Early Childhood Music Training and Associated Improvements in Music and Language Abilities

Sean Hutchins
The Royal Conservatory of Music, Toronto, Canada

Adult musicians tend to outperform nonmusicians in a variety of language and language-relevant tasks. Moreover, children who take music lessons will show an increased ability in several types of language skills. However, it is still unclear what the time course of these developmental effects might be, and the degree to which young children improve in their musical abilities. Here, we present the first year of data from an ongoing longitudinal study, aimed at finding if measurable improvements in musical and linguistic abilities can be seen among children taking music classes. We studied 90 children (age 3–6) who were enrolled to take group classes in a conservatory setting. We measured their musical, language, and perceptual abilities both at the beginning and the ending of the school year. Pre vs. post comparisons showed an increase in vocabulary size, pre-reading skills, and singing ability; these increases were beyond what could be attributed to normal development during the time. We also found that singing ability was correlated with language skills. Taken together, these results show that early childhood music training can lead to associated improvements in both musical skills and language skills, strengthening the evidence for a developmental link between these two abilities.

Received: June 21, 2017, accepted January 20, 2018.

Key words: music training, language, transfer, vocabulary, phonological processing

Music has long been considered to be closely related to intelligence. Although many of the hyperbolic claims of the mass media about Mozart turning babies into "Einsteins" can be easily dismissed, research from the past two decades has shown that there does seem to be a convincing link between music *training* and certain types of cognitive abilities. A good number of studies have shown that musicians (defined as people who have many years of music training, regardless of their current profession) perform better at a variety of different tasks that are not immediately related to music. For example, musicians outperform nonmusicians at several different subcomponents of executive function (Miyake et al., 2000), such as task switching (Hanna-Pladdy & MacKay, 2011; Moradzadeh, Blumenthal, & Wiseheart, 2015), response inhibition (Bialystok & DePape, 2009; Strait, Kraus, Parbery-Clark, & Ashley, 2010), and working memory (George & Coch, 2011; Oechslin, Ville, Lazeyras, Hauert, & James, 2013; Pallesen et al., 2010; Slevc, Davey, Buschkuehl, & Jaeggi, 2016). These executive functions have been thought to underlie the posited relationship between music lessons and IQ (Degé, Kubicek, & Schwarzer, 2011; Schellenberg, 2004, 2006). The cognitive differences between musicians and nonmusicians are reflected in neurological measurements as well, including structural differences (e.g., Bermudez, Lerch, Evans, & Zatorre, 2009; Gaser & Schlaug, 2003) and functional and electrophysiological measurements (George & Coch, 2011; Koelsch, Fritz, Schulze, Alsop, & Schlaug, 2005; Parbery-Clark, Skoe, & Kraus, 2009; Zuk, Benjamin, Kenyon, & Gaab, 2014).

However, the relationship between cognitive functions and music training is not as straightforward as it might seem. Though most studies published on the matter show an effect of musical background, many show only a limited benefit of music training. In particular, the evidence is inconclusive as it relates to many of the farthest cases of transfer, such as transfer across modalities. Visual tasks are less likely to show an advantage for musicians (Ho, Cheung, & Chan, 2003; Parbery-Clark, Strait, Anderson, Hittner, & Kraus, 2011), though they are reported in some studies (Oechslin et al., 2013; Schellenberg, 2006). In contrast, musicians seem to have an advantage when cognitive skills are assessed through the auditory modality (Bialystok & DePape, 2009; Bidelman, Hutka, & Moreno, 2013; Chan, Ho, & Cheung, 1998; Ho et al., 2003; Pallesen et al., 2010; Parbery-Clark et al., 2011).

Music and Language

However, it is not simply measurements of executive function that show an advantage for musicians. One area in particular that seems to show a consistent

Music Perception, volume 35, issue 5, pp. 579–593, issn 0730-7829, electronic issn 1533-8312. DOI: https://doi.org/10.1525/mp.2018.35.5.579

musician advantage is language skills. A considerable number of studies have shown that musicians outperform nonmusicians in many different subcomponents of language skill (e.g., Chan et al., 1998; Schellenberg, 2006; Slevc & Miyake, 2006; Thompson, Schellenberg, & Husain, 2004). In children, music training has been linked to larger vocabularies (Forgeard, Winner, Norton, & Schlaug, 2008; Moreno et al., 2011), better expressive grammar abilities (Gordon, Shivers, et al., 2015), better comprehension of emotional prosody (Thompson et al., 2004), second language learning ability (Patscheke, Degé, & Schwarzer, 2016; Swaminathan & Gopinath, 2013), and better reading ability (Anvari, Trainor, Woodside, & Levy, 2002; Goswami, 2011).

