

# UCSF

## UC San Francisco Previously Published Works

### Title

A Randomized Controlled Crossover Study of the Impact of Online Music Training on Pitch and Timbre Perception in Cochlear Implant Users

### Permalink

<https://escholarship.org/uc/item/0vr8w81h>

### Journal

Journal of the Association for Research in Otolaryngology, 20(3)

### ISSN

1525-3961

### Authors

Jiam, Nicole T  
Deroche, Mickael L  
Jiradejvong, Patpong  
[et al.](#)

### Publication Date

2019-06-01

### DOI

10.1007/s10162-018-00704-0

Peer reviewed

## Research Article



# A Randomized Controlled Crossover Study of the Impact of Online Music Training on Pitch and Timbre Perception in Cochlear Implant Users

NICOLE T. JIAM,<sup>1</sup> MICKAEL L. DEROCHE,<sup>2</sup> PATPONG JIRADEJVONG,<sup>1</sup> AND CHARLES J. LIMB<sup>1</sup> 

<sup>1</sup>*Department of Otolaryngology—Head and Neck Surgery, School of Medicine, University of California - San Francisco (UCSF), 2380 Sutter St, 1st Floor, San Francisco, CA 94115, USA*

<sup>2</sup>*Centre for Research on Brain, Language and Music, McGill University, Montreal, QC, Canada*

Received: 26 February 2017; Accepted: 17 October 2018

## ABSTRACT

Cochlear implant (CI) biomechanical constraints result in impoverished spectral cues and poor frequency resolution, making it difficult for users to perceive pitch and timbre. There is emerging evidence that music training may improve CI-mediated music perception; however, much of the existing studies involve time-intensive and less readily accessible in-person music training paradigms, without rigorous experimental control paradigms. Online resources for auditory rehabilitation remain an untapped potential resource for CI users. Furthermore, establishing immediate value from an acute music training program may encourage CI users to adhere to post-implantation rehabilitation exercises. In this study, we evaluated the impact of an acute online music training program on pitch discrimination and timbre identification. Via a randomized controlled crossover study design, 20 CI users and 21 normal hearing (NH) adults were assigned to one of two arms. Arm-A underwent 1 month of online self-paced music training (intervention) followed by 1 month of audiobook listening (control). Arm-B underwent 1 month of audiobook listening followed by 1 month of music training. Pitch and timbre sensitivity scores were taken across three visits: (1) baseline, (2) after 1 month of intervention, and (3) after 1 month of

control. We found that performance improved in pitch discrimination among CI users and NH listeners, with both online music training and audiobook listening. Music training, however, provided slightly greater benefit for instrument identification than audiobook listening. For both tasks, this improvement appears to be related to both fast stimulus learning as well as procedural learning. In conclusion, auditory training (with either acute participation in an online music training program or audiobook listening) may improve performance on untrained tasks of pitch discrimination and timbre identification. These findings demonstrate a potential role for music training in perceptual auditory appraisal of complex stimuli. Furthermore, this study highlights the importance and the need for more tightly controlled training studies in order to accurately evaluate the impact of rehabilitation training protocols on auditory processing.

**Keywords:** music training, cochlear implants, normal hearing, pitch discrimination, timbre identification

*Correspondence to:* Charles J. Limb · Department of Otolaryngology—Head and Neck Surgery, School of Medicine · University of California - San Francisco (UCSF) · 2380 Sutter St, 1st Floor, San Francisco, CA 94115, USA. Telephone: 415-353-2870; email: charles.limb@ucsf.edu

## INTRODUCTION

Cochlear implants (CIs) are surgically implanted devices that allow people with hearing loss recover some degree of hearing. Although CIs are successful in processing speech in quiet environments, present-

day systems remain far from ideal in conveying spectrally complex sounds due to their delivery of impoverished spectral cues and poor frequency resolution in signal processing (Limb 2006; Reiss et al. 2007; Nardo et al. 2007; Shannon et al. 2004). Such deficits have serious ramifications on challenging listening conditions like music perception. Among all the musical elements, pitch (Gfeller et al. 2002a; Looi et al. 2012; Limb and Roy 2014), and timbre (Gfeller et al. 2002b; McDermott 2004; Looi et al. 2012; Limb and Roy 2014; Heng et al. 2011) are the most challenging elements for CI users to process because of the device-mediated limitations in conveying spectral fine structure. As such, there is considerable variability among CI users in the perceptual sensitivity to musical stimuli (Gfeller et al. 2008, 2010; Looi et al. 2012; Galvin et al. 2009). Because of the inconsistency with cochlear implant outcomes with regard to complex sound processing, there has been ongoing but limited research on the impact of auditory training with CI users (Galvin et al. 2007; Galvin et al. 2008; Galvin et al. 2009; Gfeller et al. 2002c, 2015).

Speech recognition training for CI patients has been an early, important area of auditory rehabilitation research (Fu et al. 2005; Busby et al. 1991; Dawson and Clark 1997). Although Busby et al. (1991) failed to find significant changes in speech recognition performance among three CI users after ten 1 h sessions (1–2 sessions per week) of auditory training, the majority of the literature reports gains in speech perception tasks with longitudinal auditory training. These benefits include vowel recognition (Dawson and Clark 1997), sentence recognition (Fu and Galvin 2003), intervocalic consonants, medial vowels in monosyllables, and words in sentences (Rosen et al. 1999) with as little as 1 h a week of auditory training over 5 weeks (Nogaki et al. 2007).

Over the years, CI processing systems have evolved such that most implantees are able to achieve satisfactory to excellent speech perception scores in quiet environments. With this development, the initial focus on speech recognition training programs within auditory rehabilitation research has naturally shifted towards complex auditory training paradigms, such as music. Part of the interest in music training arises from studies demonstrating enhanced perceptual outcomes in normal hearing (NH) musicians (Barrett et al. 2013; Shahin et al. 2003; Kaganovich et al. 2013; Musacchia et al. 2008; Pantev et al. 2001). Nevertheless, even among non-musicians with NH, experience-based neural plasticity studies have shown that music training promotes functional advantages in the perception of complex auditory input, such as pitch discrimination (Tervaniemi et al. 2005; Micheyl et al. 2006; Bidelman et al. 2011; Brown et al. 2014)

and timbre specificity (Pantev et al. 2001; Strait et al. 2012; Brown et al. 2014).

These training-induced benefits on music perception among normal hearing adults (Tierney and Kraus 2015; Slater and Kraus 2016; Woodruff Carr et al. 2014; Kraus and Chandrasekaran 2010; Kraus et al. 2014), have also been reported in CI users. Gfeller et al. (2002c) observed significant improvement in CI users' ability to recognize and appraise timbre after 3 months of home-based music training (ten 30-min instruction modules, four times per week) on musical instruments. Galvin et al. (2007) demonstrated that melodic contour identification training improved all six of six CI listeners' abilities to identify familiar melodies and melodic contours. While these longitudinal studies suggest a promising role for music training in CI-mediated music perception, none of these studies employed a rigorous control group involving a non-musical acoustic intervention. Additionally, many of these studies lack a high degree of variability in auditory stimuli and the provision of feedback—parameters associated with successful perceptual training (Vandali et al. 2015; Gfeller et al. 1999). Furthermore, it is unclear how accessible the training programs would be for the clinical population, whether findings are generalizable to untrained tasks and/or conditions, and whether they are acoustically rewarding to the end user.

The minimal training period required to see benefit is not well-established in the literature. Prior auditory training paradigms have ranged from 3 months to as many as 36 months of training (Petersen et al. 2012; Chen et al. 2010; Gfeller et al. 2000, 2002c), with few exceptions of shorter training intervals. Long training periods maximize training effects. However, this may come at the cost of subject attrition and compliance rates. On the other hand, shorter training periods improves compliance and lowers attrition. A potential drawback may be reduced training effects among participants. These limitations may partially explain why relatively few studies used short training intervals. Among the handful of studies, Nogaki et al. (2007) reported adaptation to spectrally shifted speech with a training paradigm requiring 1 h a week for 5 weeks. Research by Driscoll (2012) also demonstrated significant pre-to-posttest improvement in musical instrument identification accuracy after 15 10-min lessons over 5 weeks. Again, it is important to note that in these studies, the CI users were trained on the same tasks and conditions on which they were tested. Additionally, there were no control conditions in this study.

The benefits of spaced rehearsals on learning, on the other hand, are well established in the field of psychology (Sternberg 1999) and are now being applied to CI users. Spaced practice of 12 30-min

lessons over the period of 4 weeks has been reported as an acceptable training interval by CI recipients from at least one prior study (Looi et al. 2012). Several authors have identified obstacles that limit patient use of current in-person music rehabilitation programs, as they require time, financial resources, ability to travel, and awareness of the resources available to themselves (Herholz and Zatorre 2012; Patel 2011). Again, a long, intensive training program may deter participation and lead to delays in patient rehabilitation.

As such, we designed a randomized control crossover study to evaluate the impact of an online, short music training intervention on pitch and timbre perception in CI users. Our study involved two control conditions: a hearing and a treatment control. NH listeners were enrolled in the study as a control group to CI users. This paradigm was intended to determine whether the effects of a music training intervention are specific to one type of hearing population versus generalizable to all participants. Improvements among CI users would suggest the presence of stimulus learning through impoverished spectrotemporal degradations induced by CI processors, whereas improvements in both study populations might indicate the use of auditory cues accessible to both hearing populations or perhaps reflect a form of non-stimulus learning. Additionally, our study implemented audiobook listening as a control intervention to compare the efficacy of music training on pitch and timbre perception. We used audiobooks because they contain non-musical acoustic stimuli while being both cognitively engaging and enjoyable for participants, thereby representing a very robust auditory control. By choosing audiobook listening as a control task, the quality and quantity of participation may have been biased towards the control intervention. We decided that for the purposes of our study, however, it was important to assess the impact of non-musical auditory training paradigms that were comparably robust to the musical paradigms. The study employed a short training interval (8 h over the course of 1 month) and repetition spacing (2 h a week over 4 weeks). We hypothesized that CI users participating in an online music training program would demonstrate improvements in pitch discrimination and timbre identification, two auditory perception tasks known to be especially challenging for CI users.

