniemi, "Context effects on pitch perception in musicians and non-musicians: evidence from ERP recordings," *Music Perception*, vol. 19, no. 2, pp. 199–222, 2001.
- [69] H. Honing, "Without it no music: beat induction as a fundamental musical trait," *Annals of the New York Academy of Sciences*, vol. 1252, no. 1, pp. 85–91, 2012.
- [70] A. D. Patel, J. R. Iversen, M. R. Bregman, and I. Schulz, "Experimental evidence for synchronization to a musical beat in a nonhuman animal," *Current Biology*, vol. 19, no. 10, pp. 827–830, 2009.
- [71] W. Zarco, H. Merchant, L. Prado, and J. C. Mendez, "Subsecond timing in primates: comparison of interval production between human subjects and rhesus monkeys," *Journal of Neurophysiology*, vol. 102, no. 6, pp. 3191–3202, 2009.
- [72] H. J. Honing, H. Merchant, G. P. Hádén, L. Prado, and R. Bartolo, "Rhesus monkeys (*Macaca mulatta*) detect rhythmic groups in music, but not the beat," *PLoS ONE*, vol. 7, no. 12, Article ID e51369, 2012.
- [73] O. Ladinig, H. Honing, G. Hádén, and I. Winkler, "Probing attentive and preattentive emergent meter in adult listeners without extensive music training," *Music Perception*, vol. 26, no. 4, pp. 377–386, 2009.
- [74] I. Winkler, G. P. Hádén, O. Ladinig, I. Sziller, and H. Honing, "Newborn infants detect the beat in music," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 106, no. 7, pp. 2468–2471, 2009.
- [75] F. L. Bouwer, T. L. van Zuijen, and H. Honing, "Beat processing is pre-attentive for metrically simple rhythms with clear accents: an ERP study," *PLoS ONE*, vol. 9, no. 5, Article ID e97467, 2014.
- [76] P. Virtala, V. Berg, M. Kivioja et al., "The preattentive processing of major vs. minor chords in the human brain: an event-related potential study," *Neuroscience Letters*, vol. 487, no. 3, pp. 406–410, 2011.
- [77] K. Alho, M. Tervaniemi, M. Huottilainen et al., "Processing of complex sounds in the human auditory cortex as revealed by magnetic brain responses," *Psychophysiology*, vol. 33, no. 4, pp. 369–375, 1996.
- [78] P. Vuust, E. Brattico, M. Seppänen, R. Näätänen, and M. Tervaniemi, "The sound of music: differentiating musicians using a fast, musical multi-feature mismatch negativity paradigm," *Neuropsychologia*, vol. 50, no. 7, pp. 1432–1443, 2012.
- [79] P. Vuust and A. Roepstorff, "Listen up! Polyrhythms in brain and music," *Cognitive Semiotics*, vol. 3, pp. 131–159, 2008.
- [80] T. F. Münte, C. Kohlmetz, W. Nager, and E. Altenmüller, "Superior auditory spatial tuning in conductors," *Nature*, vol. 409, no. 6820, p. 580, 2001.
- [81] W. Nager, C. Kohlmetz, E. Altenmüller, A. Rodriguez-Fornells, and T. F. Münte, "The fate of sounds in conductors' brains: an ERP study," *Cognitive Brain Research*, vol. 17, no. 1, pp. 83–93, 2003.
- [82] M. Tervaniemi, "Musicians—same or different," *Annals of the New York Academy of Sciences*, vol. 1169, pp. 151–156, 2009.
- [83] L. J. Trainor, R. N. Desjardins, and C. Rockel, "A comparison of contour and interval processing in musicians and nonmusicians using event-related potentials," *Australian Journal of Psychology*, vol. 51, no. 3, pp. 147–153, 1999.
- [84] C. Drake, A. Penel, and E. Bigand, "Tapping in time with mechanically and expressively performed music," *Music Perception*, vol. 18, no. 1, pp. 1–23, 2000.
- [85] J. Rüsseler, E. Altenmüller, W. Nager, C. Kohlmetz, and T. F. Münte, "Event-related brain potentials to sound omissions differ in musicians and non-musicians," *Neuroscience Letters*, vol. 308, no. 1, pp. 33–36, 2001.
- [86] T. L. van Zuijen, E. Sussman, I. Winkler, R. Näätänen, and M. Tervaniemi, "Auditory organization of sound sequences by a temporal or numerical regularity—a mismatch negativity study comparing musicians and non-musicians," *Cognitive Brain Research*, vol. 23, no. 2–3, pp. 270–276, 2005.
- [87] D. A. Nikjeh, J. J. Lister, and S. A. Frisch, "Hearing of note: an electrophysiologic and psychoacoustic comparison of pitch discrimination between vocal and instrumental musicians," *Psychophysiology*, vol. 45, no. 6, pp. 994–1007, 2008.
- [88] P. Vuust, K. J. Pallesen, C. Bailey et al., "To musicians, the message is in the meter: pre-attentive neuronal responses to incongruent rhythm are left-lateralized in musicians," *NeuroImage*, vol. 24, no. 2, pp. 560–564, 2005.

- [89] M. Seppänen, E. Brattico, and M. Tervaniemi, "Practice strategies of musicians modulate neural processing and the learning of sound-patterns," *Neurobiology of Learning and Memory*, vol. 87, no. 2, pp. 236–247, 2007.
- [90] M. Seppänen, J. Hämäläinen, A.-K. Pesonen, and M. Tervaniemi, "Passive sound exposure induces rapid perceptual learning in musicians: event-related potential evidence," *Biological Psychology*, vol. 94, no. 2, pp. 341–353, 2013.
- [91] S. C. Herholz, B. Boh, and C. Pantev, "Musical training modulates encoding of higher-order regularities in the auditory cortex," *The European Journal of Neuroscience*, vol. 34, no. 3, pp. 524–529, 2011.
- [92] P. Schneider, M. Scherg, H. G. Dosch, H. J. Specht, A. Gutschalk, and A. Rupp, "Morphology of Heschl's gyrus reflects enhanced activation in the auditory cortex of musicians," *Nature Neuroscience*, vol. 5, no. 7, pp. 688–694, 2002.
- [93] C. Pantev, R. Oostenveld, A. Engelien, B. Ross, L. E. Roberts, and M. Hoke, "Increased auditory cortical representation in musicians," *Nature*, vol. 392, no. 6678, pp. 811–814, 1998.
- [94] C. E. James, M. S. Oechslin, D. Van De Ville, C.-A. Hauert, C. Descloux, and F. Lazeyras, "Musical training intensity yields opposite effects on grey matter density in cognitive versus sensorimotor networks," *Brain Structure and Function*, vol. 219, no. 1, pp. 353–366, 2014.
- [95] R. Draganova, H. Eswaran, P. Murphy, M. Huotilainen, C. Lowery, and H. Preissl, "Sound frequency change detection in fetuses and newborns, a magnetoencephalographic study," *NeuroImage*, vol. 28, no. 2, pp. 354–361, 2005.
- [96] M. Holst, H. Eswaran, C. Lowery, P. Murphy, J. Norton, and H. Preissl, "Development of auditory evoked fields in human fetuses and newborns: a longitudinal MEG study," *Clinical Neurophysiology*, vol. 116, no. 8, pp. 1949–1955, 2005.
- [97] R. Näätänen, T. Kujala, C. Escera et al., "The mismatch negativity (MMN)—a unique window to disturbed central auditory processing in ageing and different clinical conditions," *Clinical Neurophysiology*, vol. 123, no. 3, pp. 424–458, 2012.
- [98] M. Cheour, E. Kushnerenko, R. Čeponiene, V. Fellman, and R. Näätänen, "Electric brain responses obtained from newborn infants to changes in duration in complex harmonic tones," *Developmental Neuropsychology*, vol. 22, no. 2, pp. 471–479, 2002.
- [99] L. J. Trainor, K. Lee, and D. J. Bosnyak, "Cortical plasticity in 4-month-old infants: specific effects of experience with musical timbres," *Brain Topography*, vol. 24, no. 3–4, pp. 192–203, 2011.
- [100] C. François, J. Chobert, M. Besson, and D. Schön, "Music training for the development of speech segmentation," *Cerebral Cortex*, vol. 23, no. 9, pp. 2038–2043, 2013.
- [101] V. Putkinen and K. Saarikivi, "Do informal musical activities shape auditory skill development in preschool-age children?" *Frontiers in Psychology*, vol. 4, article 572, 6 pages, 2013.
- [102] K. I. Krohn, E. Brattico, V. Välimäki, and M. Tervaniemi, "Neural representations of the hierarchical scale pitch structure," *Music Perception*, vol. 24, no. 3, pp. 281–296, 2007.
- [103] M. Cheour, P. H. t. Leppänen, and N. Kraus, "Mismatch negativity (MMN) as a tool for investigating auditory discrimination and sensory memory in infants and children," *Clinical Neurophysiology*, vol. 111, no. 1, pp. 4–16, 2000.
- [104] W. T. Fitch and M. D. Martins, "Hierarchical processing in music, language, and action: lashley revisited," *Annals of the New York Academy of Sciences*, vol. 1316, no. 1, pp. 87–104, 2014.
- [105] S. Koelsch, "Music-syntactic processing and auditory memory: similarities and differences between ERAN and MMN," *Psychophysiology*, vol. 46, no. 1, pp. 179–190, 2009.
- [106] J. Driemeyer, J. Boyke, C. Gaser, C. Büchel, A. May, and D. M. Eagleman, "Changes in gray matter induced by learning—revisited," *PLoS ONE*, vol. 3, no. 7, Article ID e2669, 2008.
- [107] S. J. Morrison, S. M. Demorest, E. H. Aylward, S. C. Cramer, and K. R. Maravilla, "fMRI investigation of cross-cultural music comprehension," *NeuroImage*, vol. 20, no. 1, pp. 378–384, 2003.

Research Article

A Voxel-Based Morphometry Study of the Brain of University Students Majoring in Music and Nonmusic Disciplines

Kanako Sato,¹ Eiji Kirino,^{2,3} and Shoji Tanaka¹

¹Department of Information and Communication Sciences, Sophia University, Tokyo 102-8554, Japan

²Department of Psychiatry, Juntendo University School of Medicine, Tokyo 113-8421, Japan

³Juntendo Shizuoka Hospital, Shizuoka 410-2295, Japan

Correspondence should be addressed to Shoji Tanaka; shoji.tanaka@gmail.com

Received 23 January 2015; Revised 28 March 2015; Accepted 29 March 2015

Academic Editor: Shinichi Furuya

Copyright © 2015 Kanako Sato et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The brain changes flexibly due to various experiences during the developmental stages of life. Previous voxel-based morphometry (VBM) studies have shown volumetric differences between musicians and nonmusicians in several brain regions including the superior temporal gyrus, sensorimotor areas, and superior parietal cortex. However, the reported brain regions depend on the study and are not necessarily consistent. By VBM, we investigated the effect of musical training on the brain structure by comparing university students majoring in music with those majoring in nonmusic disciplines. All participants were right-handed healthy Japanese females. We divided the nonmusic students into two groups and therefore examined three groups: music expert (ME), music hobby (MH), and nonmusic (NM) group. VBM showed that the ME group had the largest gray matter volumes in the right inferior frontal gyrus (IFG; BA 44), left middle occipital gyrus (BA 18), and bilateral lingual gyrus. These differences are considered to be caused by neuroplasticity during long and continuous musical training periods because the MH group showed intermediate volumes in these regions.

1. Introduction

The brain is not a static organ, but it changes dynamically throughout an individual's lifespan. During development, the brain changes flexibly with various kinds of experiences. These changes are characterized both structurally and functionally [1, 2]. Structural changes include alterations in the volume of specific gray matter (GM) regions, as well as volume and fiber arrangement in the white matter (WM). Functional changes include alterations in regional activation and functional connectivity between regions. The musician's brain is regarded as an ideal model of brain plasticity, because musicians usually start musical training very early in life and continue for many years into adulthood [1, 2]. Musical training requires complex information processing or functions and includes auditory and somatosensory processing, motor control, attention, working memory, executive functions, and higher-order integrative functions. Playing musical instruments requires multitasking, in which emotional expression is critical. Musical training is also an activity in which

procedural, episodic, and semantic memories and emotion converge. Therefore, many brain regions in distinct domains can be changed with musical training.

Magnetic resonance imaging (MRI) assesses anatomical structures in the GM and WM. Voxel-based morphometry (VBM) uses a voxel-wise analysis method for determining focal differences in volume [3]. Previous VBM studies have found GM volume differences between musicians and nonmusicians. A whole brain VBM study showed that professional musicians (keyboard players) have larger bilateral inferior temporal gyri (ITG), left Heschl's gyrus (HG.L), bilateral precentral gyri, right superior parietal cortex, left inferior frontal gyrus (IFG.L), right medial frontal gyrus, and left anterior cerebellar lobe than amateur musicians and nonmusicians [4]. Increased GM volumes in the HG.R, bilateral superior temporal gyri (STG), left middle temporal gyrus (MTG.L), ITG.R, right posterior cingulate gyrus (PCC.R), right central sulcus, right superior frontal gyrus (SFG.R), and IFG.R in musicians compared with those in nonmusicians were also observed in another study [5]. Other studies have

reported a larger GM volume of the left IFG pars opercularis (BA 44) in male orchestra musicians compared with non-musicians, which positively correlated with years of musical performance [6, 7]. In addition, the left hippocampus, right supplementary motor area (SMA.R), SFG.R, right middle frontal gyrus (MFG.R), PCC.L, right insula, and STG.L were positively correlated with the degree of musical practice [8]. Moreover, another VBM study [9] detected increased GM volumes with increased musical training intensity in the right fusiform gyrus (FG.R), right mid orbital gyrus, IFG.L, left inferior parietal lobule (IPL.L) ($P < 0.001$, uncorrected), and bilateral cerebellum Crus II. An increase in HG.L was also detected, but only at $P < 0.005$ (uncorrected). In contrast, decreased GM volumes were observed in bilateral perirolandic and striatal areas that are related to sensorimotor functions [9]. Thus, the brain regions that differ between musicians and nonmusicians depend on the study and have not been replicated well.