Of all of these, the language subcomponent that has shown the most reliable musician advantage is phonological processing (see Gordon, Fehd, & McCandliss, 2015, for a recent meta-analysis). Many studies with children have shown music training to be associated with better phonological processing skills (e.g., Degé & Schwarzer, 2011; Herrera, Lorenzo, Defior, Fernandez-Smith, & Costa-Giomi, 2011; Overy, 2003; Tsang & Conrad, 2011), and a recent study in our lab has shown a similar finding among adult musicians. Even in children without music training, basic musical skills have been associated with phonological abilities (Goswami, Huss, Mead, Fosker, & Verney, 2013; Huss, Verney, Fosker, Mead, & Goswami, 2011; Lamb & Gregory, 1993; Loui, Kroog, Zuk, Winner, & Schlaug, 2011); there is some evidence that this may be particularly linked with pitch or rhythm skills (Anvari et al., 2002; Huss et al., 2011; Loui et al., 2011; Moritz, Yampolsky, Papadelis, Thomson, & Wolf, 2013; Woodruff Carr, White-Schwoch, Tierney, Strait, & Kraus, 2014).

Although the mechanism for this transfer is not completely clear, one likely candidate comes from the OPERA hypothesis (Patel, 2011, 2014). This hypothesis suggests that the neural architecture supporting speech perception overlaps with that supporting musical activities, and that the greater precision required by musical processing, augmented by other factors inherent to music, drives neural plasticity that helps improve speech processing as well. This greater spectrotemporal processing ability can lead to better phonological ability, which itself can lead to improvements in vocabulary, reading ability, and other linguistic skills (Tallal & Gaab, 2006).

This line of reasoning is supported by a rich body of work showing improved spectrotemporal processing abilities in musicians. Behaviorally, musicians show increased ability to hear speech in a noisy context (Bidelman & Krishnan, 2010; Parbery-Clark, Skoe, & Kraus, 2009; Parbery-Clark, Skoe, Lam, & Kraus, 2009); this holds for musicians of all ages, from youth (Anderson, Skoe, Chandrasekaran, Zecker, & Kraus, 2010) to elderly adults (Parbery-Clark et al., 2011; Zendel & Alain, 2012). They also show greater accuracy in perceiving subtle timing differences (Gaab et al., 2005; Parbery-Clark et al., 2011) and more accurate pitch perception (Hutchins & Peretz, 2012; Kishon-Rabin, Amir, Vexler, & Zaltz, 2011), including vocal pitch perception (Hutchins & Peretz, 2012; Hutchins, Roquet, & Peretz, 2012). These abilities are also reflected in neurological measurements. The mismatch negativity response to changes in vowels, for example, is heightened in musically trained children (Chobert, François, Velay, & Besson, 2014; Chobert, Marie, François, Schön, & Besson, 2011). The brainstem frequency-following response, too (Bidelman, Gandour, & Krishnan, 2011; Bidelman, Krishnan, & Gandour, 2011; Lee, Skoe, Kraus, & Ashley, 2009; Parbery-Clark, Skoe, & Kraus, 2009), more closely tracks the time and frequency domain changes in an auditory signal in musicians than nonmusicians. This is believed to be one of the underlying mechanisms that facilitates musicians' improved behavioral performance.

These benefits of music training can also be shown to extend to less-advantaged groups. In a series of studies in collaboration with the Harmony Project, a program that provides free music education to children from low-income communities, Kraus and colleagues have demonstrated that children ages 6–9 who began to learn an instrument showed increases in both behavioral (speech in noise) and neurological (brainstem response) measurements (Kraus, Hornickel, Strait, Slater, & Thompson, 2014; Kraus, Slater, et al., 2014; Slater et al., 2015). Spanish-English bilingual children in this program were less likely to fall behind age-normalized reading levels than a matched control group, a common phenomenon in this group (Slater et al., 2014). Moreover, their results indicated that higher levels of engagement in the music training were associated with both neurological and behavioral improvements (Kraus, Hornickel, et al., 2014).