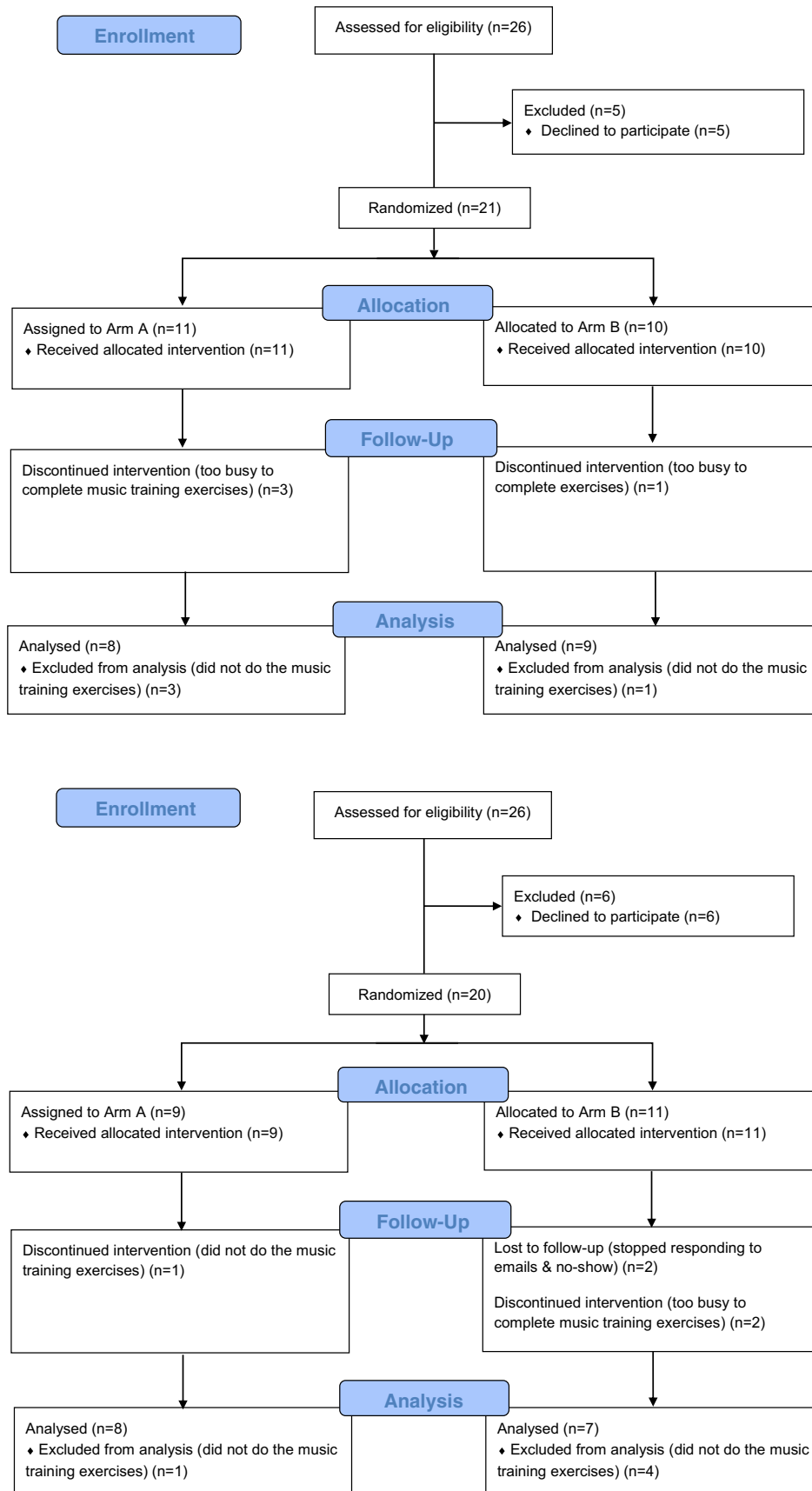
## MATERIAL AND METHODS

### Study Design

We conducted a controlled, parallel group, crossover study on adults with NH or wearing a CI at a tertiary referral center. This study initially enrolled 21 NH adults (Fig. 1—left panel) and 20 CI users

(Fig. 1—right panel). We expected a considerably large difference between NH and CI subjects in pitch and timbre sensitivity. Assuming a Cohen's  $d$  effect size of 0.8, an a priori power of 0.95 required a sample size of only 16 subjects between two groups with  $\alpha=0.05$  and a correlation among the three repeated measures of 0.5 (using G-Power version 3.1.9.2). In other words, the deficits exhibited by CI users in these tasks were expected to be easily observable. In contrast, the training effects were expected to be subtler, given the short-term nature of the proposed intervention. Assuming an effect size of 0.3, an a priori power of 0.8 required a sample size of 20 subjects in each population (NH and CI) for this within-subject factor measured over three replications. Possible within-between interactions (e.g., whether these benefits would differ depending on the arm) also required 20 subjects in each of the two groups to achieve a power of 0.8 with an effect size of 0.3. Nine participants were enrolled but eliminated from analysis due to attrition and a time-related inability to complete the training paradigm (2 h a week for 4 weeks of music training and 4 weeks of audiobook listening). Five NHA, two NHB, two CIA, and two CIB subjects did not complete a total of 8 h of music training exercises. However, we opted for a conservative approach towards our study design and included their results in the data analysis. Ultimately, a total of 17 NH adults and 15 CI users were included in the study that took place between February 2016 and May 2016. All participants were stratified by hearing status (NH versus CI) and were randomly assigned to one of two parallel arms (Arm-A or Arm-B) at the start of the study. Arm-A consisted of a 1-month period of music training (intervention) followed by a 1-month period of audiobook listening (control). Arm-B consisted of a 1-month period of audiobook listening followed by a 1-month period of music training. The audiobook intervention served as a control group for the music training intervention because it is a form of engaging non-musical auditory perception and processing.

All participants completed a standardized CI questionnaire and a musical experience questionnaire upon study enrollment. The “Cochlear Implant Questionnaire” asked for the etiology of hearing loss, age of hearing loss onset, age at profound hearing loss, years of profound deafness, pre-/post-lingual deafness, date(s) of implantation, years of implant use, unilateral/bilateral CI user, preferred ear if s/he is a bilateral implant user, presence of bimodal hearing, CI device, internal/external processor, and processing strategy. The “Musical Experience Questionnaire” asked the participants for information regarding years of formal versus self-taught music training, age at the onset of music training, learning



**FIG. 1.** CONSORT 2010 flow diagram of music training versus audiobooks for pitch and timbre perception among the normal hearing cohort (left panel) and cochlear implant cohort (right panel). The diagram includes detailed information on excluded participants

TABLE 1

Detailed cochlear implant history									
Subject	HL etiology	Age	Age at onset of HL	Age at profound HL	Pre-/post-lingually deaf	Age of implantation	Years of implant use	Uni/bilateral CI	CI device
1-CI-A	Idiopathic	65	40	61	Post	64	1	Uni (R)	Med-El Rondo
2-CI-A	Congenital	64	0	0	Pre	63	1	Uni (L)	Med-El Sonnet
3-CI-A	Meniere's	59	40	57	Post	56 (L); 57 (R)	2.5 (L); 1.75 (R)	Bi (L Pref)	AB HiRes 90 K Advantage Naida Q70
4-CI-A	Idiopathic	63	12	58	Post	57 (L); 55 (R)	6 (L); 8 (R)	Bi (L Pref)	Cochlear N6 (L); N5 (R)
5-CI-A	Congenital	52	0	0	Pre	50 (L); 48 (R)	1.5 (L); 3.5 (R)	Bi (R Pref)	Cochlear N5
6-CI-A	Congenital	30	0	0	Pre	22 (L); 2 (R)	8 (L); 28 (R)	Bi (R Pref)	Cochlear N5
7-CI-A	Genetic	73	43	68	Post	69	4	Uni (R)	Med-El Rondo/Opus II
8-CI-A	Idiopathic	61	58	58	Post	58	3	Uni (L)	Med-El Rondo/Opus II
9-CI-A	EVAS	54	3	3	Post	47 (L); 53 (R)	7 (L); 0.75 (R)	Bi (L Pref)	AB HiRes 90 K Hi Focus Mid-Scala Naida Q70
1-CI-B	Idiopathic	77	55	61	Post	75	2	Uni (L)	Cochlear
2-CI-B	Idiopathic	63	8	8	Post	59	4	Uni (L)	AB HiRes 90 K Harmony
3-CI-B	Idiopathic	79	68	71	Post	72	7	Uni (R)	Med-El Rondo/Opus II (Pulsar)
4-CI-B	Aging	84	55	65	Post	77	7	Uni (R)	Cochlear Nucleus
5-CI-B	Congenital	20	0	0	Pre	13 (L); 4 (R)	6.5 (L); 16 (R)	Bi (L Pref)	AB HiRes 90 K Naida (L); Harmony (R)
6-CI-B	Idiopathic	57	43	50	Post	55	2	Uni (R)	Cochlear N6
7-CI-B	Congenital	28	1	3	Pre	11 (L); 18 (R)	17 (L); 10 (R)	Bi (L Pref)	Cochlear N5
8-CI-B	Idiopathic	62	45	53	Post	53 (L); 56 (R)	9 (L); 6 (R)	Bi (No Pref)	AB HiRes 90 K Naida
9-CI-B	Idiopathic	79	49	59	Post	78	1	Uni (L)	AB HiRes 90 K Naida
10-CI-B	Genetic (Norrie's)	31	12	17	Post	28	2.5	Uni (L)	AB HiRes 90 K Mid-Scala Naida QX70
11-CI-B	Spinal meningitis	52	1.25	5	Post	42	10	Uni (R)	Med-El Rondo Pulsar

HL hearing loss, EVAS enlarged vestibular aqueduct syndrome, L left ear, R right ear, Pref preferred, Uni unilateral, Bi bilateral

setting, hours per week of practice, instruments or skill, whether s/he is currently practicing music, and hours per week spent listening to music.