Because of the abovementioned variability in VBM results, corroborative analysis with different sets of participants is needed. Here, we report the results of our VBM analysis of whole brain structural data from university students majoring in music and those majoring in nonmusic disciplines. We chose university students because they are all in the same generation with regard to the cultural environment and exhibit less background variability than professional musicians of a wide range of ages. Students majoring in music studied in the same university as those majoring in other nonmusic disciplines. Our aim was to investigate how education in different disciplines (music versus nonmusic) affects brain structure. Because nonmusic major students still had various degrees of extracurricular music activities, we divided these students into two groups; therefore, we examined three groups in total: music expert (ME) group (music major students), music hobby (MH) group (nonmusic major students having active extracurricular music lessons), and nonmusic (NM) group (nonmusic major students with no or less musical training). Differences between the ME group and NM group suggest a direct effect of musical training. The MH group was included to provide insight into the effect of different degrees of musical training on brain structure.

2. Methods

2.1. Participants. This study was approved by the local ethics committees, the Sophia University Ethics Committee, and the Juntendo University Ethics Committee. All participants provided written informed consent before the study commenced. University students majoring in music ($n = 23$) and nonmusic disciplines ($n = 32$) were recruited. All participants were healthy right-handed Japanese females and had no history of neurological disorders. Students majoring in music started musical training at around 3–5 years of age, which continued to the present. They were screened on enrollment in a college of music. All of them specialized in classical music and played various instruments (piano, violin, cello, clarinet, or trumpet). Students majoring in nonmusic disciplines showed various degrees of extracurricular music

TABLE 1: Gray matter (GM), white matter (WM), and total brain volumes (all in mL) and P values obtained by ANOVA between all three groups. NM: nonmusic; MH: music hobby; ME: music expert.

	NM		MH		ME		P value
	Mean	SD	Mean	SD	Mean	SD	
GM	676.6	46.3	697.9	65.8	685.6	48.1	0.533
WM	441.5	54.8	444.5	41.1	465.8	47.0	0.225
Total	1118.1	82.2	1142.4	83.9	1151.4	79.9	0.469

activities and were therefore divided into two groups: students with active extracurricular music lessons and those with less musical training. Students in the former group had taken music lessons for nine years or longer, whereas those in the latter group had taken lessons for seven years or shorter. Thus, there were three groups of students: ME: music expert group ($n = 23$; age: 18–26 years; mean = 21.2 years), MH: music hobby group ($n = 17$; age: 18–23 years; mean = 20.9 years), and NM: nonmusic group ($n = 15$; age: 19–23 years; mean = 21.6 years).

2.2. Data Acquisition. Structural MRI was performed using a 3T Philips Achieva scanner at Juntendo Hospital, Tokyo, with a MPRAGE sequence (TE = 3.3 ms, TR = 15 ms, TI = 955.5 ms, and flip angle = 15°). MRI included a 3D set of T1 images (voxel size = $1 \times 1 \times 1$ mm; FOV = $180 \times 232 \times 256$ mm). Imaging time was 3 min 31 s. The scan protocol was identical for all participants.

2.3. Analysis. T1 images were preprocessed by VBM using SPM12 (Wellcome Department of Imaging Neuroscience, London, United Kingdom), which segments T1 brain images into GM, WM, and cerebrospinal fluid. GM images were then subjected to spatial normalization to standard MNI space using the DARTEL toolbox (Wellcome Department of Imaging Neuroscience) [10]. Individual voxel volumes were modulated by whole brain volume with proportional scaling. All images were smoothed using the Gaussian kernel of $8 \times 8 \times 8$ mm FWHM. VBM analyzed differences in local GM and WM volumes across the whole brain using two-sample t -tests between ME and NM groups, and analysis of variance (ANOVA) across all three groups.

3. Results

GM, WM, and total brain volumes are summarized in Table 1. The P values obtained by ANOVA of all three groups are also provided. GM, WM, and total volume did not differ between groups. The surface map of brain regions with differing GM volumes is shown in Figure 1. Sagittal and horizontal cross sections are shown in Figures 2 and 3. Brain regions with different volumes between ME and NM groups are summarized in Table 2. The ME group had larger volumes in the bilateral LiG, left MOG (BA 18), and right IFG than in the NM group. The ME group also tended to have larger volumes in the left MOG (BA 19), right STG (BA 22), right anterior insula (AI), and precuneus. These regions showed

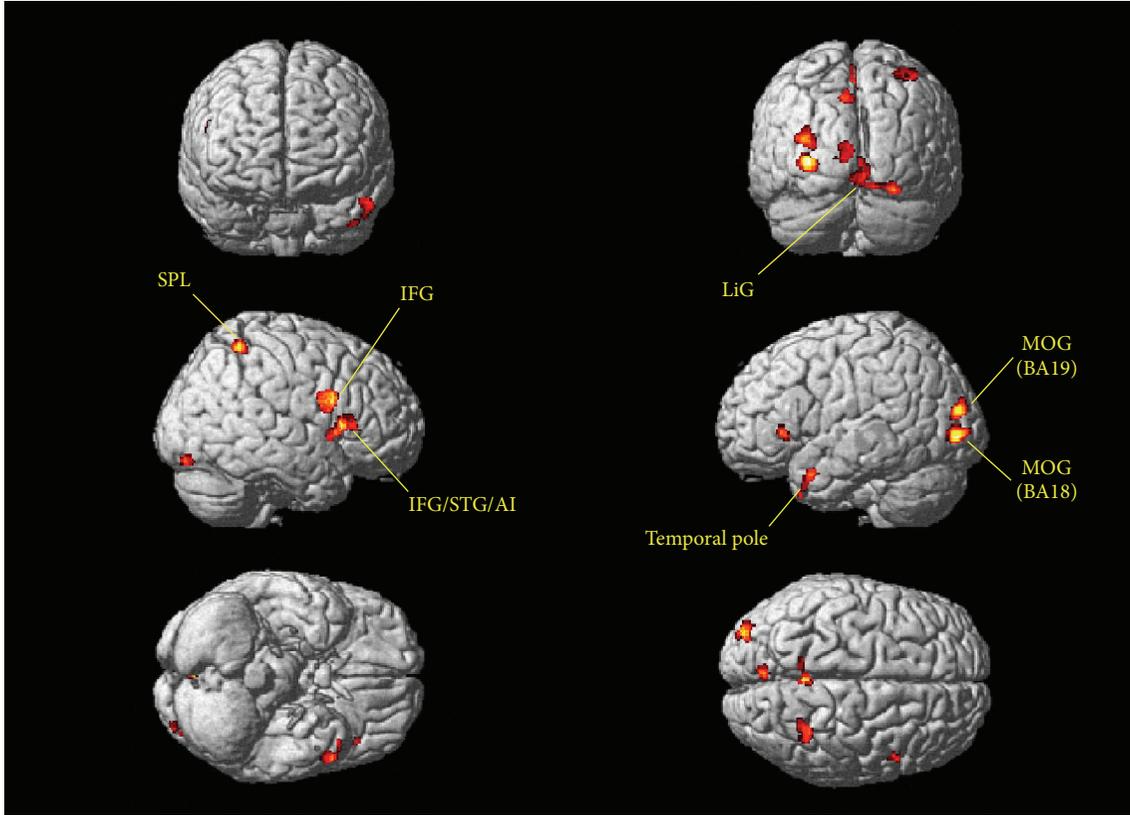


FIGURE 1: Brain regions with larger gray matter volumes in the ME group than in the NM group ($P < 0.01$, uncorrected). Only cluster sizes >100 voxels are shown. No surface regions were smaller in the ME group than in the NM group.

TABLE 2: Brain regions and peak locations (in MNI coordinates) that differed between music expert and nonmusic groups. Brain regions with cluster sizes, $k > 252$ voxels (expected voxels per cluster), are shown. * $P < 0.001$ (uncorrected) for two-sample t -tests; otherwise, $P < 0.01$ (uncorrected).

Brain region	x	y	z	Cluster size (in voxels)	t value
Music experts > Nonmusic					
Left MOG (BA18)	-27	-86	-2	383	4.10*
Right IFG (BA44)	44	11	23	566	4.00*
Bilateral LiG	-2	-77	-6	1389	3.74*
Left MOG (BA19)	-33	-89	17	313	3.34
Right STG (BA22)	53	11	-2		3.25
Right AI	38	24	12	734	3.14
Right AI/IFG	41	17	6		3.00
Precuneus	0	-48	53	373	3.14
Nonmusic > Music experts					
Right caudate nucleus	18	14	9	253	2.87

larger clusters than the threshold size of 252. In contrast, the right caudate nucleus showed a trend to be smaller in the ME group than in the NM group. Comparison of relative GM volumes between the three groups is shown in Figure 4. This figure also included the right SPL (peak MNI: 30, -53, 59)

($k = 242$; $t = 2.87$) and left temporal pole (peak MNI: -44, 17, -42) ($k = 203$; $t = 2.76$), although below the cluster size threshold. In all regions ((a) through (i)), volumes for the MH group were in between the other two groups.

4. Discussion

In our study, the brain regions with the largest volumes in the ME group were higher-order sensory and association areas, which are not inherently music-proper areas. However, our finding is consistent with the fact that musical training requires higher-order cognitive and attentional functions [9, 11]. In the following section, we discuss the information processing required in musical training by associating the functions of the altered regions.

4.1. Inferior Frontal Gyrus. A study found that the left posterior IFG (including Broca's area) was larger in male symphony orchestra musicians [7]. The subjects (both musicians and controls) in this study showed a wide range of ages (26–66 years). Although the volume of Broca's area decreased with age, the decrement was much smaller or nullified in musicians ($P = 0.44$); therefore, the volume differed between groups at >45 years of age. In addition, VBM and deformation-based morphometry (DBM) studies detected significant differences in the right IFG between musicians and nonmusicians [5, 12]. Consistently with these studies, we

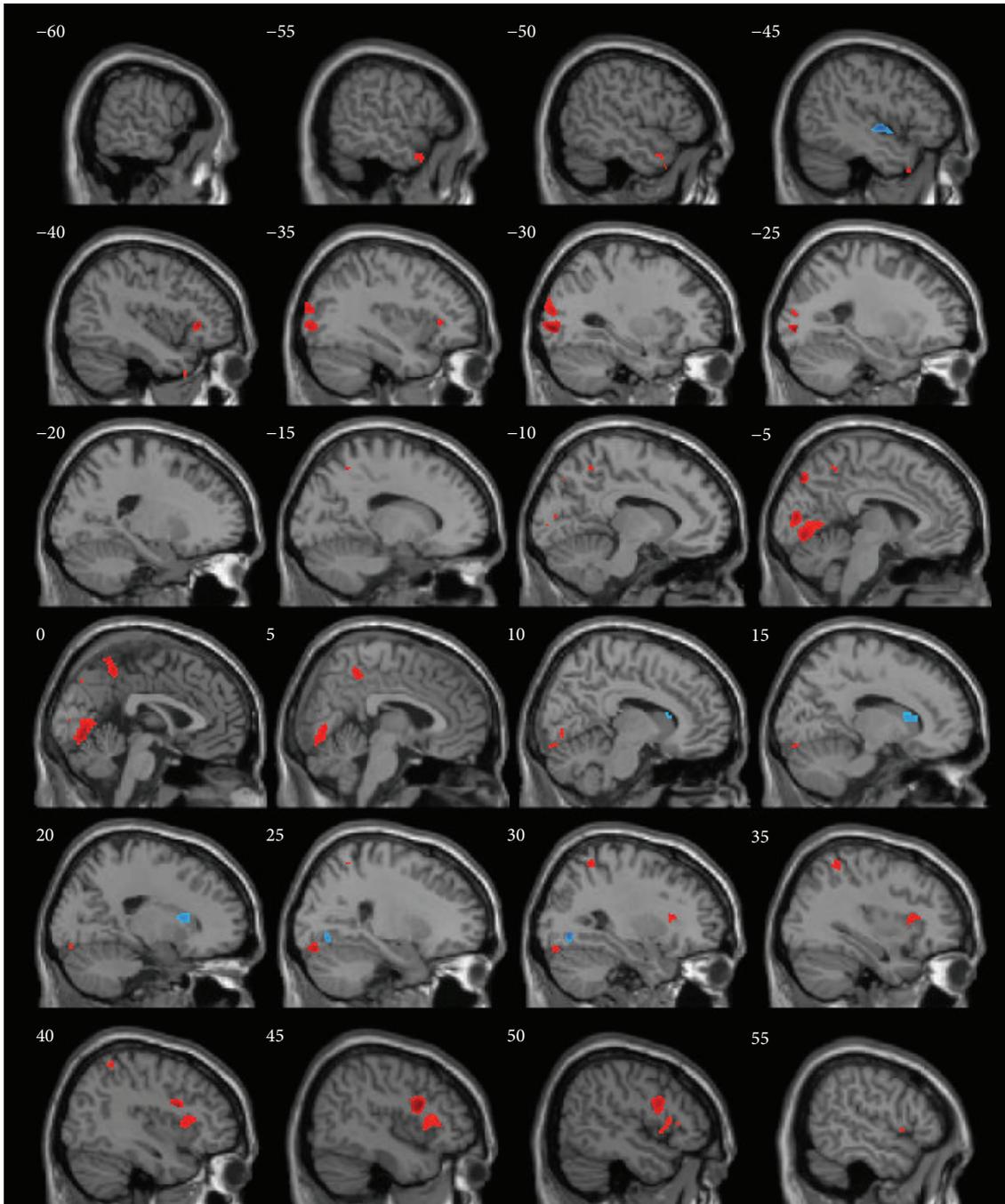


FIGURE 2: Sagittal-plane locations of brain regions with gray matter volumetric differences between ME and NM groups. Only clusters sizes >100 voxels are shown. Regions in red or blue color were larger or smaller, respectively, in the ME group.

observed a volumetric difference in the right posterior IFG. There could be several interpretations for this, as discussed below.