Children with language difficulties, too, may be aided by music training. Both dyslexia and SLI have been argued to arise from core auditory deficits in spectrotemporal processing (see Goswami, 2011, for a review), the same ability that is generally increased in musicians. Dyslexia is often coincident with musical impairments, such as poor rhythmic entrainment (Huss et al., 2011; Thomson, Fryer, Maltby, & Goswami, 2006; Thomson & Goswami, 2008; Wolff, 2002). Musical interventions have been shown to help dyslexics in some cases (Bhide, Power, & Goswami, 2013; Habib et al., 2016; Overy, 2003), though the existence of musicians with dyslexia (Bishop-Liebler, Welch, Huss, Thomson, & Goswami, 2014; Zuk et al., 2017) indicates that music training alone is not sufficient to overcome dyslexia. These dyslexic musicians were, however, found to have better core auditory abilities and slightly better phonological abilities than a control group of dyslexics without music training.

STARTING EARLY

Given this musician advantage for language and language-relevant processing, a natural question to ask is when does this advantage begin? How much training is required to see this benefit? In addition, musicians start their studies at different ages—many start well before school age—while other equally accomplished musicians do not begin their training until much later in adolescence. Is there a difference between early- and late-trained musicians? Some research has tried to address these questions.

Several studies show evidence that neurological structural differences between musicians and nonmusicians are mediated by the duration of the musicians' training (Bermudez et al., 2009; Elbert, Pantev, Wienbruch, Rockstroh, & Taub, 1995; Foster & Zatorre, 2010; Schlaug, Jäncke, Huang, Staiger, & Steinmetz, 1995). Moreover, these structural and performance differences have also been linked to the age of training onset: Early-trained musicians outperform late-trained musicians in music-relevant tasks (such as synchronization), even after controlling for total duration of training (Bailey & Penhune, 2010; Bailey, Zatorre, & Penhune, 2013; Steele, Bailey, Zatorre, & Penhune, 2013). Absolute pitch (AP, commonly referred to as perfect pitch), too, shows a strong effect of age of training onset, with the large majority of AP possessors having begun their training before the age of 6, and those beginning earlier being more likely to develop the ability (Baharloo, Johnston, Service, Gitschier, & Freimer, 1998; Takeuchi & Hulse, 1993).

This early music training also seems to affect the development of language skills. A meta-analysis of 13 studies showed a significant, positive effect of music training on the development of phonological awareness skills in children age 4-9 (Gordon, Fehd, & McCandliss, 2015). A model based on this analysis concluded that approximately 40 hours of training were needed to show significant effects of music training (though the authors urged caution in its interpretation, due to small sample sizes). Thus, it seems likely that music training can affect language skills from the very beginning.

These experiments give some useful clues as to the timeline of language-relevant benefits from music training. However, one of the more compelling ways to study this issue would be to use a longitudinal design, tracking young children's linguistic and musical skills as they progress through several years of study. Longitudinal designs pose certain problems for researchers, however. In addition to the cost and the difficulty of maintaining consistent training and contact across years, it can be difficult to find experimental tests and materials that are appropriate for a wide age group, to enable comparisons. Another major problem is the controls for testing. Most studies of the effects of music training include a control group that either receives an alternative type of training, or no contact. This allows the researchers to compare the two groups, as a control for normal development. The control group is asked to not participate in extracurricular musical activities during this time. However, in a longitudinal study, this becomes problematic. The desires and situations of children and their caregivers can change dramatically over the course of several years, and to maintain a large group of children who are directed to not take music classes over the course of 3-5 years seems ethically problematic. In addition, dropout rates can be high over the course of even one year of such a study, and the number and types of families that drop out from a control group would likely be quite different from those dropping from the music training group.

AIMS AND LONGITUDINAL DESIGN

In this study, we aim to find if measurable improvements in musical and linguistic abilities can be seen among children taking music classes. The current manuscript examines the first year of results from our ongoing longitudinal study, where we examine the change over the course of one school year. We hypothesize that during this timeframe, we should see measurable benefits to musical abilities (tapping and singing accuracy), and small but measurable benefits to linguistic abilities. These improvements should be independent of aging, as would be shown by improvement to the standardized scores when available, and a lack of interaction with age for those tests without standardized scores. We also hypothesize that there should be a relationship between some musical abilities and language related skills.

This current study is part of a larger longitudinal study that attempts to deal with many of the problems inherent to these types of studies. Several of the logistical problems of a longitudinal design are eased due to working with a prestigious conservatory. The Royal Conservatory has been operating early childhood music education since the 1920s, and is the most well-known conservatory in Canada, located in downtown Toronto. The total enrollment in the conservatory's early childhood music education program is typically over 300 children, making it much easier to find willing participants. These classes maintain a consistent curriculum (and also lead into individualized music lessons later on, for many children), effectively controlling the types of training received, and many children continue in the same course of study from preschool up through kindergarten. Thus, there is a willing and stable base of children and families motivated to continue music classes for several years.