The study involved a total of three visits. At the first visit, we obtained the participants' baseline pitch and timbre performance scores. Depending on the arm assignment, participants would undergo 1 month of music training or audiobook listening after the first visit. After 1 month of music training or audiobook listening, the subjects returned to the laboratory to undergo pitch and timbre testing. Then, they were instructed to undergo the second form of training (1 month of either music training or audiobook listening) in this crossover study design. After completing the sequence of training interventions, participants would return for their third and last visit for pitch and timbre testing. Subjects were paid by-the-hour for the three testing sessions (pre-intervention testing; after the first month; after the second month). They were not compensated for at-home training sessions or for using the music training program.

Subjects were allowed to quit the study at any point without repercussions to their clinical care or payment. The research protocol was reviewed and approved by the local institutional review board and we obtained written consent from all participants.

## Participants

The criteria for participant eligibility included adults (aged 18 or over) with either NH status, a unilateral CI, or bilateral CIs. The exclusion criteria included any medical conditions that may prevent participation in this study (e.g., vision problems, cognitive disorders) and failure to comply with the study requirements of music training and audiobook listening 2 h each week for 4 weeks.

Among the participants who complied with the study requirements, Arm-A was comprised of eight NH adults (NHA) and eight CI adults (CIA) and Arm-B was comprised of nine NH adults (NHB) and seven CI adults (CIB). Within the NH cohort, the average

participant age was 37 years (SD 16; range 22–75). The average age of the CI participants was 63 years (SD 13; range 28–84). The duration of implant use ranged from 1 to 28 years (Table 1). The age at profound hearing loss varied depending on the hearing loss etiology (3 to 71 years old). Six CI users were pre-lingually deafened. In this study, we categorized individuals who were deaf at birth and/or acquired hearing loss prior to their ability to speak as pre-lingually deafened.

Of the 15 CI users, 9 were unilateral CI users and 6 were bilateral CI users. We recruited both unilateral and bilateral CI users to ensure we had an adequate sample size for the study. However, to maximize homogeneity among our test subjects and our testing conditions, we tested all CI users using only one CI at a time. Bilateral CI users were tested only on the preferred side; usually this was the side implanted first. Seven unilateral CI users used hearing aids in their contralateral ear. Regardless of the hearing status of the contralateral ear, all participants were asked to remove any assistive hearing device (e.g., hearing aids or contralateral CI devices for bimodal users) and occlude the contralateral ear with an earplug to avoid any confounding effect that could arise from residual hearing. This was asked for all tasks including both the laboratory tests and the at-home training sessions. We had representation from three CI manufacturers in this study (Med-El, Advanced Bionics, Cochlear Americas).

### Protocol for Testing Pitch Sensitivity and Timbre Discrimination

At the start of each round of testing, the study investigator explained the testing protocol to the research participant. Each participant completed the pitch and timbre tasks in a soundproof acoustic chamber. Loudspeakers (Sony SS-MB150H) were placed directly in front of the subject's face in a seated position, and the user-interface was displayed on a touch-screen monitor located inside the sound booth. All acoustic stimuli were sampled at 44.1 kHz and 16-bit resolution. The stimuli were presented via a mixer (Alesis Multimix 6 USB), a stereo power amplifier (Pyle PCA2), and a single calibrated loudspeaker, located approximately 2 ft from the subject at an average level of 65-dB sound pressure level (dB SPL). Listeners manually selected their response to the task at-hand using the touch-screen monitor. Subjects were instructed to be as accurate as possible in a timely manner. Immediate feedback was provided to the participant regarding whether his or her answer was accurate or inaccurate in all the tests. Typical experimental sessions lasted between 1.5 and 2 h for CI users and about 1 h for NH subjects, as CI

users generally required more time to navigate the testing interface and experimental setup.

**Pitch Task.** The pitch perception protocol was adapted from Deroche et al. (2014). Each pitch testing session was comprised of a practice block (20 trials in a practice block) followed by two test blocks (140 trials in each test block). In each trial, participants were presented with two complex tones of varying intervals in seven steps of power of 2 (from 1/16th of a semitone to 4 semitones for NH subjects and from half a semitone to 32 semitones for CI subjects) and asked to report which of the two stimuli was higher in pitch. The seven pitch change conditions were randomly ordered within a block, and the order in which the target (higher pitched tone) was presented in a given trial was also randomly determined. We used different parameters for CI versus NH listeners because the differences in pitch perceptions between the two groups were known to be considerable. Using the same interval parameters for both cohorts would either create a task that would be too easy for NH listeners (hence, an observable ceiling effect) or a task that would be too difficult for CI users (hence, a floor effect).

Practice blocks differed from the test blocks in that the trials were presented at (a) a longer duration (500-ms versus 300-ms long) separated by longer inter-stimulus intervals (500-ms versus 300-ms long), (b) without level roving, and (c) consisted only 20 trials. We used the practice session to familiarize the listener to the task, to ensure that the subject understood the instructions clearly, and to optimize data collection.

The tones were broadband harmonic complexes, with partials in sine phase with equal amplitude, low-pass filtered at 8 kHz with a sixth order Butterworth filter, and gated by 10-ms onset and offset ramps. We incorporated roving of the reference fundamental frequency ( $F_0$ ) between 100 and 150 Hz to ensure that the pitch changes did not target a particular  $F_0$  range where sensitivities could have been particularly fine or poor in an individual. All acoustic signals presented in the test block were 300-ms long, interspaced by 300 ms between them. Regardless of the  $F_0$ , the level of each complex was equalized at 65 dB SPL prior to the start of the study and presented with a  $\pm 3$  dB rove to hinder the use of potential loudness cues for pitch directionality (particularly in the case of CI users).

**Timbre Task.** Similar to the pitch task, each timbre testing session was comprised of a practice block (16 trials in a practice block, one trial for each instrument) followed by two test blocks (160 trials in each). The practice block allowed participants to familiarize themselves with the task. In each trial, participants were presented with a 750-ms-long acous-

tic stimulus and asked to identify the source of the sound from a set of 16 musical instruments illustrated by corresponding images. The list of 16 musical instruments equally represented the four instrument classes (four woodwinds, four percussions, four brasses, and four strings). The instruments represented in the timbre task were the French horn, tuba, trumpet, trombone, bassoon, oboe, piccolo, clarinet, cello, guitar, ukulele, banjo, xylophone, marimba, steel drum, and piano. We used Logic Pro X software (Apple; Cupertino, CA) to create the instrumental sounds. All acoustic signals were presented at the same pitch ( $F_0$  of 261.6 Hz) and level (65 dB SPL). We elected to present our timbre stimuli at a pitch of middle C (C4), because this is a commonly used vocal and instrumental frequency in Western music. Many studies have used base frequencies in C4 (261.6 Hz) because of its regular use in music (Gfeller et al. 2002c; Vandali et al. 2005; Laneau et al. 2006; Kang et al. 2009). Our preference to use musical instruments whose written range includes the middle C limited our timbre stimuli to a select list of instruments. However, we decided it would be more appropriate to select instruments that naturally had a performance range that included 261.6 Hz rather than to artificially create an acoustic range for a more well-known instrument.

## Intervention

The intervention was a commercially available, online, self-paced music training program (Meludia; Paris, France; [www.meludia.com](http://www.meludia.com)) with more than 600 musical exercises covering a range of musical features, such as micro-melody (ascending/descending melody), melodic patterns, pitch direction, pitch identification, density (harmonics), stable/unstable (consonance/dissonance), chord quality, rhythm, and arpeggios. Participants were provided with an account for the duration of the study. Study coordinators were also able to access the subjects' accounts for practice-log verification and technical assistance. There was standardization among users on the genre of exercises they underwent for the music training month. However, individuals had control over which modules they did within those specific genres of exercises. All participants started their music training exercises at the beginner level (the easiest level available) and had access to the same set of music exercises. Users could only access the next exercise (in increasing level of difficulty) after correctly completing the previous one, ensuring proficiency and sequential advancement. Study participants varied on how quickly they proceeded through the learning modules over 4 weeks.

The control arm consisted of 1 month of non-musical audiobook listening. Participants were allowed to choose an audiobook of their choice in order to ensure auditory engagement (rather than tuning out or sleeping through the audiobook) during the control arm segment of the study. The only requirement was that the audiobook listening was a purely auditory but non-musical experience, meaning that subjects were not permitted to read along with the book. The no-reading criterion was implemented to ensure that audiobook listening served as a form of auditory processing and that participants did not replace the audiobook paradigm with visual processing (aka reading a book with auditory sounds in the background), and thereby make this segment of non-musical auditory processing moot. Some subjects chose to listen to the talk radio or to podcasts if the medium was more accessible than an audiobook.