4.1.1. Syntax Processing. The left posterior IFG is crucial for syntax processing in language and music [13, 14]. Although lateralization of the posterior IFG has been recognized for language [15], lateralization for music remains less clear. A cortical network comprising the IFG (BA 44), ventrolateral

premotor cortex (PMC), and anterior STG has been implicated in the processing of musical structure [16], which is also a principal component of the language processing network [17]. Bilateral IFG activation was observed during a harmonic processing task in an fMRI study of nonmusicians with moderate musical training [18]. Furthermore, a magnetoencephalography study showed bilateral BA 44 activation for harmonically inappropriate chords [19]. Interestingly, a chord sequence paradigm activated BA 44 with right-hemispheric

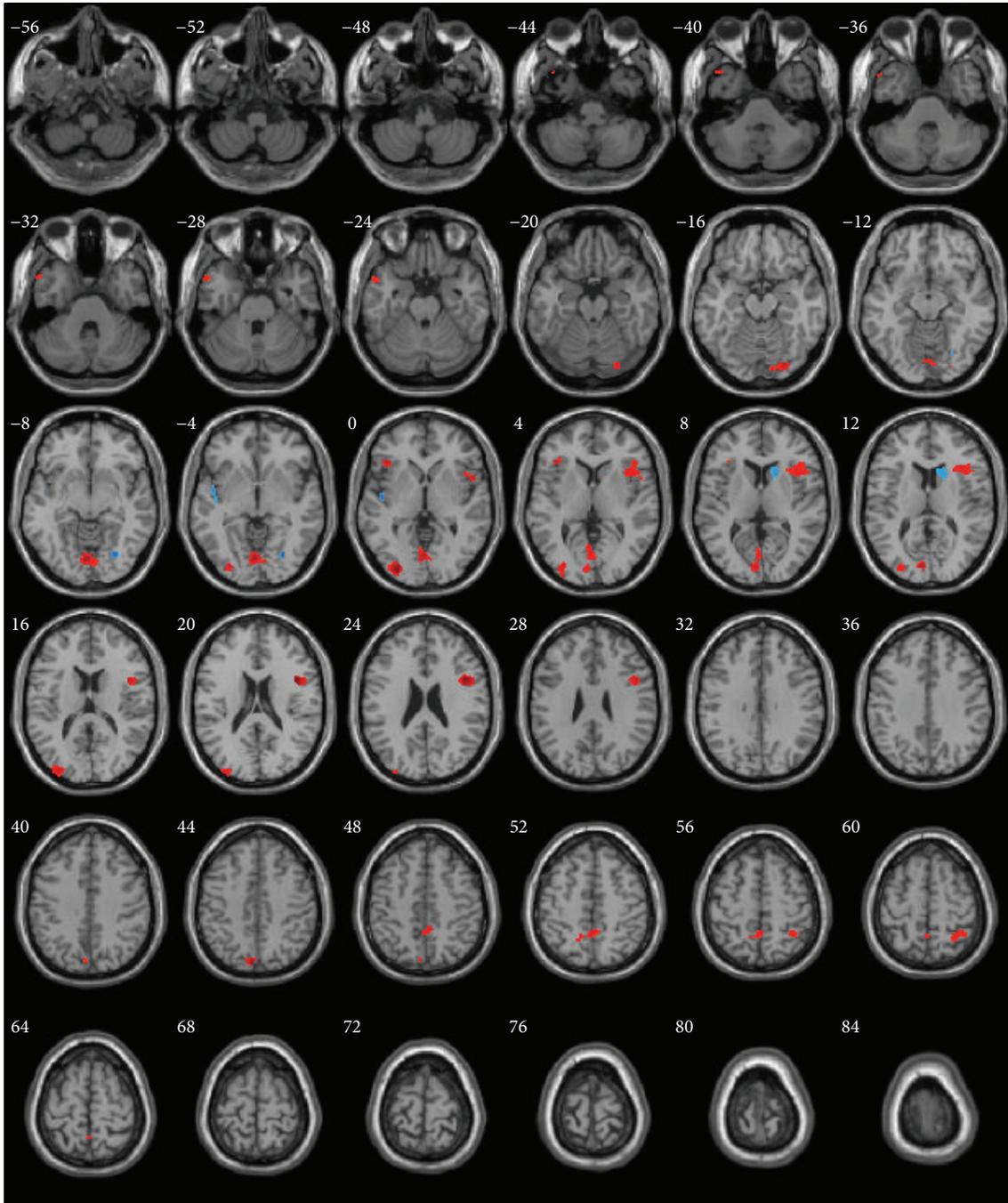


FIGURE 3: Horizontal-plane locations of brain regions with gray matter volumetric differences between ME and NM groups. Only cluster sizes >100 voxels are shown. Regions in red or blue color were larger or smaller, respectively, in the ME group.

weighting [20]. Taken together, these previous studies suggest that music and language use a common network but with different lateralization tendencies, with language favoring the left hemisphere [13, 21]. Our analysis found a difference between the ME and NM groups in the right BA 44, the right-hemisphere homologue of Broca's area. A plausible interpretation of this is that the right posterior IFG is enlarged in the ME group due to the highly demanding syntax processing of musical training. However, this does not

simply imply that left BA 44 processes lingual syntax, while right BA 44 processes musical syntax. There is the possibility that the left BA 44 did not differ because NM students were using language intensively. Consequently, only the right BA 44 would show a significant difference. Therefore, our result suggests that musical training has increased the right posterior IFG volume, potentially due to the involvement of this brain region in music syntax processing, and thereby supporting the previous suggestion.

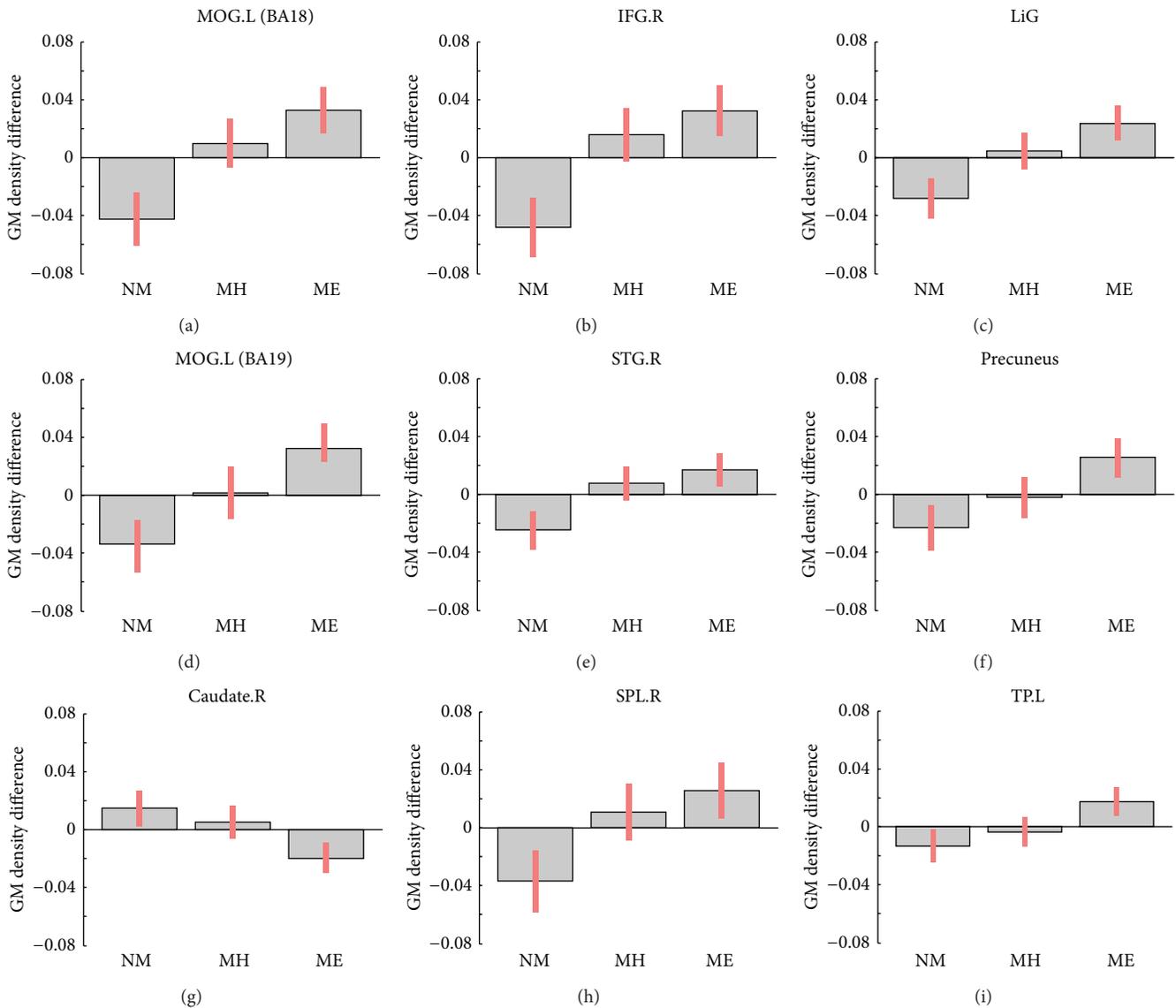


FIGURE 4: Gray matter density differences between nonmusic (NM), music hobby (MH), and music expert (ME) groups at the selected voxel (MNI coordinates are given in Table 2 for (a)–(g) and in the text for (h)–(i)). Vertical red bars indicate 90% confidence intervals. (a) Left middle occipital gyrus (BA 18); (b) right inferior frontal gyrus; (c) lingual gyrus; (d) left middle occipital gyrus (BA 19); (e) right superior temporal gyrus; (f) precuneus; (g) right caudate nucleus; (h) right superior parietal lobule; (i) left temporal pole.

4.1.2. Cognitive Control. An fMRI study using a three-dimensional mental rotation task found that orchestral musicians had significantly increased activity in Broca's area, in addition to the visuospatial network, which was activated in both musicians and age-matched nonmusicians [22]. This finding suggests that Broca's area also contributes to the control of cognitive operations. The IFG (in particular the posterior part) is a node of the frontoparietal mirror neuron system that is suggested to play an important role in playing and listening to music through the observation of others' actions for understanding emotion and intention [23]. This brain region is also activated by the perception of faces with emotion [24]. More broadly, the IFG has been associated with cognitive control of memory [25], response

selection/inhibition and other self-control functions [26], and cognitive flexibility [27]. Working memory is critical in music performance. Verbal and tonal working memory appear to share a common network comprising the IFG, PMC, and inferior parietal lobule, and musicians activate tonal working memory more strongly than nonmusicians [28]. Interestingly, musicians, but not nonmusicians, recruit other brain regions for tonal as well as verbal working memory. For tonal working memory, these regions include the left cuneus, right globus pallidus, right caudate nucleus, and left cerebellum [28]. Working memory and cognitive control in music connect many brain regions, which enables complex multitask performance. Therefore, hypothetically, enlargement of the posterior IFG in the ME group may reflect

inherent involvement of cognitive control in musical training and performance.

4.1.3. Creativity. Another interesting construct that may account for the difference in right BA 44 volume is “creativity.” Both the left and right IFG are correlated with verbal creativity [29]. Moreover, a functional connectivity study suggested that subjects with higher creativity have increased connectivity at rest between BA 44 in both hemispheres and the default mode network (DMN) [30]. They also showed that the right BA 44 has stronger connectivity with the left dorsolateral prefrontal cortex in subjects with higher divergent thinking ability. Therefore, the right BA 44 volume difference observed in our study may reflect, at least partially, a difference in creativity between the groups, although we did not assess creativity in our participants.