Participants were instructed to participate in a minimum of 2 h a week of music training exercises or audiobook listening over the course of 4 weeks, totaling at least 8 h on each arm. During the study, subjects were asked to keep a practice log of their music training sessions and audiobook listening intervals to ensure study compliance. The log asked for detailed information including the start and stop time, the completed exercise, and notes on their experience for each practice session. Every participant completed and provided the study investigator with his or her practice logs. The music training software also records the time users spend on exercises and the specific exercise completed as a part of the data analytics. Study investigators had access to the data analytics related to the study. Finally, the participants' time-keeping sheets were randomly cross-verified with the software time-keeping system for accuracy.

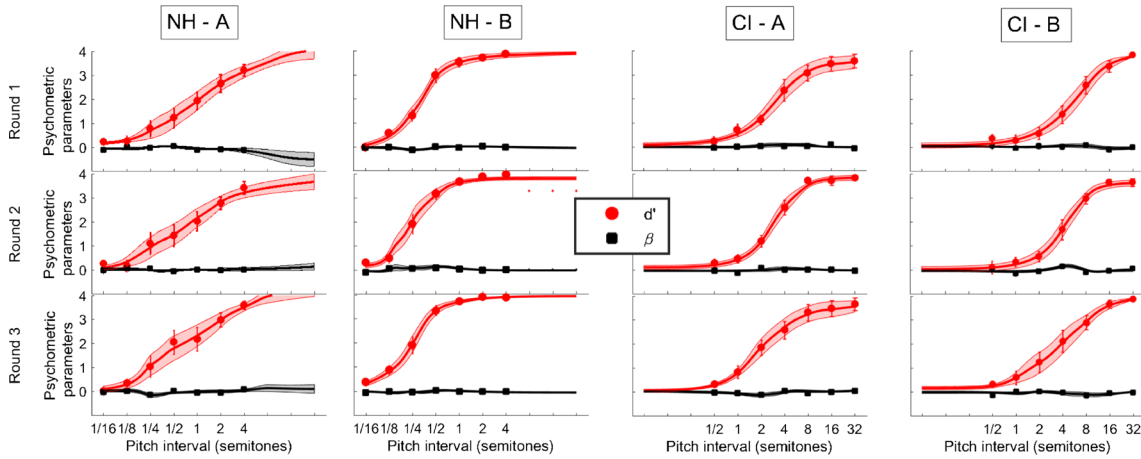
## Randomization

Research subjects were first stratified by hearing loss status (NH versus CI). For allocation of the participants within this sub-group, a computer-generated list of random numbers was used. Participants were randomly assigned following simple randomization procedures (computerized random numbers) to 1 of 2 crossover arms. Randomization by a computer-generated random list was prepared by one of the study investigators and this assignment sequence was kept concealed in the subject's folder.

## Data Analysis

The primary outcome with respect to efficacy of music training was the change in pitch and timbre percep-





**FIG. 2.**  $d'$  and  $\beta$  for each group of subjects in each pitch round. In all panels, psychometric functions ( $d'$  values) showed how performance increased with the pitch interval. Each data point consists of 40 trials, totaling 280 trials per round

tion from pre-intervention performance to post-intervention performance. The primary endpoint was change in pitch and timbre perception after each month of intervention. All data analyses were performed using MATLAB (MATLAB Release 2015b, The MathWorks, Inc., Natick, MA) and SPSS (IBM SPSS Statistics Version 23, IBM Corp., Armonk, NY). We used type III sums of squares for the analyses because the number of subjects per study group was uneven, due to attrition.

## RESULTS

### Pitch Task: Psychometric Findings

We found that the performances for the upwards and downwards changes in pitch were very similar and increased nicely along the intervals used in each population, namely from 1/16th of a semitone to 4 semitones for NH subjects and from half a semitone to 32 semitones for CI subjects. Hit and false alarm rates were translated into d-prime ( $d'$ ) and beta ( $\beta$ ) data. These psychometric parameters ( $d'$  and  $\beta$ ) were obtained for each group of subjects in each round. A

maximum likelihood technique was then used to fit a Weibull function to the performance data, and a  $d'$  and  $\beta$  fit was reconstructed from these performance fits. As illustrated by the red lines in Fig. 2, among all subject groups (NHA, NHB, CIA, CIB) and the three rounds of testing,  $d'$  increased steadily with the pitch interval reaching a plateau between 3 and 4 for the largest (easiest) intervals. The  $\beta$  functions (black) were overall flat at 0, revealing no apparent bias for upwards or downwards changes in pitch. A performance threshold was extracted from the  $d'$  fit, at a  $d'$  value of 0.77, which in a 2AFC task corresponds to a performance of 70.7%, for each subject and each round.

In order to examine more specifically the effect of music training compared with the effect of audiobooks, the pitch direction thresholds at  $d' = 0.77$  were passed through a repeated-measures analysis of variance with one within-subject factor (round, with three levels) and two between-subjects factors (hearing status and arm). The results are displayed in Table 2. The thresholds were log-transformed (base 2) as represented in Fig. 3, so that the assumption of homogeneity of variance between NH and CI subjects could be respected. There was a main effect of round, reflecting that thresholds improved over time. Pairwise comparisons showed that thresholds were similar between round 1 and 2 ( $p = 0.694$ ) but were lowered (enhanced) at round 3 ( $p < 0.001$  and  $p = 0.018$  relative to rounds 1 and 2, respectively). Critically, we hypothesized that the effect of round would have been different for the subjects undergoing arm-A (i.e., thresholds improving considerably from rounds 1 to 2) or arm-B (i.e., thresholds improving considerably from rounds 2 to 3) and perhaps differently for NH and CI subjects, but none of these hypotheses were supported. Round did not interact with arm, hearing status, or both. There was also a main effect of hearing status ( $p < 0.001$ ). This was expected; CI users required much larger pitch intervals

**TABLE 2**

Statistics of thresholds in the pitch discrimination task

Factors	F value	p value
Round	$F(2, 56) = 9.3$	$< 0.001^*$
Hearing	$F(1, 28) = 96.5$	$< 0.001^*$
Arm	$F(1, 28) = 1.0$	0.317
Round $\times$ hearing	$F(2, 56) = 0.4$	0.673
Round $\times$ arm	$F(2, 56) = 0.2$	0.860
Hearing $\times$ arm	$F(1, 28) = 11.1$	$0.002^*$
3-way	$F(2, 56) < 0.1$	0.921

The asterisks signifies a p-value less than 0.05 - the significance threshold used in this study

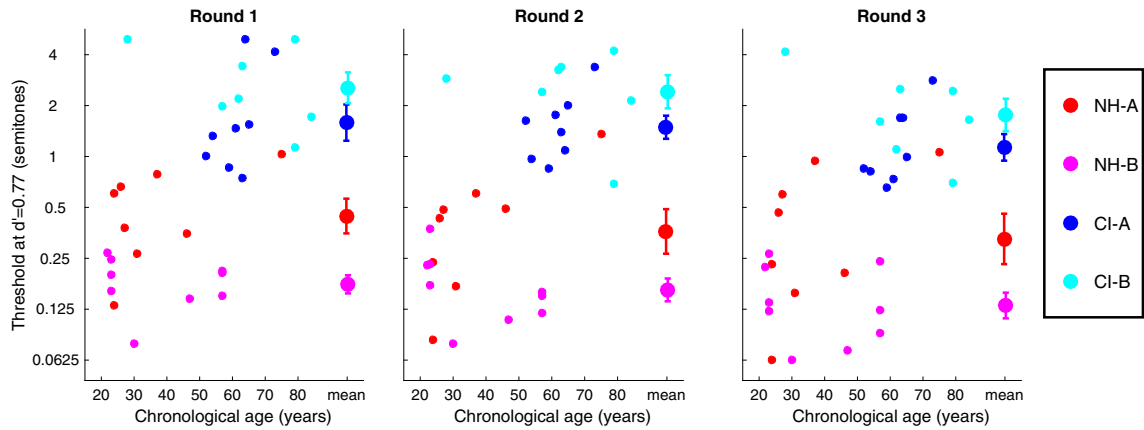


FIG. 3. Individual and aggregated pitch discrimination  $d'$  thresholds with each round of testing

than NH subjects for the same level of performance, on the order of two semitones versus a quarter of a semitone, on average over the two arms (Fig. 3). Finally, there was an interaction between hearing status and arm: namely, the NHB group outperformed the NHA group ( $p=0.004$ ), while the CIA group did not differ from the CIB group ( $p=0.125$ ). Because this was true in all three rounds of testing, this interaction had presumably nothing to do with the type of auditory intervention received, but was instead simply the result of differential abilities intrinsic to the subjects selected randomly in each group. It is worth noting that these trends were no different when we repeated the same analysis with  $d' = 1.44$  (pitch intervals being twice more discriminable).

TABLE 3

Statistics of the  $d'$  data for classes of instruments, in the timbre task

Factors	$F$ value	$p$ value
Class	$F(3, 84) = 111.7$	$< 0.001^*$
Round	$F(2, 56) = 21.4$	$< 0.001^*$
Hearing	$F(1, 28) = 109.3$	$< 0.001^*$
Arm	$F(1, 28) < 0.1$	0.886
Class $\times$ round	$F(6, 168) = 3.4$	0.004*
Class $\times$ hearing	$F(3, 84) = 37.4$	$< 0.001^*$
Class $\times$ arm	$F(3, 84) = 1.8$	0.155
Round $\times$ hearing	$F(2, 56) = 0.8$	0.431
Round $\times$ arm	$F(2, 56) = 3.0$	0.059
Hearing $\times$ arm	$F(1, 28) = 6.3$	0.018*
Class $\times$ round $\times$ hearing	$F(6, 168) = 1.9$	0.091
Class $\times$ round $\times$ arm	$F(6, 168) = 1.0$	0.429
Class $\times$ hearing $\times$ arm	$F(3, 84) = 2.8$	0.045*
Round $\times$ hearing $\times$ arm	$F(2, 56) = 0.5$	0.633
4-way	$F(6, 168) = 0.9$	0.503

The asterisks signifies a  $p$ -value less than 0.05 - the significance threshold used in this study

### Timbre Task: $d'$ for Instrument Class Identification

In order to determine the effects of the online music training program on timbre identification, we performed a repeated-measures analysis of variance on the  $d'$  calculated for instrument classes (woodwinds, brass, percussion, strings), with two within-subject factors (class, round) and two between-subject factors (hearing status, arm) (Table 3). There were several interesting observations to report. First, there was a main effect of round and a near significant interaction between round and arm ( $p=0.059$ ). Subjects allocated to Arm-A improved from rounds 1 to 2 ( $p<0.001$ ) but not from rounds 2 to 3 ( $p=0.649$ ). On the other end, subjects that were allocated to Arm-B did not improve sufficiently from rounds 1 to 2 ( $p=0.119$ ), but performed better on round 3 ( $p=0.013$ , for the comparison rounds 1 to 3). This distinct pattern of benefit is promising because it taps into an effect that may be specific to music training as opposed to procedural training alone.

We also observed an interaction between round and class. Post hoc tests (with Bonferroni corrections) revealed that the effect of round was significant for each instrumental class ( $p=0.009, 0.004, <0.001, <0.001$ ). In each class, the improvement occurred from rounds 1 to 2 ( $p<0.008$ ) but not from rounds 2 to 3 ( $p>0.999$ ). In other words, this interaction was primarily due to the effect sizes attributed to round as a function of instrument class. The total improvement in  $d'$  (i.e., from rounds 1 to 3) amounted to 0.20, 0.30, 0.51, and 0.29, respectively, for brasses, percussions, strings, and woodwinds.

Similar to the previous analyses, there was a main effect of hearing status. NH subjects obtained  $d'$  values that were 2.0 higher (3.3 versus 1.3, on average) than CI users. Hearing status interacted with arm since NHB tended to outperform NHA ( $p=0.095$ ) while CIA tended to outperform CIB ( $p=0.079$ ), across all rounds of testing. These observations are

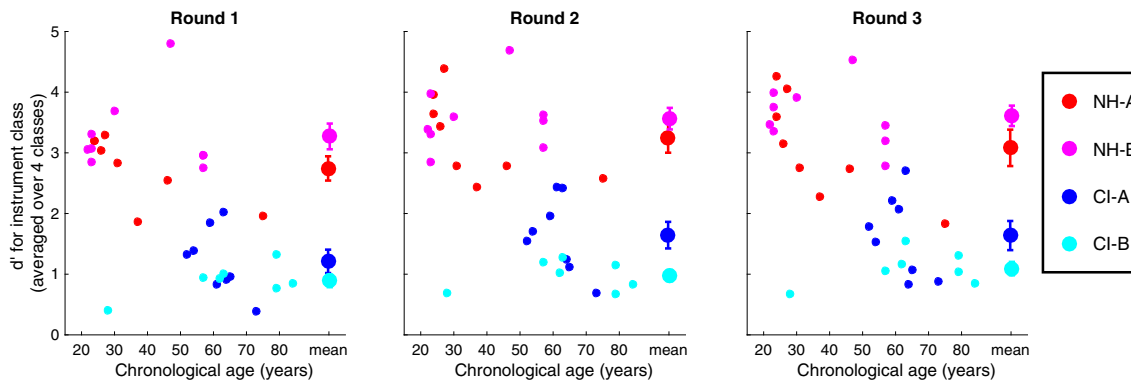


FIG. 4. Averaged  $d'$  values in the timbre task for NH and CI subjects as a function of instruments' class

simply a product of differences in the initial abilities of subjects allocated to each group.

Finally, there was a main effect of instrument class, which was dependent on the hearing status (Fig. 4 and Table 4). NH subjects discriminated percussive instruments the best ( $d' = 4.7$ ,  $p < 0.001$  relative to the three other classes)—followed by strings ( $d' = 3.9$ ,  $p < 0.001$  relative to brasses and woodwinds), brasses ( $d' = 2.3$ ,  $p > 0.999$ ), and woodwinds ( $d' = 2.3$ ,  $p > 0.999$ ). CI subjects also discriminated percussive instruments better ( $d' = 1.9$ ,  $p < 0.001$  relative to the three other classes) than string, brass, and woodwind instruments ( $1.0 < d' < 1.1$ ,  $p > 0.999$  among them). With respect to hearing status, CI users displayed lower  $d'$  values than NH subjects with every instrument class ( $p < 0.001$  in every case). The size of these deficits in instruments' class discrimination was qualitatively larger for strings (2.8) and percussions (2.7) than for woodwinds (1.3) and brasses (1.2). This observation is noteworthy because it is often thought that CI users perform better with percussive sounds due to the preservation of salient temporal cues. That pattern was reflected in this study; however, implantees demonstrate greater deficits within this instrument class than with brasses or woodwinds (effect of hearing status on the  $d'$  difference between percussions and brasses, or between percussions and woodwinds:  $p < 0.001$  in both cases) relative to their NH counterparts.

### Examination of Plasticity-Related Factors

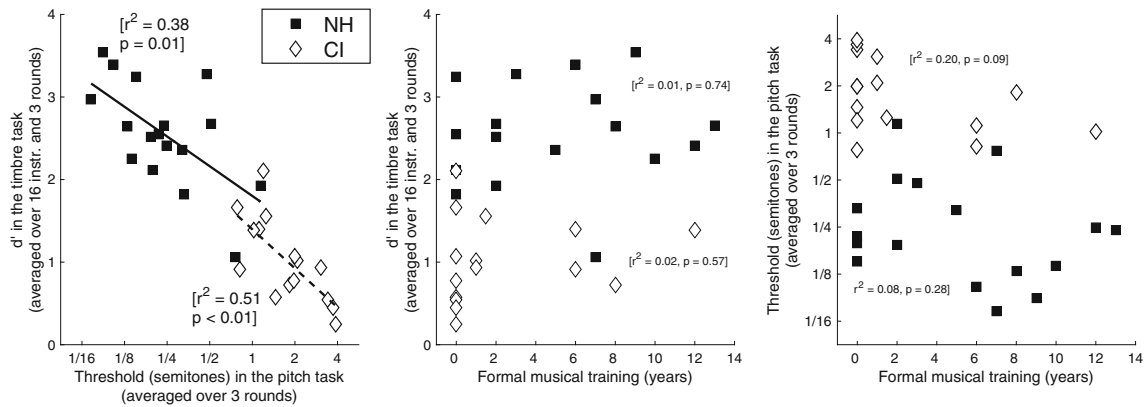
Subjects' mean performance (the averaged  $d'$  across all three testing sessions) in the pitch and timbre tasks were strongly correlated (NH listeners  $r^2 = 0.38$ ,  $p = 0.01$ ; CI users  $r^2 = 0.51$ ,  $p < 0.01$ ). Listeners with finer pitch sensitivity were also better at discriminating among instruments. This could certainly have been expected when considering all subjects at once (given that the data points occupied different corners of the correlation space, depending on the hearing status), but this result held within NH subjects alone and also within CI subjects alone (Fig. 5, left-most panel). We attempted a similar correlation between (1) mean  $d'$  in the timbre task and the music training-induced improvement in pitch thresholds; (2) the music training-induced improvement in  $d'$  for the timbre task and mean pitch thresholds; (3) the music training-induced improvement in  $d'$  for the timbre task with the music training-induced improvement in pitch thresholds. None of these correlations were significant, meaning that the correlation observed in the left panel of Fig. 5 were unlikely to be related to the training effects examined in this study but more likely stemmed from baseline differences among subjects that existed prior to the study. We also attempted to correlate mean threshold in the pitch task (averaged over the three rounds) and mean  $d'$  in the timbre task (averaged over 16 instruments and three rounds) with four different plasticity-related factors: age at onset of hearing loss (pitch task—CI users  $r^2 = 0.00$ ,  $p = 0.86$ ; timbre task—CI users  $r^2 = 0.02$ ,  $p = 0.66$ ); age at onset of

TABLE 4

Confusion matrix averaged over rounds and subjects of same hearing status for instruments classes

		Instruments' class responded by NH subjects				Instruments' class responded by CI subjects			
		Brasses	Percussions	Strings	Woodwinds	Brasses	Percussions	Strings	Woodwinds
Presented instrumental class	Brasses	61.7	0.1	3.0	17.7	44.5	2.4	11.1	33.2
	Percussions	0.1	78.2	2.6	0.0	1.6	52.4	21.9	1.2
	Strings	1.5	1.6	72.5	1.4	7.0	19.6	38.1	6.6
	Woodwinds	16.7	0.1	1.9	60.8	25.1	3.8	7.1	37.2

NH normal hearing, CI cochlear implant



**FIG. 5.** Correlations between mean pitch and timbre performances (left panel). A closer look at inter-individual differences revealed a positive trend between pitch (right panel) and timbre performance (middle panel) and the years of music training, but these factors did not reach statistical significance. The black squares denote the normal hearing participants and the white diamonds represent the cochlear implant users

profound hearing loss (pitch task—CI users  $r^2 = 0.02$ ,  $p = 0.60$ ; timbre task—CI users  $r^2 = 0.00$ ,  $p = 0.82$ ); years of profound deafness (pitch task—CI users  $r^2 = 0.00$ ,  $p = 0.91$ ; timbre task—CI users  $r^2 = 0.01$ ,  $p = 0.79$ ); and years of CI use (pitch task—CI users  $r^2 = 0.20$ ,  $p = 0.09$ ; timbre task—CI users  $r^2 = 0.05$ ,  $p = 0.44$ ). None of these correlations reached significance. Finally, we analyzed the association between music training-induced improvements in pitch thresholds or timbre  $d'$  with these four factors, but did not observe any statistically significant correlations ( $p > 0.16$ ).