4.2. Visual and Visuospatial Areas. The ME group had larger volumes in bilateral LiG, left MOG (BA 18), and left MOG (BA 19) than in the NM group. This finding is novel to the best of our knowledge. These visual association areas have multiple functions that are considered relevant to music. Visual attention ability is enhanced in musicians compared with nonmusicians [31]. Harmonic processing activates different visual association areas for musicians and nonmusicians [32]. Bilateral LiG activation is associated with harmonic processing in musicians, but not naive subjects, suggesting that musical training changes visual representation of harmonies [33]. In a positron emission tomography study, LiG activation was associated with visual imagery of a subjects’ hands playing an electronic piano with their eyes closed [34]. For playing musical instruments, dynamic spatial mapping from visual information is also critical. In this regard, the dorsal visual pathway (including the SPL) is likely important for music performance. Our present study shows that the SPL tends to be larger in the ME group. This is consistent with a previous study suggesting that the SPL mediates spatial mapping of reading music scores to fingers tapping a piano key [35]. This region was also activated in music perception, which requires selective attention [36]. Furthermore, melody processing and sight reading activate the SPL and intraparietal sulcus (IPS) [37, 38]. The bilateral IPS has been implicated as part of a multimodal network for systematic transformation of stimulus information [39]. Specifically, this study shows a positive correlation between bilateral IPS GM volume and relative pitch performance in healthy volunteers with a wide range of musical experiences.

4.3. Other Areas. The right STG (BA 22), or secondary auditory cortex, was larger in the ME group at a trend level. The right STG has been associated with pitch processing [5, 39]. Musical sophistication appears to cause a shift of musical perception from the right to the left hemisphere [40]. Musical memory bilaterally activates the STG [41]. STG activity is also modulated bilaterally by music with an affective tone, but with the right hemisphere weighted [42].

Right AI volume had a trend to be larger in the ME group than in the NM group. AI involvement in music

has been consistently observed [34, 43]. Both singing and speech activate the AI, although singing activates the right AI and speech activates the left [44]. Processing of music-evoked emotions also activates the AI, with increased activation for emotionally mismatched displays compared with emotionally matched displays [45], suggesting that the AI plays an active role in monitoring the consistency of musical emotional information. Furthermore, this area is broadly associated with emotional awareness, interoception, body movement awareness, auditory and visual awareness of the moment, time perception, perceptual decision making, cognitive control, and performance monitoring [43]. All of these functions are relevant to music performance.

The ME group had larger precuneus volumes than the NM group, albeit at a trend level. Along with the posterior cingulate cortex, the precuneus is a core DMN node [46], and its function has been associated with visual imagery, episodic memory, and self-processing [47]. A recent study suggested that mental representation of an auditory scene involves the precuneus [33]. Moreover, scene construction consistently activates the precuneus [48]. Considered along with these findings, our result suggests that mental imagery or scene construction plays an important role in musical training.

We found that the right caudate nucleus was the only region smaller in the ME group than in the NM group. Interestingly, a recent VBM study reported that skilled ballet dancers have smaller motor areas [49]. The caudate nucleus constitutes the cortico-striato-pallido-thalamo-cortical loop for sensorimotor control [50]. There is the possibility that ballet training shapes the loop circuit so that efficient signal transmission through the circuit is enabled with a smaller volume. Another study suggested that perirolandic areas and striatal volumes are reduced by musical training, while several cortical areas such as the IFG, Heschl’s gyrus, IPL, and cerebellum show increased volumes in expert musicians [9]. The reduced areas are sensorimotor areas and may follow the principle of parsimony [9]. Although further study is needed, our result supports this notion.

4.4. Limitations. A major limitation of our study is the sample size. Our analysis found group differences at the level of $P < 0.001$, uncorrected. Increasing the threshold by correcting for multiple comparisons found no differences. Therefore, we employed this threshold level conventionally, according to several previous studies [8, 9, 39, 51]. We divided the nonmusic students into two groups because we could then determine if the volumetric differences were graded depending on musical training intensity. The degree of the differences will be ascertained using a larger sample size.

5. Conclusions

Using VBM, we observed larger GM volumes in the right IFG, left MOG (BA 18), and bilateral LiG in the ME group than in the NM group. Right IFG enlargement may be associated with musical syntax processing or, more generally, cognitive control in musical training and performance. Visual area enlargement in the ME group may be associated with multiple

functions relevant to music including visual attention, harmonic processing, and visual imagery. These volumetric differences are considered to be caused by neuroplasticity during long and continuous musical training periods because the MH group showed intermediate volumes in these regions.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

The authors acknowledge continuous assistance from Toho Gakuen School of Music during the execution of the project in this study. The authors thank Fukuta, Inami, and Hirayama at Juntendo University School of Medicine for their technical assistance in the experiments. This study was supported by the Collaborative Research Grant at Sophia University.

References

- [1] T. F. Münte, E. Altenmüller, and L. Jäncke, "The musician's brain as a model of neuroplasticity," *Nature Reviews Neuroscience*, vol. 3, no. 6, pp. 473–478, 2002.
- [2] S. C. Herholz and R. J. Zatorre, "Musical training as a framework for brain plasticity: behavior, function, and structure," *Neuron*, vol. 76, no. 3, pp. 486–502, 2012.
- [3] J. Ashburner and K. J. Friston, "Voxel-based morphometry—the methods," *NeuroImage*, vol. 11, no. 6, pp. 805–821, 2000.
- [4] C. Gaser and G. Schlaug, "Brain structures differ between musicians and non-musicians," *The Journal of Neuroscience*, vol. 23, no. 27, pp. 9240–9245, 2003.
- [5] P. Bermudez, J. P. Lerch, A. C. Evans, and R. J. Zatorre, "Neuroanatomical correlates of musicianship as revealed by cortical thickness and voxel-based morphometry," *Cerebral Cortex*, vol. 19, no. 7, pp. 1583–1596, 2009.
- [6] I. A. Abdul-Kareem, A. Stancak, L. M. Parkes, and V. Sluming, "Increased gray matter volume of left pars opercularis in male orchestral musicians correlate positively with years of musical performance," *Journal of Magnetic Resonance Imaging*, vol. 33, no. 1, pp. 24–32, 2011.
- [7] V. Sluming, T. Barrick, M. Howard, E. Cezayirli, A. Mayes, and N. Roberts, "Voxel-based morphometry reveals increased gray matter density in Broca's area in male symphony orchestra musicians," *NeuroImage*, vol. 17, no. 3, pp. 1613–1622, 2002.
- [8] M. Groussard, F. Viader, B. Landeau, B. Desgranges, F. Eustache, and H. Platel, "The effects of musical practice on structural plasticity: the dynamics of grey matter changes," *Brain and Cognition*, vol. 90, pp. 174–180, 2014.
- [9] C. E. James, M. S. Oechslin, D. Van De Ville, C.-A. Hauert, C. Descloux, and F. Lazeyras, "Musical training intensity yields opposite effects on grey matter density in cognitive versus sensorimotor networks," *Brain Structure and Function*, vol. 219, no. 1, pp. 353–366, 2014.
- [10] Y. Shigemoto, H. Matsuda, K. Kamiya et al., "In vivo evaluation of gray and white matter volume loss in the parkinsonian variant of multiple system atrophy using SPM8 plus DARTEL for VBM," *NeuroImage: Clinical*, vol. 2, no. 1, pp. 491–496, 2013.
- [11] M. S. Oechslin, D. van de Ville, F. Lazeyras, C.-A. Hauert, and C. E. James, "Degree of musical expertise modulates higher order brain functioning," *Cerebral Cortex*, vol. 23, no. 9, pp. 2213–2224, 2013.
- [12] J. A. Bailey, R. J. Zatorre, and V. B. Penhune, "Early musical training is linked to gray matter structure in the ventral premotor cortex and auditory-motor rhythm synchronization performance," *Journal of Cognitive Neuroscience*, vol. 26, no. 4, pp. 755–767, 2014.
- [13] S. Koelsch, T. C. Gunter, D. Y. V. Cramon, S. Zysset, G. Lohmann, and A. D. Friederici, "Bach speaks: a cortical 'language-network' serves the processing of music," *NeuroImage*, vol. 17, no. 2, pp. 956–966, 2002.
- [14] D. Sammler, S. Koelsch, and A. D. Friederici, "Are left fronto-temporal brain areas a prerequisite for normal music-syntactic processing?" *Cortex*, vol. 47, no. 6, pp. 659–673, 2011.
- [15] C. J. Price, "The anatomy of language: contributions from functional neuroimaging," *Journal of Anatomy*, vol. 197, no. 3, pp. 335–359, 2000.
- [16] S. Koelsch, "Significance of Broca's area and ventral premotor cortex for music-syntactic processing," *Cortex*, vol. 42, no. 4, pp. 518–520, 2006.
- [17] A. D. Friederici, "The brain basis of language processing: from structure to function," *Physiological Reviews*, vol. 91, no. 4, pp. 1357–1392, 2011.
- [18] B. Tillmann, P. Janata, and J. J. Bharucha, "Activation of the inferior frontal cortex in musical priming," *Cognitive Brain Research*, vol. 16, no. 2, pp. 145–161, 2003.
- [19] B. Maess, S. Koelsch, T. C. Gunter, and A. D. Friederici, "Musical syntax is processed in Broca's area: an MEG study," *Nature Neuroscience*, vol. 4, no. 5, pp. 540–545, 2001.
- [20] S. Koelsch and W. A. Siebel, "Towards a neural basis of music perception," *Trends in Cognitive Sciences*, vol. 9, no. 12, pp. 578–584, 2005.
- [21] S. Brown, M. J. Martinez, and L. M. Parsons, "Music and language side by side in the brain: a PET study of the generation of melodies and sentences," *European Journal of Neuroscience*, vol. 23, no. 10, pp. 2791–2803, 2006.
- [22] V. Sluming, J. Brooks, M. Howard, J. J. Downes, and N. Roberts, "Broca's area supports enhanced visuospatial cognition in orchestral musicians," *The Journal of Neuroscience*, vol. 27, no. 14, pp. 3799–3806, 2007.
- [23] I. Molnar-Szakacs and K. Overy, "Music and mirror neurons: from motion to 'emotion,'" *Social Cognitive and Affective Neuroscience*, vol. 1, no. 3, pp. 235–241, 2006.
- [24] P. Johnston, A. Mayes, M. Hughes, and A. W. Young, "Brain networks subserving the evaluation of static and dynamic facial expressions," *Cortex*, vol. 49, no. 9, pp. 2462–2472, 2013.
- [25] D. Badre and A. D. Wagner, "Left ventrolateral prefrontal cortex and the cognitive control of memory," *Neuropsychologia*, vol. 45, no. 13, pp. 2883–2901, 2007.
- [26] V. M. Goghari and A. W. MacDonald III, "The neural basis of cognitive control: response selection and inhibition," *Brain and Cognition*, vol. 71, no. 2, pp. 72–83, 2009.
- [27] D. G. Ghahremani, J. Monterosso, J. D. Jentsch, R. M. Bilder, and R. A. Poldrack, "Neural components underlying behavioral flexibility in human reversal learning," *Cerebral Cortex*, vol. 20, no. 8, pp. 1843–1852, 2010.
- [28] K. Schulze and S. Koelsch, "Working memory for speech and music," *Annals of the New York Academy of Sciences*, vol. 1252, no. 1, pp. 229–236, 2012.