### Examination of Musical Background

To further delve into the inter-individual differences relating to musical background, these measures were then correlated among four other factors: years of musical training, age at onset training, instrument practice time (in hours/week), and simply listening to music (in hours/week). None of them reached significance. The closest correlations arose for years of musical training (timbre task—NH listeners  $r^2 = 0.01$ ,  $p = 0.74$ ; timbre task—CI users  $r^2 = 0.02$ ,  $p = 0.57$ ; pitch task—NH listeners  $r^2 = 0.06$ ,  $p = 0.28$ ; pitch task—CI users  $r^2 = 0.20$ ,  $p = 0.09$ ). One could have expected subjects with more years of musical training to display higher  $d'$  in the timbre task and lower pitch thresholds. As depicted in the middle and right panel of Fig. 5, the linear trends are towards the expected direction, but they did not reach significance for these tasks.

### Examination of Gender, Pre-lingual Versus Post-lingual Deafness, and Unilateral Versus Bilateral CIs Effects

We included gender, pre-/post-lingual deafness, and uni-/bi-lateral CIs as additional between-subjects factors in the analysis reported in Table 2 for the pitch

task. None of these three factors resulted in a main effect ( $p = 0.944$ ;  $p = 0.424$ ;  $p = 0.389$ , respectively), nor did they interact with any of the other factors ( $p > 0.257$ ;  $p > 0.064$ ;  $p > 0.054$ , respectively). We also included gender in our analysis of the timbre task. Again, gender, pre-/post-lingual deafness, and uni-/bi-lateral CIs did not result in a main effect ( $p = 0.591$ ;  $p = 0.302$ ;  $p = 0.261$ , respectively), nor did it interact with any of the other factors ( $p > 0.630$ ;  $p > 0.317$ ;  $p > 0.162$ , respectively).

### Examination of Procedural Learning Effects

In this study, we refer to the procedural learning effect as learning associated with the training experience itself rather than learning devoted to the specific features of the stimulus (e.g., fine structure) or perceptual judgment (e.g., frequency discrimination versus timbre identification, task learning) (Ortiz and Wright 2009). In cases involving procedural learning, the act of completing a task multiple times allows a participant to better understand the study environment (e.g., practice session), the testing method (e.g., two-interval forced choice versus identification), the response demands, and/or develop general strategies to perform the assigned tasks.

To determine the degree to which procedural learning may have contributed to our results, we split the two test blocks that occurred in each round and reiterated an analysis of variance using two within-subject factors (replication, round) and two between-subjects factors (hearing status, arm). This was done with both pitch and timbre tasks. In addition to all the effects reported in Table 3, there was a main effect of replication ( $F(1, 28) = 6.0$ ,  $p = 0.021$  in the pitch task;  $F(1, 28) = 57.3$ ,  $p < 0.001$  in the timbre task). Subjects performed better on the second block than on the first block. In the pitch task, replication did not interact with round ( $F(2, 56) = 2.2$ ,  $p = 0.124$ ) or with

any other between-subjects factors ( $p > 0.126$ ). This means that there were immediate practice effects that were beneficial at each round, in addition to long-term effects that carried over from 1 month to the next. In the timbre task, replication interacted with round and hearing status ( $F(2, 56) = 5.4$ ,  $p = 0.007$ ): the increase in  $d'$  from the first to the second block (immediate practice benefits) was 0.26, 0.12, and 0.09 for NH subjects in rounds 1–3; while it was 0.10, 0.18, and 0.19 for CI subjects.

Post hoc tests were performed on the  $d'$  difference between the two blocks in each round, and confirmed that the benefit obtained from immediate practice effects was larger for NH listeners than CI subjects on the first round of testing ( $F(1, 28) = 5.4$ ,  $p = 0.028$ ). This effect was similar between the two groups on the second round of testing ( $F(1, 28) = 0.7$ ,  $p = 0.396$ ), but tended to be smaller for NH listeners than CI subjects on the third round of testing ( $F(1, 28) = 3.4$ ,  $p = 0.074$ ). This opposite pattern essentially reflects a delay in acquiring immediate procedural learning for CI users in the timbre identification task, relative to NH subjects. However, the improvements observed from month to month were not dependent on the population. Regardless of whatever form of training intervention occurred over a 1-month period (music intervention or listening to audiobooks), performance improved for both subject populations (NH listeners and CI users) and at somewhat similar degrees of magnitude. Based on these findings, it is indeed possible that long-term improvements observed over the 2-month study period are strongly influenced by procedural learning occurring across both tasks.

## DISCUSSION

In this study, we observed small improvements in pitch discrimination and timbre identification following three visits spanning a two-month period across both CI and NH participants. The fact that these improvements were seen in both experimental arms (i.e., whether subjects underwent music training first or underwent audiobook listening first) means that both music training and audiobook listening may be beneficial, to some degree, for auditory performance enhancement. For both NH and CI research participants, the primary benefit specific to music training (and not seen with audiobook training) is best exemplified by performance in the recognition of musical instrument class, a complex auditory task that is independent of pitch and loudness detection yet requires the listener to interpret short segments of highly dynamic spectrotemporal information. The music training program contained some exercises

related to instrument identification (but different from those used in the test), so it is plausible that CI users could have learned to use certain cues in the degraded input they receive (for example, a very sharp attack for percussions, or a high spectral centroid for brasses) via task learning. The audiobook training stimuli were limited to the spoken voice, which theoretically should not enhance detection of timbre cues. However, prosodic cues that are pervasive in spoken language may also influence pitch perception for CI users in a positive manner, potentially accounting for the benefits in pitch perception observed in both the music arms and audiobook arms. Hearing status also did not interact with round or round and arm. A possible explanation for this finding is that CI users did not become better at using auditory cues that were any different from those used by NH listeners and that these cues are unrelated to the particular spectro-temporal degradations induced by cochlear implant processors. As discussed earlier, there may have been a task learning effect where both CI users and NH listeners became more adept at determining the specific perceptual judgment to be made, as opposed to generalized procedural learning-based improvements in performance (Ahissar and Hochstein 1993; Karni and Bertini 1997; Meinhardt and Grabbe 2002).

It is worth noting that the audiobook control task we employed here may have been asymmetrically advantaged due to the fact that it employed speech stimuli, a listening category that both NH and CI users generally do well with, whereas the experimental task employed musical stimuli, a category with which both NH non-musicians and CI users generally find more difficult than speech processing. Hence, engagement may have been biased towards the audiobook arm. Taken within this context, the fact that performance improved at all in this short training period and that these improvements were relatively similar for both NH and CI subjects is encouraging. Both NH listeners with excellent spectral information and CI users with severe spectro-temporal degradations are still capable of extracting the cues necessary for learning to occur. It is indeed plausible that both hearing groups may not have learned to use all the auditory information offered to them, and that continued music rehabilitation would provide an opportunity to reinforce such learning.

These benefits are consistent with Driscoll (2012)'s published findings demonstrating that auditory rehabilitation improved timbre identification in CI users. It is important to note that the present training paradigm deviated from the previous study in that Driscoll required a longer training commitment (12 weeks), did not involve a control group or task, and trained participants directly on the tested task

(Driscoll 2012). Driscoll also acknowledged that improvement in instrument recognition could have been partially attributable to a procedural learning effect of practicing on the same items. In our timbre identification task, we did not train research participants on the same items involved in the test, and therefore, our study theoretically should be less vulnerable to a repetition effect. However, as discussed in section “Examination of procedural learning effects” of our results, we did observe a repetition effect in this study. There was a small improvement in CI users’ and NH listeners’ timbre scores (a  $d'$  gain of 0.15) after 1 month of using audiobooks, but this improvement was less than the timbre scores (a  $d'$  gain of 0.3) obtained after 1 month of music training and was only present for Arm B and not Arm A. Therefore, while part of the  $d'$  gain reflected by the main effect of music training on timbre perception may be attributed to procedural training, our study results cannot be entirely explained by this reasoning.

We find it intriguing that performance on the timbre identification task improved more dramatically than performance on the pitch task after 1 month. This perhaps suggests that training benefits (for both music and audiobook arms) are more efficient for timbre identification than for pitch sensitivity, that frequency-discrimination tasks are poor representations of musical training, or that ceiling effects for pitch perception may be observable at lower thresholds. Furthermore, there are prior training studies on NH listeners suggesting stimulus training on fundamental-frequency discrimination to be somewhat specific with only partial generalization (Wright and Zhang 2009). This explanation may partially account for why performance on the pitch task was slower to improve after a month of broad music training. It is important to note that the training that occurred in these studies (Wright and Zhang 2009) tended to focus on a single stimulus, unlike our study which observed training effects across multiple stimuli. Additional differences between the pitch and timbre tasks are the type of learning and performance requirements placed on the participant. The pitch task demands discrimination between two stimuli whereas the timbre task requires identification of the stimulus. Prior speech perception studies have found auditory discrimination to be significantly easier than identification (Blumstein and Cooper 1971). As such, the pitch discrimination task may be more susceptible to ceiling effects in comparison to the instrument identification task.