- [29] F. Zhu, Q. Zhang, and J. Qiu, "Relating inter-individual differences in verbal creative thinking to cerebral structures: an optimal voxel-based morphometry study," *PLoS ONE*, vol. 8, no. 11, Article ID e79272, 2013.
- [30] R. E. Beaty, M. Benedek, R. W. Wilkins et al., "Creativity and the default network: a functional connectivity analysis of the creative brain at rest," *Neuropsychologia*, vol. 64, pp. 92–98, 2014.
- [31] A. C. Rodrigues, M. A. Loureiro, and P. Caramelli, "Long-term musical training may improve different forms of visual attention ability," *Brain and Cognition*, vol. 82, no. 3, pp. 229–235, 2013.
- [32] V. J. Schmithorst and S. K. Holland, "The effect of musical training on music processing: a functional magnetic resonance imaging study in humans," *Neuroscience Letters*, vol. 348, no. 2, pp. 65–68, 2003.
- [33] D. Spada, L. Verga, A. Iadanza, M. Tettamanti, and D. Perani, "The auditory scene: an fMRI study on melody and accompaniment in professional pianists," *NeuroImage*, vol. 102, pp. 764–775, 2014.
- [34] L. M. Parsons, J. Sergent, D. A. Hodges, and P. T. Fox, "The brain basis of piano performance," *Neuropsychologia*, vol. 43, no. 2, pp. 199–215, 2005.
- [35] L. Stewart, "A neurocognitive approach to music reading," *Annals of the New York Academy of Sciences*, vol. 1060, pp. 377–386, 2005.
- [36] M. Satoh, K. Takeda, K. Nagata, J. Hatazawa, and S. Kuzuhara, "Activated brain regions in musicians during an ensemble: a PET study," *Cognitive Brain Research*, vol. 12, no. 1, pp. 101–108, 2001.
- [37] R. J. Zatorre, A. R. Halpern, and M. Bouffard, "Mental reversal of imagined melodies: a role for the posterior parietal cortex," *Journal of Cognitive Neuroscience*, vol. 22, no. 4, pp. 775–789, 2010.
- [38] D. Schön, J. L. Anton, M. Roth, and M. Besson, "An fMRI study of music sight-reading," *NeuroReport*, vol. 13, no. 17, pp. 2285–2289, 2002.
- [39] N. E. V. Foster and R. J. Zatorre, "Cortical structure predicts success in performing musical transformation judgments," *NeuroImage*, vol. 53, no. 1, pp. 26–36, 2010.
- [40] T. Ohnishi, H. Matsuda, T. Asada et al., "Functional anatomy of musical perception in musicians," *Cerebral Cortex*, vol. 11, no. 8, pp. 754–760, 2001.
- [41] M. Groussard, G. Rauchs, B. Landeau et al., "The neural substrates of musical memory revealed by fMRI and two semantic tasks," *NeuroImage*, vol. 53, no. 4, pp. 1301–1309, 2010.
- [42] C. Pehrs, L. Deserno, J. Bakels et al., "How music alters a kiss: superior temporal gyrus controls fusiform-amygdalar effective connectivity," *Social Cognitive and Affective Neuroscience*, vol. 9, no. 11, pp. 1770–1778, 2014.
- [43] A. D. Craig, "How do you feel—now? The anterior insula and human awareness," *Nature Reviews Neuroscience*, vol. 10, no. 1, pp. 59–70, 2009.
- [44] H. Ackermann and A. Riecker, "The contribution of the insula to motor aspects of speech production: a review and a hypothesis," *Brain and Language*, vol. 89, no. 2, pp. 320–328, 2004.
- [45] K. Petrin, F. Crabbe, C. Sheridan, and F. E. Pollick, "The music of your emotions: neural substrates involved in detection of emotional correspondence between auditory and visual music actions," *PLoS ONE*, vol. 6, no. 4, Article ID e19165, 2011.
- [46] P. Fransson and G. Marrelec, "The precuneus/posterior cingulate cortex plays a pivotal role in the default mode network: evidence from a partial correlation network analysis," *NeuroImage*, vol. 42, no. 3, pp. 1178–1184, 2008.
- [47] A. E. Cavanna and M. R. Trimble, "The precuneus: a review of its functional anatomy and behavioural correlates," *Brain*, vol. 129, no. 3, pp. 564–583, 2006.
- [48] D. Hassabis and E. A. Maguire, "The construction system of the brain," *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 364, no. 1521, pp. 1263–1271, 2009.
- [49] J. Hänggi, S. Koeneke, L. Bezzola, and L. Jäncke, "Structural neuroplasticity in the sensorimotor network of professional female ballet dancers," *Human Brain Mapping*, vol. 31, no. 8, pp. 1196–1206, 2010.
- [50] Y. Chudasama and T. W. Robbins, "Functions of frontostriatal systems in cognition: comparative neuropsychopharmacological studies in rats, monkeys and humans," *Biological Psychology*, vol. 73, no. 1, pp. 19–38, 2006.
- [51] B. Fauvel, M. Groussard, G. Chételat et al., "Morphological brain plasticity induced by musical expertise is accompanied by modulation of functional connectivity at rest," *NeuroImage*, vol. 90, pp. 179–188, 2014.

Research Article

Melodic Contour Training and Its Effect on Speech in Noise, Consonant Discrimination, and Prosody Perception for Cochlear Implant Recipients

Chi Yhun Lo,^{1,2} Catherine M. McMahon,^{1,2,3} Valerie Looi,⁴ and William F. Thompson^{3,5}

¹Department of Linguistics, Macquarie University, Sydney, NSW 2109, Australia

²HEARing Cooperative Research Centre, Melbourne, VIC 3053, Australia

³ARC Centre of Excellence in Cognition and Its Disorders, Macquarie University, Sydney, NSW 2109, Australia

⁴SCIC Cochlear Implant Program-An RIDBC Service, Sydney, NSW 2109, Australia

⁵Department of Psychology, Macquarie University, Sydney, NSW 2109, Australia

Correspondence should be addressed to Chi Yhun Lo; chi.lo@mq.edu.au

Received 24 April 2015; Revised 24 July 2015; Accepted 29 July 2015

Academic Editor: Kentaro Ono

Copyright © 2015 Chi Yhun Lo et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Cochlear implant (CI) recipients generally have good perception of speech in quiet environments but difficulty perceiving speech in noisy conditions, reduced sensitivity to speech prosody, and difficulty appreciating music. Auditory training has been proposed as a method of improving speech perception for CI recipients, and recent efforts have focussed on the potential benefits of music-based training. This study evaluated two melodic contour training programs and their relative efficacy as measured on a number of speech perception tasks. These melodic contours were simple 5-note sequences formed into 9 contour patterns, such as “rising” or “rising-falling.” One training program controlled difficulty by manipulating interval sizes, the other by note durations. Sixteen adult CI recipients (aged 26–86 years) and twelve normal hearing (NH) adult listeners (aged 21–42 years) were tested on a speech perception battery at baseline and then after 6 weeks of melodic contour training. Results indicated that there were some benefits for speech perception tasks for CI recipients after melodic contour training. Specifically, consonant perception in quiet and question/statement prosody was improved. In comparison, NH listeners performed at ceiling for these tasks. There was no significant difference between the posttraining results for either training program, suggesting that both conferred benefits for training CI recipients to better perceive speech.

1. Introduction

Cochlear implants (CIs) are surgically implanted hearing devices that enable the perception of sound for most persons diagnosed with severe to profound deafness. Designed primarily for the purpose of speech perception, they are generally effective in quiet environments but less effective for perceiving speech in noisy environments [1] or for perceiving prosody [2, 3]. Prosody provides information such as the emotional state of a speaker and is used to transform linguistic content from statements to questions with the use of intonation. These difficulties are, in part, due to the lack of fine-structure processing in CI processing

strategies that utilise temporal envelope, such as the Advance Combination Encoder (ACE) [4, 5]. As oral communication often occurs in the presence of complex and noisy acoustic environments, and prosodic components of speech convey important aspects of expression, these limitations can have a direct impact on social interactions and quality of life outcomes [6].

Advances in our understanding of neuroplasticity and learning capacity have led to interest in formal auditory training, with investigators proposing that it may form a component of comprehensive (re)habilitation [6, 7]. As some studies have demonstrated that normal hearing (NH) musicians are particularly adept listeners under challenging

listening conditions such as noise [8, 9], the incorporation of music as a tool for improving language based tasks is a focus for many studies [9–11].

Using auditory brainstem responses (ABR), Parbery-Clark et al. [12] showed that NH musicians exhibited enhanced neural responses at the phoneme level for stop consonants /ba, da, and ga/, relative to nonmusicians. In addition, Strait and Kraus [13] also found that musicians were both faster and more precise than nonmusicians with encoding voice onset time (VOT) and second formant (F2) transitions, both of which contribute to the perception of stop consonants.

A study by Galvin III et al. [14] trained 11 CI recipients with an adaptive, PC-based melodic contour program for 30 minutes a day, with a time course varying between 1 week and 2 months for each participant. Posttraining results using a Melodic Contour Identification (MCI) task indicated that CI recipients' performance was improved between 15.5 and 45.4 percentage points. Recent research extending from this paradigm has investigated the use of melodic contours as training stimuli for CI recipients to improve speech perception tasks. In a preliminary study described in Patel [11], CI recipients were trained to play melodic contours on a piano keyboard 30 minutes a day, 5 days a week, for 1 month. The training stimuli consisted of 9 patterns of 5-note sequences that varied in the size of the intervals between consecutive tones (as used in [14]). Participants trained with intervals between 1 and 3 semitones, with the hypothesis that practise in this task should develop greater precision for MCI. While preliminary results (from two CI recipients) suggest that melodic contour training may improve intonation prosody and speech in noise perception, more evidence is needed to substantiate this finding. Thus, the present study is motivated by a need to provide additional evidence with a larger sample size and explore the transfer effects that melodic contour training may provide for enhancing the speech perception of CI recipients.

The OPERA hypothesis provides a theoretical framework that suggests why music training may drive perceptual speech gains [15]. These are as follows: overlap: acoustic features relevant to both speech and music are encoded on overlapping brain networks; precision: the requirements for music perception are higher than those for speech; emotion: the musical activity should elicit a strong positive emotion; repetition: the promo promotion of plasticity from repeated engagement of the neural network; attention: the focussed attention toward the task. When these conditions are met, there should be a flow on effect resulting in performance gains for speech perception.

Pitch-based tasks are a focus for many studies measuring music perception [16, 17], and music training programs have been used to improve pitch perception for CI recipients [18]. However, sounds are dynamic and multidimensional by nature, and different forms of music training may affect speech and music perception differentially. Therefore, a wider range of potential benefits, such as speed of processing, should be considered. An analysis of 5000 MIDI melodies suggests that mean note durations are approximately 280 ms [19], and an analysis of 16,000 syllables in American English

suggests that mean syllable utterances are approximately 191 ms in length [20]. Thus the time available to extract cues is generally much shorter in speech than in music. Such a comparison can only be evaluated broadly, as there are many redundant cues that make speech accessible. The perception of various consonants that use VOT contrasts (e.g., the distinction between voiced and unvoiced stops /b/ and /p/) or formant trajectory discrimination (e.g., to identify stops within the voiced class such as /b/ from /g/) also relies on the extraction of cues across very short periods, between 5 and 50 ms, for effective perception [21]. As such, the exploration of shorter (and thus more difficult) note durations may be a mechanism for effective training. An emphasis on speed of processing is a differentiating factor for the present study and allows for an exploration of transferred skills, beyond the typical approach of manipulating pitch to adjust difficulty.

The purpose of the present study was to develop and evaluate two take-home, PC-based melodic contour training programs for CI recipients. The programs were adaptive and differentiated by two types of changes introduced in the stimuli: Interval: the interval size was adjusted and note duration was fixed; and Duration: the note durations were adjusted and interval size was fixed. As Patel's [11] results cannot disentangle effects related to the motor requirement of the piano playing task, we designed a purely perceptual training protocol. A key goal was to explore the transfer of nonlinguistic musical skills to specific aspects of speech perception. Using a baseline and posttraining paradigm, the relative efficacy of the two training programs was compared. It was hypothesised that both training programs should confer speech perception benefits to tasks that utilised pitch contours. Specifically, both programs would enhance speech in noise perception and prosodic cue perception due to improved F0 tracking, while consonant perception would be improved for participants assigned the duration program, due to greater speed of processing of VOT and F2 trajectories. The rationale was based on how short the transition period of VOT and F2 is (approximately 50 ms or less). Hence, identifying F2 is reliant on tracking the pitch over a short duration. As such, improvement in identifying melodic contours with shorter durations may have benefits for consonant perception, providing a specific advantage for consonant stops such as /b, d, g, p, t, k, m, and n/.

2. Materials and Methods

Approval for this study was granted by the Macquarie University Faculty of Human Sciences Human Research Ethics Subcommittee (reference: 5201400348).

2.1. Participants. Sixteen adult postlingually deafened CI recipients (11 female, 5 male) ranging in age from 26 to 86 ($M = 58$, $SD = 15$) and CI experience from 1 to 20 years ($M = 9$, $SD = 7$) participated in the study. All CI recipients used Cochlear Ltd. implants in unilateral, bilateral, or bimodal (with a hearing aid (HA)) configuration and were recruited from the Sydney Cochlear Implant Centre (SCIC). Eligibility required full time use of a CI, and at

TABLE 1: Demographic information for Cochlear implant (CI) recipients.

ID	Age	Gender	CI/HA	Processor	Strategy	Number of electrodes activated	Unilateral/bilateral/bimodal	Number of years implanted	Training program
1	80	Female	L-CI24M R-CI24RE (CA)	L-CP810 R-CP810	ACE	L-16 R-22	Bilateral	20	Interval
2	26	Female	L-CI24RE (ST) R-HA	L-CP810	ACE	L-22	Bimodal	1	Interval
3	66	Male	L-CI422 R-HA	L-CP810	ACE	L-21	Bimodal	2	Duration
4	56	Female	L-CI24M R-CI24RE (CA)	L-CP810 R-CP810	ACE	L-18 R-22	Bilateral	14	Duration
5	35	Female	R-CI24RE (ST)	R-CP810	ACE	R-22	Unilateral	1	Duration
6	61	Male	L-CI24R (ST) R-CI24RE (ST)	L-CP810 R-CP810	ACE	L-22 R-14	Bilateral	12	Duration
7	47	Female	L-CI24RE (CA) R-CI24R (ST)	L-CP810 R-CP810	ACE	L-22 R-20	Bilateral	10	Interval
8	86	Female	L-CI24RE (CA) R-CI24RE (ST)	L-CP810 R-CP810	ACE	L-22 R-18	Bilateral	8	Interval
9	52	Female	L-CI24RE (CA) R-CI24RE (CA)	L-Freedom R-Freedom	ACE	L-22 R-18	Bilateral	10	Interval
10	54	Male	L-HA R-CI422	R-CP810	ACE	R-21	Bimodal	2	Duration
11	48	Male	R-CI512	R-CP810	ACE	R-22	Unilateral	4	Interval
12	69	Female	L-CI24RE R-CI24M	L-CP910 R-Freedom	ACE	L-21 R-22	Bilateral	15	Interval
13	66	Female	L-CI512 R-CI24RE (CA)	L-CP910 R-CP810	ACE	L-19 R-22	Bilateral	18	Duration
14	60	Male	L-CI24RE (CA) R-HA	L-CP810	ACE	L-22	Unilateral	2	Duration
15	67	Female	L-CI24RE (CA) R-CI24M	L-CP810 R-CP810	ACE	L-22 R-20	Bilateral	15	Interval
16	55	Female	L-CI422 R-CI22	L-CP900 R-Freedom	ACE	L-22 R-15	Bilateral	19	Duration

least 6-month CI experience. For performance reference purposes, 12 NH adults (6 female, 6 male) ranging in age from 21 to 42 ($M = 27$ years) were recruited from Macquarie University. All NH adults had hearing thresholds ≤ 30 dB hearing level (HL) measured in octave steps between 500 and 4,000 Hz, tested in a sound proof room. All participants were native Australian English speakers and did not have a significant impairment (such as a learning or cognitive disorder) that affected their ability for testing or training. Relevant demographic information can be found in Table 1.