As expected, the greatest post-training gains in music perception scores are often observed for the trained task. In fact, there is evidence that difficult listening tasks (e.g., spectrally complex sound process-

ing) are more likely to require specific and spectrally complex exercises for improvement (Loebach and Pisoni 2008; Loebach et al. 2009). Whether the impact of music training may transfer to untrained tasks, particularly with speech perception, is an extremely interesting issue that lies at the core of educational neuroscience. Similarly, whether speech-based auditory rehabilitation training benefits may extend to untrained musical tasks (as suggested by the audiobook arm findings presented here) remains intriguing and requires further investigation. There are many studies that demonstrate the benefits and the potential of music training on untrained conditions such as speech processing (Tierney and Kraus 2015; Slater and Kraus 2016; Woodruff Carr et al. 2014; Kraus and Chandrasekaran 2010; Kraus et al. 2014) and familiar melody recognition (Galvin et al. 2007). On a related note, this study evaluated whether music training may improve perception scores for untrained but spectrally complex auditory tasks. We observed a small effect in pitch after both interventions and greater improvement in timbre identification after music training as compared to audiobook listening. We recognize that our training intervention demands a significant degree of generalizability and fast-learning by prioritizing the use of highly variable acoustic stimuli (Vandali et al. 2015; Moore and Amitay 2007; Boothroyd 2010; Gfeller 2001) and a relatively short-training interval. While this decision may have limited the training effects we observed, it also likely led to increased active user engagement, an important aspect of any training paradigm where patient fatigue and attrition are common barriers to success.

Some studies suggest that the benefits of auditory training are mainly due to cognitive factors such as attention, concentration, and memory (Amitay et al. 2006; Moore et al. 2009; Strait et al. 2010) rather than improved auditory perception. In order to elucidate whether CI auditory training studies are primarily due to improved auditory perception or cognitive processing, Oba et al. (2013) conducted a training study among 10 CI users using a forward visual digit span (VDS) task, a non-auditory task that targets cognitive processing. Although VDS training significantly improved VDS scores, it provided little-to-no benefit for speech and music perceptual measures, suggesting that post-training gains in CI training studies cannot be solely explained by cognitive factors and raising the possibility that training-induced bottom-up processes are involved. Although the present study was not designed to clarify the mechanisms that may account for our observations, we suspect that there is a combination of cognitive, perceptual, and procedural effects that are responsible for our findings. Further studies are needed to disambiguate the specific roles

played by each, likely requiring longer training paradigms and broader cognitive assessments.

There are several limitations to the present study. Nine participants dropped out of the study, reporting a lack of time to do the music training exercises. Their reasons for quitting were related to the required time commitment (2 h/week for 8 weeks) for study eligibility rather than the ease of use or level of enjoyment offered by the music training program. Of the remaining 17 NH subjects and 15 CI users, 11 of them did not meet the minimum of 2 h a week of music training exercises. Five NHA, two NHB, two CIA, and two CIB subjects did not complete a total of 8 h of music training exercises; these subjects reported some use of music training program (range = 0.88 to 1.89 h a week), so we decided to include their datasets in our analysis. Again, personal commitments were cited as a main reason for dropping out, highlighting the importance of using an enjoyable intervention and short training intervals that account for a trainee's lifestyle for successful rates of participation and rehabilitation (Gfeller et al. 2015). Despite the nine participant drop-outs, there seemed to be a general shift in the attitude towards music among the CI population. A CI subject summarizes her mental shift at her last visit: "I'm more open to music than I ever had been in my life. For a long time, I didn't listen to music because it didn't sound like what it used to. I would never have listened to music again. But this study and Meludia has helped me be more open to listening to music that I normally would not have in the past. And I'm grateful for that." It should be noted that these personal reports were unsolicited, and therefore may not be representative of everyone's experience with the music training exercises. It is possible that participants who did not have a positive experience either did not mention it to the study investigator or dropped out of the study. Furthermore, many of the CI research participants were prelingually deafened. As such, self-reported music processing outcomes and overall music satisfaction levels may be relatively high within this group of participants (as opposed to a study with only post-lingually deafened CI users). Nonetheless, we found testimonies like this enlightening in that they offered a glimpse of what this music training paradigm may have offered from the perspective of a CI participant regardless of the end outcome. Furthermore, it should be acknowledged that including a small group of patients that did not complete the full training paradigm in its entirety could have also limited the size of the training effects we observed here.

Another limitation to this study is related to our control intervention. Participants were not allowed to read along while listening to audiobooks. Some CI

users rely heavily on visual cues (e.g., lip reading or reading words) while listening, and we were worried that subjects would depend on visual cues rather than auditory information if given the choice (Desai et al. 2008), thereby negating the original intent of this control paradigm to serve as a non-musical listening task. Delis et al. (2000) found that lip reading predicted 42 % of the variance in 6-month post-implant, monosyllabic word scores and we wanted to eliminate this potential confound from our study analyses. The decision to not provide feedback along with the control audiobook task could contribute to some of the learning differences between the two arms (Driscoll 2012). Furthermore, it is worth noting that our study did not include a third control group that received no training at all. Therefore, while our study demonstrated benefit after audiobook listening and music training, we cannot rule out the possibility that these reported benefits may have occurred even without any training intervention, simply from taking the pitch and timbre tasks multiple times, or listening in natural environments.

In summary, we evaluated the impact of an online music training approach on pitch and timbre perception using an independent passive listening control program in CI users and NH individuals. We tested participants on the two most difficult music perception tasks for CI users: pitch discrimination and timbre identification (Limb and Roy 2014; Heng et al. 2011). In what we believe represents the most rigorous and carefully controlled music training study to date in CI users, we observed small benefits in pitch and timbre perception among both hearing populations with both training interventions. However, timbre identification scores improved to a greater degree after music training as compared to audiobook listening. We suspect that these benefits are due to a combination of training-induced stimulus learning and also procedural learning. These findings speak to the importance of appropriate outcome measures and rigorous study paradigms in future investigations involving long-term auditory training for both musical and non-musical stimuli.

## ACKNOWLEDGMENTS

We would like to thank Meludia for providing access to the online music training accounts for the research participants in this music training study.

## FUNDING

This present study did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

## COMPLIANCE WITH ETHICAL STANDARDS

*Conflicts of Interest* CJL serves as a member of the Medical Advisory Board and receives research funding and support from Advanced Bionics Corporation. He is also a consultant and has served as Scientific Chair of the Music Advisory Board for Med-El Corporation. He is also a consultant for Oticon Medical, Spiral Therapeutics, and Frequency Therapeutics. None of these relationships were relevant to this study. For the remaining authors, no conflicts of interests were declared.