2.2. Melodic Contour Training Program (MCTP). Two take-home PC-based training programs were created: MCTP (Interval) and MCTP (Duration). The training paradigm was adaptive with the stimuli becoming more difficult after a correct response and easier after every incorrect response. The program began at the easiest setting, and each change in difficulty was modulated by 1 level (one-up, one-down procedure), with a total of 7 levels of difficulty. The task was to identify a randomly selected melodic contour, at the designated difficulty level, using a four-alternative

forced choice paradigm (4AFC). The melodic contours were sequences of 5 consecutive notes that formed a total of 9 patterns as used in Galvin III et al. [14]. Figure 1 shows the melodic contours used in the training programs.

The two programs differed by how difficulty was controlled. In the MCTP (Interval), note duration was fixed at 250 ms, and the interval size between consecutive notes was manipulated between 1 and 7 semitones that increased or decreased by 1 semitone. In the MCTP (Duration), interval size was fixed at 5 semitones, while the duration of each note was manipulated between 7 durations: 450, 350, 250, 200, 150, 100, and 50 ms. The lowest note in all stimuli was A4 for both programs; these are marked in light grey in Figure 1. The F0 range for the MCTP (Interval) was 440 to 2218 Hz, and the MCTP (Duration) was 440 to 1397 Hz. The stimuli were created using a Yamaha Disklavier Pro, providing a fairly realistic MIDI representation of an acoustic grand piano.

The program had two modes: “Practice” and “Training.” In Practice, participants were provided with all 9 melodic contours on their screen, and it was designed so that participants could practise listening to (and seeing) all 9 melodic

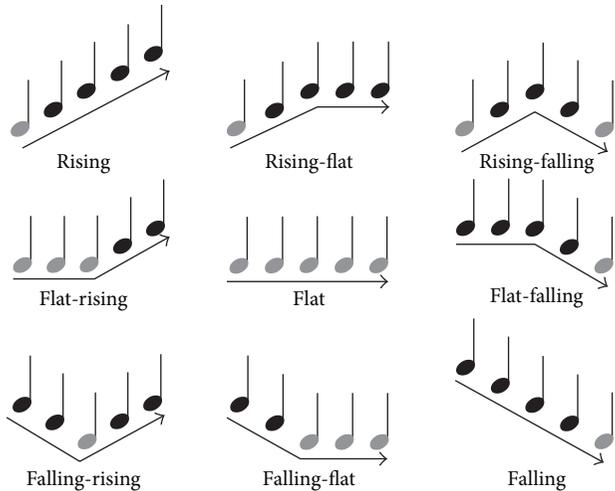


FIGURE 1: The 9 melodic contours used in the Melodic Contour Training Program. The lowest notes are marked in grey. From Galvin III et al. [14].

contours available. The main task was the Training mode in which participants were presented with a melodic contour sound stimulus (which they could repeat), and four buttons representing answers, with one correct answer matching the presented contour, and three other options that were randomly selected from the pool of 9 contours. Feedback was provided after each response. If they were incorrect, the correct response would be highlighted, and they were then permitted (and encouraged) to listen for the differences between their selected and correct responses.

Data logging tracked the progress of each participant's session. For the MCTP (Interval), a melodic contour interval threshold was calculated, the interval size (measured in semitones) at which 50% of contours were correctly perceived. Similarly, for the MCTP (Duration), a melodic contour duration threshold was calculated. The thresholds for each session were averaged across each week of training.

2.3. Materials. The Australian Sentences Test in Noise (AuSTIN) is an adaptive speech in noise test developed specifically for Australian CI recipients [22]. Sixteen sentences were randomly selected and spoken by a female speaker in the presence of time-locked four-talker babble (4TB). In each session, two lists were completed, and a speech reception threshold (SRT, the signal to noise ratio at which 50% of words were correctly perceived) was calculated.

A short Consonant Discrimination Test was developed for the purposes of this study, using a set of 12 commonly used consonants /pa, ta, ka, ba, da, ga, fa, va, sa, za, ma, and na/. The speech materials consisted of one male speaker and were validated for clarity and level-balance by two professional linguists. Lists consisting of 60 consonants in random order were created in two conditions: quiet and noise with 4TB (10 dB SNR). Spectrograms for voiced stop consonants are presented in Figure 2, highlighting F2 as the primary contrastive feature.

An individual subtest (turn-end reception) was selected from the Profiling Elements of Prosody in Speech-Communication (PEPS-C) [23], as a means to assess simple question and statement prosodic discrimination. Participants were presented with 16 single word utterances such as "carrot" or "milk" spoken by a female speaker that varied with intonation. Rising intonations indicated questions, while falling intonations indicated statements. With a 2AFC paradigm, participants were asked to indicate if the utterance was perceived as a question or a statement.

2.4. Procedures. Testing occurred in an acoustically treated test booth in the Macquarie University Speech and Hearing Clinic and in an acoustically treated room at SCIC, Gosford, NSW. The test battery was administered using a Toshiba Tecra R850 laptop. A Yamaha Audiogram 3 USB audio interface provided the sound signal and was connected to a Behringer Truth B3030A loudspeaker. Stimuli were presented at 65 dBA as measured with a sound level metre from the participant's listening position, located 1 metre in front of the loudspeaker. CI recipients were asked to use their regular, everyday settings and adjust their volume to a comfortable sound level on their Cochlear device and hearing aid. Once set, participants were requested to refrain from modifying any settings.

Following the baseline battery, participants were randomly assigned either the Interval or Duration program for the MCTP and provided instructions. There was an equal distribution of participants in each program. The training required the completion of one set of the "Training mode" (25 melodic contours, requiring approximately 15 to 30 minutes, depending on the participants' ability), 4 days a week, for a total duration of 6 weeks. All participants were provided with a set of Edifier M1250 USB powered loudspeakers to use during their training and instructed to train with their regular, everyday settings. Progress was monitored at 2 and 4 weeks, with contact through phone calls and email.

2.5. Statistical Methods. Analysis was performed with IBM SPSS Statistics version 21. Unless stated otherwise, each test was analysed using a repeated measures analysis of variance (ANOVA), with session (baseline and posttraining) as the within-subject factor and program (Interval or Duration) as the between-group factor. Additionally, the posttraining scores were compared between the CI group and the NH group using independent sample *t*-tests. All statistical tests used a criterion of 0.05 and all tests were 2-tailed.

3. Results

Group means and statistical data have been tabulated and are presented in Table 2.

While participants were randomly assigned a training program, to confirm there were no statistically significant differences in key variables between those assigned the Interval program compared with the Duration program, independent sample *t*-tests were calculated across age, CI experience, and all baseline scores. There were no statistically significant differences found; therefore the two groups were considered broadly equivalent prior to the training program.

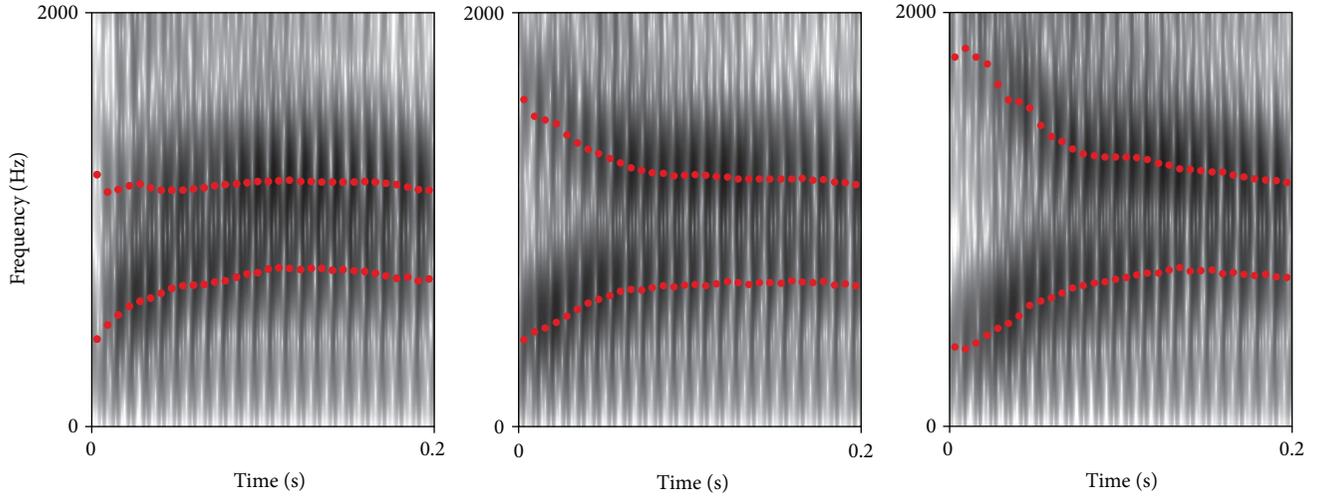


FIGURE 2: Spectrograms for voiced stop consonants /ba, da, and ga/ with F1 and F2 labelled.

TABLE 2: Main effects of session, program, and interactions for all tests.

Test	<i>t</i> or <i>F</i> (df)	<i>p</i>
MCTP (Interval)		
Session	2.75 (6)	0.033*
MCTP (Duration)		
Session	3.35 (7)	0.012*
AuSTIN		
Session	2.46 (1, 14)	0.139
Program	0.01 (1, 14)	0.925
Session/program	0.01 (1, 14)	0.914
Consonant discrimination (quiet)		
Session	6.00 (1, 14)	0.028*
Program	0.03 (1, 14)	0.868
Session/program	2.69 (1, 14)	0.123
Consonant discrimination (4TB)		
Session	0.48 (1, 14)	0.500
Program	0.08 (1, 14)	0.779
Session/program	0.62 (1, 14)	0.444
PEPS-C		
Session	9.31 (1, 14)	0.009*
Program	0.01 (1, 14)	0.978
Session/program	0.90 (1, 14)	0.359

*Indicates Significance at alpha = 0.05.

Compliance was high, with data-logged results indicating that 13 participants completed the full 6 weeks of training. Additionally, there were no drop-outs. Two participants (1 and 7) were inconsistent, completing 4 weeks of the required training, but did compensate with extra sessions in those weeks trained. As such, their data was still included in the analyses. Unfortunately, the data-log recording training performance was corrupted and thus unavailable for participant 9. In summary, performance in the training programs was analysed for 15 participants (excluding participant 9),

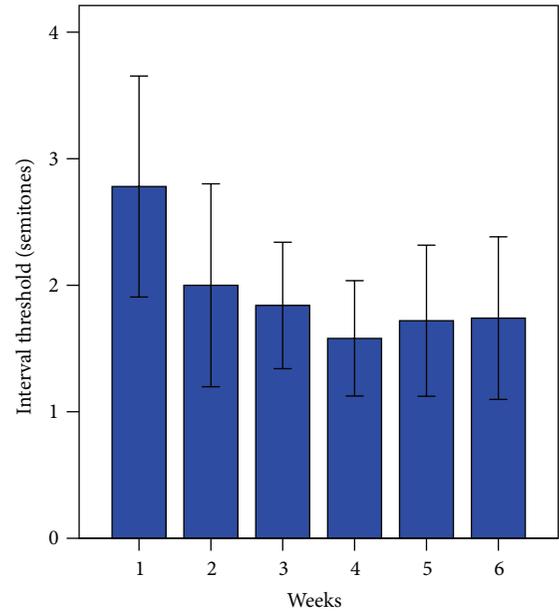


FIGURE 3: Week-to-week interval threshold scores for the Melodic Contour Training Program (Interval group). Error bars indicate 1 standard error.

while baseline and posttraining speech perception measures included all 16 participants.

3.1. Melodic Contour Training Program (Interval). Figure 3 shows the mean interval threshold (semitones) for each week of training. Using paired *t*-tests, the posttraining session threshold (measured at week 6, $M = 1.7 \pm 1.2$ semitones) was significantly better compared with baseline (measured at week 1, $M = 2.5 \pm 1.7$ semitones), $t(6) = 2.75$, $p = 0.033$, indicating that CI recipients were able to identify melodic contours with smaller interval sizes at posttraining than at baseline, with the greatest improvement found at week 4.

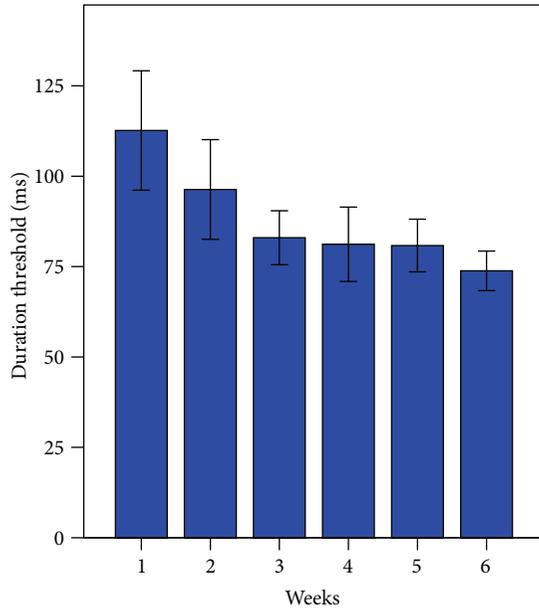


FIGURE 4: Week-to-week duration threshold scores for the Melodic Contour Training Program (Duration group). Error bars indicate 1 standard error.

3.2. *Melodic Contour Training Program (Duration)*. Figure 4 shows the mean duration threshold (ms) for each week of training. Using paired t -tests, the posttraining session threshold ($M = 79 \pm 23$ ms) was significantly better compared with baseline ($M = 115 \pm 39$ ms), $t(7) = 3.35$, $p = 0.012$. These results indicate that CI recipients were able to identify melodic contours with shorter note durations at posttraining than at baseline. Ceiling performance was observed in 3 participants.

3.3. *Australian Sentence Test in Noise (AuSTIN)*. Figure 5 shows the mean SRTs for baseline and posttraining on speech in noise. The main effect of session was nonsignificant [$F(1, 14) = 2.46$, $p = 0.139$], the main effect of program was nonsignificant [$F(1, 14) = 0.01$, $p = 0.925$], and there were no interaction effects [$F(1, 14) = 0.01$, $p = 0.914$]. SRT scores at the posttraining session showed that the CI group was significantly higher ($M = 4.4 \pm 2.2$ dB) compared with the NH group ($M = -4 \pm 0.9$ dB), $t(26) = 11.85$, $p < 0.001$.

3.4. *Consonant Discrimination in Quiet*. The main effect of session was statistically significant [$F(1, 14) = 6.00$, $p = 0.028$], the main effect of program was nonsignificant [$F(1, 14) = 0.03$, $p = 0.868$], and there were no interaction effects [$F(1, 14) = 2.69$, $p = 0.123$]. Consonant scores in quiet at the posttraining session showed that the CI group was significantly lower ($M = 87 \pm 15\%$) compared with the NH group, with all NH individuals performing at ceiling ($M = 100\%$), $t(26) = -3.58$, $p = 0.003$. Figure 6 shows the mean scores (percent correct) for baseline and posttraining for consonant discrimination in quiet.

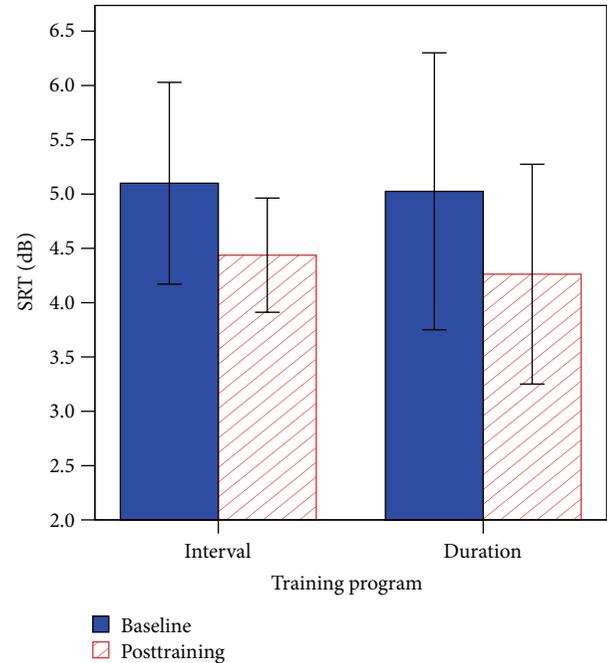


FIGURE 5: Baseline and posttraining SRTs for AuSTIN. Error bars indicate 1 standard error.

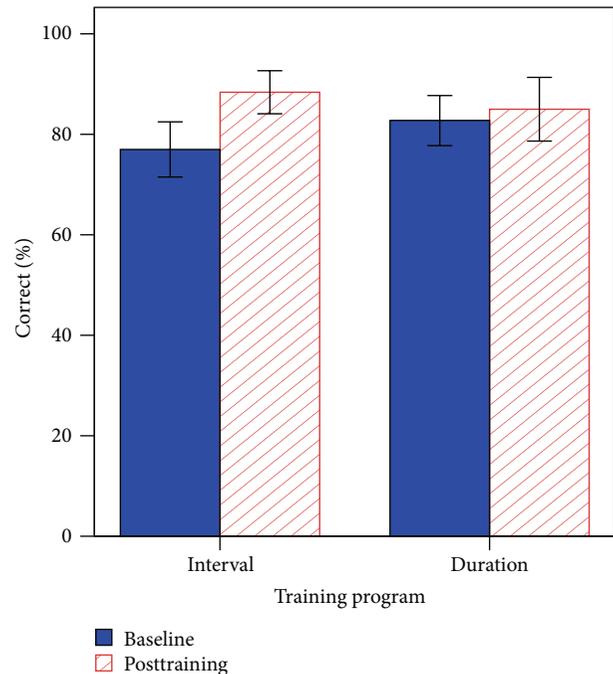


FIGURE 6: Baseline and posttraining performance for consonant discrimination in quiet. Error bars indicate 1 standard error.

Further analysis using confusion matrices for individual consonants revealed that perceiving place of articulation was most improved for both training programs. To reconcile the analysis, only confusions greater than 10% (5 or more confusions) at baseline were considered.

		Response													
		p	t	k	b	d	g	f	v	s	z	m	n	—	
Stimulus	p	23		11			1							5	
	t		40												
	k			30	1			1						8	
	b				37	2		1							
	d			1	4	22	12		1						
	g			1		4	35								
	f	2			1			34		1	1			1	
	v				5					28			3	4	
	s							3			29	8			
	z						3		1			36			
	m												31	3	
	n					3							9	25	

		Response													
		p	t	k	b	d	g	f	v	s	z	m	n	—	
Stimulus	p	35	3	2											
	t		35	5											
	k			36										4	
	b				34	2				4					
	d				2	31	7								
	g				1		39								
	f	1			3			36							
	v		1		1					32		2	2	2	
	s							2			35	3			
	z								1			39			
	m												35	2	
	n												2	38	

FIGURE 7: Confusion matrix for Interval group. Baseline is on top, posttraining at the bottom. Significant confusions in baseline have been marked in grey, and this is carried over to the posttraining matrix for easier visual identification of confusion decreases.

In the Interval group, analysis of individual consonants showed large improvements in the perception of stop consonants in which a 30% increase in accuracy was observed for /p/, a 23% increase for /d/, and an increase of 33% for /n/. A large reduction of confusions was observed for stop consonants, in which a 23% decrease was observed for /p/ perceived as /k/ and a 13% decrease was observed for /d/ perceived as /g/, in fricatives a 13% decrease was observed for /s/ perceived as /z/, and in the nasal stop an 18% decrease was observed for /m/ perceived as /n/. Pooled confusion matrices at baseline and posttraining for the Interval group are presented in Figure 7.

In the Duration group, analysis of individual consonants showed large improvements in the perception of stop consonants in which a 25% increase in accuracy was observed for /p/, a 33% increase was observed for /n/, and a 25% increase was observed for the fricative /v/. A large reduction of confusions was observed for stop consonants, in which a 13% decrease was observed for /g/ perceived as /d/, in fricatives an 18% decrease was observed for /v/ perceived as /m/, and in the nasal stop a 13% decrease was observed for /m/ perceived as /n/. Pooled confusion matrices at baseline and posttraining for the Duration group are presented in Figure 8.

3.5. Consonant Discrimination with 4TB. The main effect of session was nonsignificant [$F(1, 13) = 0.48, p = 0.500$],

		Response													
		p	t	k	b	d	g	f	v	s	z	m	n	—	
Stimulus	p	18	3	9		1	1	1						7	
	t		39	1											
	k	5		35											
	b	1			33	3			2		1				
	d					32	8								
	g			1	2	5	31							1	
	f	1		1				34	1	3					
	v			1		1		2	22		7	7			
	s										38	2			
	z											40			
	m												26	11	
	n												2	38	

		Response													
		p	t	k	b	d	g	f	v	s	z	m	n	—	
Stimulus	p	28	1	6										5	
	t		39	1											
	k	6		30										4	
	b	1			34	2	1			2					
	d				1	34	5								
	g			5		0	35								
	f	1		2		1	30	2	4						
	v							1	35		4	0			
	s							4		36					
	z									4	36				
	m											32	6	1	
	n											1	39		

FIGURE 8: Confusion matrix for Duration group. Baseline is on top, posttraining at the bottom. Significant confusions in baseline have been marked in grey, and this is carried over to the posttraining matrix for easier visual identification of confusion decreases.

the main effect of program was nonsignificant [$F(1, 13) = 0.08, p = 0.779$], and there were no interaction effects [$F(1, 13) = 0.62, p = 0.444$]. Consonant scores with 4TB at the posttraining session showed that the CI group was significantly lower ($M = 63 \pm 16\%$) compared with the NH group, with all NH individuals performing near ceiling performance ($M = 99 \pm 1\%$), $t(26) = -9.08, p < 0.001$. In the baseline session, participant 12 did not complete the task citing difficulty perceiving any consonants in noise. However, in the posttraining session after completion of training, the participant was able to complete the task, scoring 57% correct. Figure 9 shows the mean scores (percent correct) for baseline and posttraining on consonant perception, with participant 12 included.

3.6. Profiling Elements of Prosody in Speech-Communication (Turn-End Reception). Figure 10 shows the mean percent correct for baseline and posttraining on question/statement prosody. The main effect of session was statistically significant [$F(1, 14) = 9.31, p = 0.009$], the main effect of program was nonsignificant [$F(1, 14) = 0.01, p = 0.978$], and there were no interaction effects [$F(1, 14) = 0.90, p = 0.359$]. Prosody scores at the posttraining session showed that the CI group was significantly lower ($84 \pm 18\%$) compared with the NH group, with all NH individuals performing at ceiling (100%)

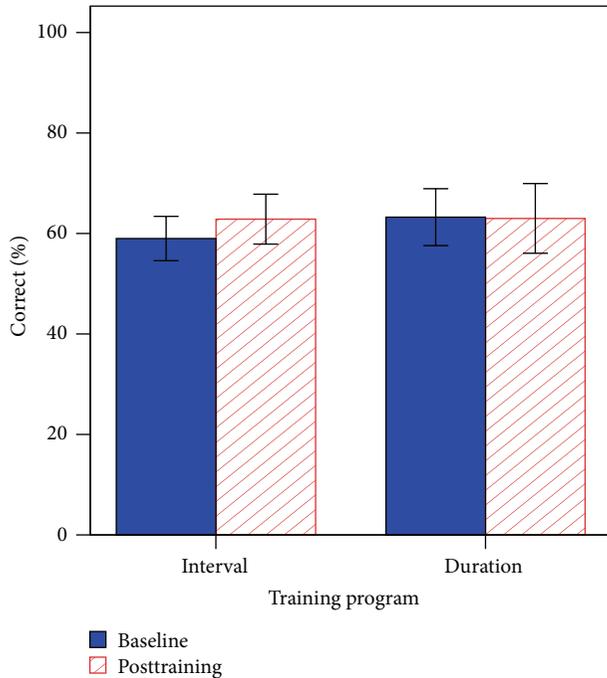


FIGURE 9: Baseline and posttraining performance for the consonant discrimination with 4TB. Error bars indicate 1 standard error.

accuracy), $t(26) = -3.42$, $p = 0.004$. These results indicate a significant posttraining improvement for prosody perception using intonation cues.

4. Discussion

The results indicate that melodic contour training can significantly improve some, but not all, aspects of speech perception in CI recipients. In particular, significant improvements for the perception of consonants in quiet and for the identification of questions and statements using only speech intonation cues were observed. Despite this, there were no significant group gains for speech in noise perception, or consonant perception in 4TB. Finally, and as expected, CI recipients performed more poorly than NH listeners in all tasks at pre- and posttraining measures.

Data-logged results from CI recipients indicate that MCI performance was significantly improved after six weeks of training in both Interval and Duration programs. However, for all tests, there was no significant effect for the type of program assigned to each participant. The greatest improvement was seen from week 1 to week 2 for both training programs, which may be an effect of familiarisation with the program. Maximum improvement with respect to interval and duration threshold was observed at weeks 4 and 6, respectively.