## REFERENCES

- AHISSAR M, HOCHSTEIN S (1993) Attentional control of early perceptual learning. *PNAS* 90(12):5718–5722
- AMITAY S, IRWIN A, MOORE DR (2006) Discrimination learning induced by training with identical stimuli. *Nat Neurosci* 9(11):1446–1448
- BARRETT KC, ASHLEY R, STRAIT DL ET AL (2013) Art and science: how musical training shapes the brain. *Front Psychol* 4:713
- BIDELMAN GM, GANDOUR JT, KRISHNAN A (2011) Musicians and tone-language speakers share enhanced brainstem encoding but not perceptual benefits for musical pitch. *Brain Cogn* 77(1):1–10
- BLUMSTEIN S, COOPER W (1971) Identification versus discrimination of distinctive features in speech perception. *Q J Exp Psychol A* 24(2):207–214
- BOOTHROYD A (2010) Adapting to changed hearing: the potential role of formal training. *J Am Acad Audiol* 21:601–611
- BROWN CJ, GFELLER K, ABBAS P ET AL (2014) Musical training: effects on perception and electrophysiologic measures of discrimination. In: American Auditory Society Scientific and Technology Meeting, Scottsdale, Arizona
- BUSBY PA, ROBERTS SA, TONG YC, CLARK GM (1991) Results of speech perception and speech production training for three prelingually deaf patients using a multiple-electrode cochlear implant. *Br J Audiol* 25(5):291–302
- CHEN JK, CHUANG AY, McMAHON C ET AL (2010) Music training improves pitch perception in prelingually deafened children with cochlear implants. *Pediatrics* 125:e793–e800
- DAWSON PW, CLARK GM (1997) Changes in synthetic and natural vowel perception after specific training for congenitally deafened patients using a multichannel cochlear implant. *Ear Hear* 18(6):488–501
- DELIS DC, KRAMER JH, KAPLAN E ET AL (2000) California verbal learning test. 2. The Psychological Corporation, San Antonio
- DEROCHE ML, LU HP, LIMB CJ ET AL (2014) Deficits in the pitch sensitivity of cochlear-implanted children speaking English or Mandarin. *Front Neurosci* 8:282
- DESAI S, STICKNEY G, ZEN FG (2008) Auditory-visual speech perception in normal-hearing and cochlear-implant listeners. *J Acoust Soc Am* 123(1):428–440
- DRISCOLL VD (2012) The effects of training on recognition of musical instruments by adults with cochlear implants. *Semin Hear* 33(4):410–418
- FU QJ, GALVIN JJ III (2003) The effects of short-term training for spectrally mismatched noise-band speech. *J Acoust Soc Am* 113(2):1065–1072
- FU QJ, GALVIN JJ III, WANG X, NOGAKI G (2005) Moderate auditory training can improve speech performance of adult cochlear implant users. *Acoust Res Lett Online* 6(3):106–111
- GALVIN JJ III, FU QJ, OBA S (2008) Effect of instrument timbre on melodic contour identification by cochlear implant users. *J Acoust Soc Am* 124:189–195
- GALVIN JJ III, FU QJ, NOGAKI G (2007) Melodic contour identification by cochlear implant listeners. *Ear Hear* 28(3):302–319
- GALVIN JJ III, FU QJ, SHANNON RV (2009) Melodic contour identification and music perception by cochlear implant users. *Ann N Y Acad Sci* 1169:518–533
- GFELLER K (2001) Aural rehabilitation of music listening for adult cochlear implant recipients: addressing learner characteristics. *Music Ther Perspect* 19:88–95
- GFELLER K, WITT S, KIM K, ADAMEK M, COFFMAN D (1999) Preliminary report of a computerized music training program for adult cochlear implant recipients. *J Acad Rehabil Audiol* 32:11–27
- GFELLER K, WITT S, STORDAHL J, MEHR M, WOODWORTH G (2000) The effects of training on melody recognition and appraisal by adult cochlear implant recipients. *JARA* 33:115–138
- GFELLER K, TURNER C, MEHR M ET AL (2002A) Recognition of familiar melodies by adult cochlear implant recipients and normal-hearing adults. *Cochlear Implants Int* 3(1):29–53
- GFELLER K, WITT S, WOODWORTH G ET AL (2002B) Effects of frequency, instrumental family, and cochlear implant type on timbre recognition and appraisal. *Ann Otol Rhinol Laryngol* 111(4):349–356
- GFELLER K, WITT S, ADAMEK M ET AL (2002C) Effects of training on timbre recognition and appraisal by postlingually deafened cochlear implant recipients. *J Am Acad Audiol* 13(3):132–145
- GFELLER K, OLESON J, KNUTSON JF ET AL (2008) Multivariate predictors of music perception and appraisal by adult cochlear implant users. *J Am Acad Audiol* 19(2):120–134
- GFELLER K, JIANG D, OLESON JJ ET AL (2010) Temporal stability of music perception and appraisal scores of adult cochlear implant recipients. *J Am Acad Audiol* 21(1):28–34
- GFELLER K, GUTHE E, DRISCOLL V ET AL (2015) A preliminary report of music-based training for adult cochlear implant users: rationales and development. *Cochlear Implants Int* 16(Suppl 3):S22–S31
- HENG J, CANTARERO G, ELHILALI M ET AL (2011) Impaired perception of temporal fine structure and musical timbre in cochlear implant users. *Hear Res* 280(1–2):192–200
- HERHOLZ SC, ZATORRE RJ (2012) Musical training as a framework for brain plasticity: behavior, function, and structure. *Neuron* 76(3):486–502
- KAGANOVICH N, KIM J, HERRING C ET AL (2013) Musicians show general enhancement of complex sound encoding and better inhibition of irrelevant auditory change in music: an ERP study. *Eur J Neurosci* 37(8):1295–1307
- KANG R, NIMMONS GL, DRENNAN W ET AL (2009) Development and validation of the University of Washington Clinical Assessment of Music Perception Test. *Ear Hear* 30(4):411–418
- KARNI A, BERTINI G (1997) Learning perceptual skills: behavioral probes into adult cortical plasticity. *Curr Opin Neurobiol* 7(4):530–535
- KRAUS N, CHANDRASEKARAN B (2010) Music training for the development of auditory skills. *Nat Rev Neurosci* 11:599–605
- KRAUS N, SLATER J, THOMPSON EC ET AL (2014) Music enrichment programs improve the neural encoding of speech in at-risk children. *J Neurosci* 34(36):11913–11918
- LANEAU J, WOUTERS J, MOONEN M (2006) Improved music perception with explicit pitch coding in cochlear implants. *Audiol Neurootol* 11(1):38–52
- LIMB CJ (2006) Cochlear implant-mediated perception of music. *Curr Opin Otolaryngol Head Neck Surg* 14(5):337–340
- LIMB CJ, ROY AT (2014) Technological, biological, and acoustical constraints to music perception in cochlear implant users. *Hear Res* 308:13–26



- LOEBACH JL, PISONI DB (2008) Perceptual learning of spectrally degraded speech and environmental sounds. *J Acoust Soc Am* 123:1126–1139
- LOEBACH JL, PISONI DB, SVIRSKY M (2009) Transfer of auditory perceptual learning with spectrally reduced speech to speech and nonspeech tasks: implications for cochlear implants. *Ear Hear* 30:662–674
- LOOI V, GFELLER K, DRISCOLL V (2012) Music appreciation and training for cochlear implant recipients: a review. *Semin Hear* 33(4):307–334
- McDERMOTT HJ (2004) Music perception with cochlear implants: a review. *Trends Amplif* 8(2):49–82
- MEINHARDT G, GRABBE Y (2002) Attentional control in learning to discriminate bars and gratings. *Exp Brain Res* 142(4):539–550
- MICHEYL C, DELHOMMEAU K, PERROT X ET AL (2006) Influence of musical and psychoacoustical training on pitch discrimination. *Hear Res* 219(1–2):36–47
- MOORE DR, AMITAY S (2007) Auditory training: rules and applications. *Semin Hear* 28:99
- MOORE DR, HALLIDAY LF, AMITAY S (2009) Use of auditory learning to manage listening problems in children. *Phil Trans R Soc B* 364(1515):409–420
- MUSACCHIA G, STRAIT D, KRAUS N (2008) Relationships between behavior, brainstem and cortical encoding of seen and heard speech in musicians and non-musicians. *Hear Res* 241(1):34–42
- NARDO WD, CANTORE I, CIANFRONE F ET AL (2007) Differences between electrode-assigned frequencies and cochlear implant recipient pitch perception. *Acta Otolaryngol* 127(4):370–377
- NOGAKI G, FU QJ, GALVIN JJ III (2007) Effect of training rate on recognition of spectrally shifted speech. *Ear Hear* 28(2):132–140
- OKA SI, GALVIN JJ 3RD, FU QJ (2013) Minimal effects of visual memory training on auditory performance of adult cochlear implant users. *J Rehabil Res Dev* 50(1):99–110
- ORTIZ J, WRIGHT B (2009) Contributions of procedure and stimulus learning to early, rapid perceptual improvements. *J Exp Psychol Hum Percept Perform* 35(1):188–194
- PANTEV C, ROBERTS LE, SCHULZ M ET AL (2001) Timbre-specific enhancement of auditory cortical representations in musicians. *Neuroreport* 12(1):169–174
- PATEL AD (2011) Why would musical training benefit the neural encoding of speech? The OPERA hypothesis. *Front Psychol* 2:142
- PETERSEN B, MORTENSEN MV, HANSEN M ET AL (2012) Singing in the key of life: a study on the effects of musical ear training after cochlear implantation. *Psychomusicology: Music, Mind, and Brain* 22(2):134–151
- REISS LA, TURNER CW, ERENBERG SR ET AL (2007) Changes in pitch with a cochlear implant over time. *J Assoc Res Otolaryngol* 8(2):241–257
- ROSEN S, FAULKNER A, WILKINSON L (1999) Adaptation by normal listeners to upward spectral shifts of speech: implications for cochlear implants. *J Acoust Soc Am* 106(6):3629–3636
- SHAHIN A, BOSNYAK DJ, TRAINOR LJ ET AL (2003) Enhancement of neuroplastic P2 and N1c auditory evoked potentials in musicians. *J Neurosci* 23(13):5545–5552
- SHANNON RV, FU QJ, GALVIN JJ III (2004) The number of spectral channels required for speech recognition depends on the difficulty of the listening situation. *Acta Otolaryngol Suppl* 552(552):50–54
- SLATER J, KRAUS N (2016) The role of rhythm in perceiving speech in noise: a comparison of percussionists, vocalists and non-musicians. *Cogn Process* 17(1):79–87
- STERNBERG RJ (1999) *Cognitive psychology*. Harcourt Brace & Compnay, Orlando
- STRAIT DL, KRAUS N, PARBERY-CLARK A, ASHLEY R (2010) Musical experience shapes top-down auditory mechanisms: evidence from masking and auditory attention performance. *Hear Res* 261(1–2):22–29
- STRAIT DL, CHAN K, ASHLEY R ET AL (2012) Specialization among the specialized: auditory brainstem function is tuned in to timbre. *Cortex* 49(3):360–362
- TERVANIEMI M, JUST V, KOELSCH S ET AL (2005) Pitch discrimination accuracy in musicians vs nonmusicians: an event-related potential and behavioral study. *Exp Brain Res* 161(1):1–10
- TIERNEY A, KRAUS N (2015) Evidence for multiple rhythmic skills. *PLoS One* 10(9):e0136645
- VANDALI AE, SUCHER C, TSANG DJ ET AL (2005) Pitch ranking ability of cochlear implant recipients: a comparison of sound-processing strategies. *J Acoust Soc Am* 117(5):3126–3138
- VANDALI A, SLY D, COWAN R, VAN HOESEL R (2015) Training of cochlear implant users to improve pitch perception in the presence of competing place cues. *Ear Hear* 36(2):e1–e13
- WOODRUFF CARR K, WHITE-SCHWOCH T, TIERNEY AT, STRAIT DL, KRAUS N (2014) Beat synchronization predicts neural speech encoding and reading readiness in preschoolers. *Proc Natl Acad Sci U S A* 111(40):14559–14564
- WRIGHT B, ZHANG Y (2009) A review of the generalization of auditory learning. *Philos Trans R Soc Lond Ser B Biol Sci* 364(1515):301–311

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.