On all tests of speech perception there was no statistical difference between either of the training programs. These findings indicate that CI recipients were able to improve their pitch perception and temporal processing abilities in the context of MCI. While the relative efficacy between

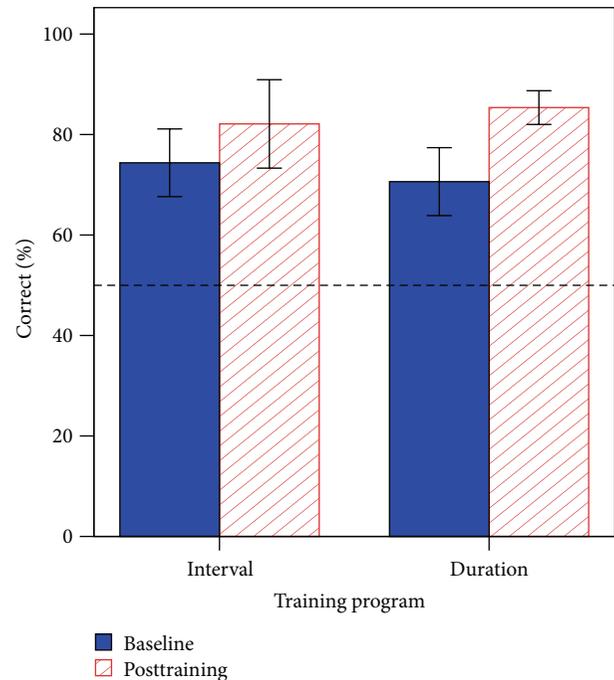


FIGURE 10: Baseline and posttraining performance for the PEPS-C (turn-end reception task). The dashed line indicates the chance score. Error bars indicate 1 standard error.

both mechanisms of interval size and note duration was nonsignificant, comparisons were difficult to make due to the small sample size, resulting in a lack of statistical power.

As the two training programs used significantly different musical mechanisms, it was surprising that the improvement in consonant perception in quiet had similar patterns for both training groups. In particular, confusions between place of articulation cues in voiced, unvoiced, and nasal stops were the most reduced, despite these cues being typified as the poorest speech production feature for CI recipients [24], which is an encouraging finding. As most improvement was found for stop consonant discrimination and as the F2 trajectory is the primary cue contrast, it is likely that recipients were better able to track F2 after training.

Both groups also showed significant improvement for the question-statement task that required cues of speech intonation. Firstly, it must be noted that the stimuli were single words consisting of one or two syllables, and the intonation pattern occurred over the final (or only) syllable. As such, there were no syntactic or semantic cues available and the improvement from training is most likely due to the mechanism of enhanced F0 tracking. However, it is possible that recipients also used duration and intensity cues across syllable boundaries as a distinction. Additionally, as question utterances rarely consist of just one word, the applicability of this enhancement to a more realistic question-statement identification task such as that with sentences, or in adverse conditions, is limited.

Based on preliminary results by Patel [11] that indicated the possibility of improvement for speech in noise perception

as a result of melodic contour training, similar gains were anticipated for the current study. On the other hand, our findings indicate that, as a group, there was no significant improvement for consonant perception in noise, or with the perception of sentences in noise. Despite this, certain individuals showed large improvement in SRTs, although these were both bimodal listeners using a contralateral HA. This suggests that HA users, with more access to acoustic, F0, and fine-structure cues, may find melodic contour training particularly effective for speech in noise improvement. Aside from presenting data from a larger sample of participants, a key difference between Patel [11] and the present study was the removal of piano playing as the training paradigm. As such, the improvements found in the present study are inherently perceptual, as sensory-motor interactions (through the mapping of finger movements and musical notes) were not explicitly trained.

While our findings indicate some level of F0 improvement, primarily for intonation, such enhancement is only accessible in quiet, indicating that maskers significantly disrupt F0 cues for CI recipients that only have access to gross temporal envelope. Effective speech in noise perception is also reliant on auditory stream segregation processes to perceptually group and separate multiple sources [25]. As the melodic contours were a single-stream melody, it is unlikely that it would confer any benefit for segregation tasks.

The OPERA hypothesis suggests that music-driven speech gains are likely dependent on the type of training stimuli itself. Our results indicate that improvement to MCI, with an emphasis on pitch through the Interval program, and speed of processing with the Duration program both provide cues that transfer to more effective perception of stop consonants and speech intonation. As such, a training program manipulating both pitch and speed of processing difficulty may yield even greater improvement.

While there were overall group improvements for both training programs, there was considerable variation among individual participants, a common finding for CI studies. Ten of our participants were bilaterally implanted, and 3 participants were bimodal users. Two-ear listening allows for a binaural advantage, primarily improving spatially separated speech in noise tasks that require access to interaural cues to enhance localisation and segregation ability, relative to a unilateral CI [26]. On the other hand, as each of our speech perception tasks was delivered via one loudspeaker located at 0-degree azimuth, the main benefit of binaural devices was negated. Another benefit is binaural redundancy, whereby two ears (and binaural processing within the auditory system) integrate cues into a more salient speech signal, providing a small advantage of about 1 to 2 dB that may improve speech perception in adverse listening conditions [27]. It was not a main objective to evaluate differences between unilateral, bilateral, and bimodal configurations in this study, but these advantages should be noted. However, to maximise statistical power and generalisability, the inclusion criteria were extended to include all of these configurations, and we assumed that the difference between these groups would be nonsignificant for the measures evaluated, in a repeated measures design. Future studies could assess

the effect of unilateral, bilateral, and bimodal devices on training efficacy.

The program had two tasks: Practice and Training, but data logging was only taken in the training mode. As such, the week-to-week improvements can only be interpreted broadly, as it is impossible to determine how much practice an individual completed. Additionally, participants were required to do at least 4 training sessions a week but were not discouraged from doing more. Nonetheless, irrespective of the rate of improvement, there were significant gains from baseline to posttraining.

This study was limited by a small sample, reducing the ability to evaluate subtle differences in the benefits of the two training protocols. Additionally, more robust baseline measures should be adopted ensuring stable asymptotic performance prior to training, such as introducing two or more spaced sessions prior to the training, as well as follow-up testing without training to ascertain if improvements are retained. Interpretation of cues is also made difficult without objective measures as complimentary evidence. The CI and NH groups were not age-matched, as the purpose was to provide a broad comparison across the speech tasks between the groups. However, it may be of interest to evaluate whether melodic contour training may improve older NH listeners' speech perception in noise. Certainly, cognitive abilities decline with age [28], and several studies show that music training is correlated with increased cognitive abilities [29]. Therefore it is possible that greater gains in speech perception might be found in older adults through improvements in cognitive ability.

5. Conclusion

In conclusion, the findings suggest that both musical mechanisms (intervals and durations) have had a beneficial outcome for CI recipients' perception of transition cues in quiet. These cues are most relevant for stop consonant distinctions and speech intonation, both of which derive the most advantage from melodic contour training. Masking effects, such as noise, significantly disrupt access to these cues, reducing the efficacy of melodic contour training in adverse listening situations.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This study was supported by the Macquarie University Research Excellence Scheme and the HEARing CRC, established and supported under the Cooperative Research Centres Program, Business Australia. The authors acknowledge and thank all participants for their time and effort. Many thanks are due to Ian Deng for programming assistance, Brett Swanson and Cochlear Ltd. for technical assistance with the AuSTIN materials, Felicity Cox for assistance with confusion

matrices, and Jane Brew, Samantha Barrett, and the SCIC Cochlear Implant Program for their considerable recruitment efforts. The authors are grateful to Mariana Reis and two anonymous reviewers for their constructive and valuable comments.

References

- [1] G. S. Stickney, F.-G. Zeng, R. Litovsky, and P. Assmann, "Cochlear implant speech recognition with speech maskers," *The Journal of the Acoustical Society of America*, vol. 116, no. 2, article 1081, 2004.
- [2] M. Marx, C. James, J. Foxton et al., "Speech prosody perception in cochlear implant users with and without residual hearing," *Ear and Hearing*, vol. 36, no. 2, pp. 239–248, 2015.
- [3] M. Chatterjee and S.-C. Peng, "Processing F0 with cochlear implants: modulation frequency discrimination and speech intonation recognition," *Hearing Research*, vol. 235, no. 1-2, pp. 143–156, 2008.
- [4] C. Lorenzi, G. Gilbert, H. Carn, S. Garnier, and B. C. J. Moore, "Speech perception problems of the hearing impaired reflect inability to use temporal fine structure," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 103, no. 49, pp. 18866–18869, 2006.
- [5] Z. M. Smith, B. Delgutte, and A. J. Oxenham, "Chimaeric sounds reveal dichotomies in auditory perception," *Nature*, vol. 416, no. 6876, pp. 87–90, 2002.
- [6] A. Boothroyd, "Adult aural rehabilitation: what is it and does it work?" *Trends in Amplification*, vol. 11, no. 2, pp. 63–71, 2007.
- [7] Q.-J. Fu and J. J. Galvin III, "Perceptual learning and auditory training in cochlear implant recipients," *Trends in Amplification*, vol. 11, no. 3, pp. 193–205, 2007.
- [8] C. D. Fuller, J. J. Galvin III, B. Maat, R. H. Free, and D. Başkent, "The musician effect: does it persist under degraded pitch conditions of cochlear implant simulations?" *Frontiers in Neuroscience*, vol. 8, article 179, 2014.
- [9] N. Kraus and B. Chandrasekaran, "Music training for the development of auditory skills," *Nature Reviews Neuroscience*, vol. 11, no. 8, pp. 599–605, 2010.
- [10] M. Besson, J. Chobert, and C. Marie, "Transfer of training between music and speech: common processing, attention, and memory," *Frontiers in Psychology*, vol. 2, article 94, 2011.
- [11] A. D. Patel, "Can nonlinguistic musical training change the way the brain processes speech? The expanded OPERA hypothesis," *Hearing Research*, vol. 308, pp. 98–108, 2014.
- [12] A. Parbery-Clark, A. Tierney, D. L. Strait, and N. Kraus, "Musicians have fine-tuned neural distinction of speech syllables," *Neuroscience*, vol. 219, pp. 111–119, 2012.
- [13] D. L. Strait and N. Kraus, "Biological impact of auditory expertise across the life span: musicians as a model of auditory learning," *Hearing Research*, vol. 308, pp. 109–121, 2014.
- [14] J. J. Galvin III, Q.-J. Fu, and G. Nogaki, "Melodic contour identification by cochlear implant listeners," *Ear and Hearing*, vol. 28, no. 3, pp. 302–319, 2007.
- [15] A. D. Patel, "Why would musical training benefit the neural encoding of speech? The OPERA hypothesis," *Frontiers in Psychology*, vol. 2, article 142, 2011.
- [16] J. Laneau, J. Wouters, and M. Moonen, "Improved music perception with explicit pitch coding in cochlear implants," *Audiology & Neurotology*, vol. 11, no. 1, pp. 38–52, 2006.
- [17] W. Wang, N. Zhou, and L. Xu, "Musical pitch and lexical tone perception with cochlear implants," *International Journal of Audiology*, vol. 50, no. 4, pp. 270–278, 2011.
- [18] K.-C. C. Joshua, C. C. Ann Yi, C. McMahon, J.-C. Hsieh, T.-H. Tung, and P.-H. L. Lieber, "Music training improves pitch perception in prelingually deafened children with cochlear implants," *Pediatrics*, vol. 125, no. 4, pp. e793–e800, 2010.
- [19] R. Watt and S. Quinn, "Analysis of local and global timing and pitch change in ordinary melodies," in *Proceedings of the 9th International Conference on Music Perception & Cognition (ICMPC '09)*, M. Baroni, A. Addressi, R. Caterina, and M. Costa, Eds., pp. 30–37, Bologna, Italy, 2006.
- [20] S. Greenberg, J. Hollenback, and D. Ellis, "Insights into spoken language gleaned from phonetic transcription of the switchboard corpus," in *Proceedings of the International Conference on Spoken Language Processing*, Philadelphia, Pa, USA, 1996.
- [21] A. E. Vandali, "Emphasis of short-duration acoustic speech cues for cochlear implant users," *The Journal of the Acoustical Society of America*, vol. 109, no. 5, pp. 2049–2061, 2001.
- [22] P. W. Dawson, A. A. Hersbach, and B. A. Swanson, "An adaptive Australian Sentence Test in Noise (AuSTIN)," *Ear and Hearing*, vol. 34, no. 5, pp. 592–600, 2013.
- [23] S. Peppé and J. McCann, "Assessing intonation and prosody in children with atypical language development: the PEPS-C test and the revised version," *Clinical Linguistics & Phonetics*, vol. 17, no. 4-5, pp. 345–354, 2003.
- [24] C. M. McKay and H. J. McDermott, "Perceptual performance of subjects with cochlear implants using the spectral maxima sound processor (SMSP) and the mini speech processor (MSP)," *Ear and Hearing*, vol. 14, no. 5, pp. 350–367, 1993.
- [25] A. S. Bregman, *Auditory Scene Analysis: The Perceptual Organization of Sound*, MIT Press, 1994.
- [26] R. Y. Litovsky, A. Parkinson, and J. Arcaroli, "Spatial hearing and speech intelligibility in bilateral cochlear implant users," *Ear and Hearing*, vol. 30, no. 4, pp. 419–431, 2009.
- [27] T. Y. C. Ching, P. Incerti, M. Hill, and E. Van Wanrooy, "An overview of binaural advantages for children and adults who use binaural/bimodal hearing devices," *Audiology and Neurotology*, vol. 11, supplement 1, pp. 6–11, 2006.
- [28] T. A. Salthouse, "What and when of cognitive aging," *Current Directions in Psychological Science*, vol. 13, no. 4, pp. 140–144, 2004.
- [29] A. S. Chan, Y.-C. Ho, and M.-C. Cheung, "Music training improves verbal memory," *Nature*, vol. 396, no. 6707, article 128, 1998.