The lack of a control group is more problematic, though. Our current study deals with this in multiple ways. First, we are implementing a relatively new type of design in which incoming participants are used as controls for those who have been in the program for longer. We term this the "cascade design," and it is possible because our programs have new students beginning at various ages; there is no regimented starting age. Thus, in the longer term, newly starting 4-year-old children (for example) can act as controls for 4-year-old children who have been in the program for the prior year. Those same children can then serve as a test group during the next year for incoming 5-year-old children, as well as be compared to the 5-year-old children who have been in the program for two years, and so on. While this design does not guarantee equal numbers in each condition, and does not serve as a true experimental design, it does provide relatively similar demographic groups to compare between, while avoiding the problems of having a long-term control group forbidden from music classes.

Over the shorter term, the other major way we deal with this problem is through the use of age-standardized tests. Some of the tests we use provide age-standardized scores, based on the abilities of the general population. These are normalized to have a mean of 100 and a standard deviation of 15 at all age levels, allowing comparison across different ages. Using this, then, we can provide some control for the differences between different age groups in our study, and more importantly, for the aging of individual participants during our study, including the aging between pre and post testing sessions. Again, this does not control for 100% of the changes across the time periods, but it does help control for normal aging-related development considerably.

For those tests that do not have age-standardized scores (e.g., our in-house musical tests), we use statistical controls to help parse out the variance due to normal aging-related development and those due to other factors. Showing better performance in the post test among 3-year-old children than in the pre-test among 4-year-old children, for example, can help indicate that the former's improvement might not have been due to aging alone.

Method

In this study, we tested children ages 3–6 who were enrolled in group music classes. Tests of vocabulary, phonological processing, speech perception in noise, singing ability, and tapping ability were conducted at the beginning and the end of a school year.

PARTICIPANTS

Participants were recruited from the population of families from the greater Toronto area who had already enrolled for Smart Start classes (The Royal Conservatory's early childhood music education program). These classes are divided by age group (roughly equivalent to the school year); children in classes for ages 3-6 were eligible for the research study. One-hundred-five children (21% of the total eligible population) signed up for the study. Of these, 90 were included in the final sample (41 female, 49 male), with 23 in the 3-year-old group, 38 in the 4-year-old group, 19 in the 5-year-old group, and 10 in the 6-year-old group. The main reasons for non-inclusion were not showing up for the post test (13 children) and non-compliance of the child (2 children, both refused to wear headphones). Of these exclusions, 9 were from the youngest age group, and 4 from the second youngest age group. Children were not excluded from the analysis for performance or language issues, however. The mean age of included children was 4.55 years (range = 3.09–6.83 years) at the pre-session, and 5.13 years (range = 3.68–7.43 years) at the post-session. Of the 90 total children, 54 were reported by their primary caregiver as having at least one parent with a prior musical background, 44 had taken a prior Smart Start class, and 15 were concurrently learning an instrument. In addition, 17 were reported as speaking a primary language other than English in the home. These children began learning English at a mean age of 22 months and were fluent in English at the time of data collection, as reported by their caregivers (all 17 were rated as "good" or "very good" in their understanding of English). No caregivers indicated any hearing problems with their children. Informed consent was obtained from all parents.

As compensation for their time, families were given 10% of the program's enrollment fee (a value of approximately $50) and a certificate of recognition with their child's name on it. The children were given a small toy at the end of each session.

TRAINING

All children in this study participated in group music education classes, held at The Royal Conservatory's downtown Toronto facility. These classes lasted eight months (from September until May), with normal breaks for holidays; there were a total of 28 classes. Individual attendance records were not kept by the teachers, but other records indicate that attendance tended to be high, near 75%, with the majority of those who did not attend regularly dropping out altogether. These classes used the Smart Start curriculum, a new curriculum developed in-house. This curriculum, applicable for children ages 0–6, is intended to incorporate current understandings of cognitive psychology into the activity design. Each activity was constructed by subject-matter specialists in music or other arts and reviewed and edited by the Royal Conservatory's early childhood music specialist (Catherine West). Curriculum activities were designed to teach specific musical skill while incorporating the use of one or more cognitive domains, which included attention, memory, perception, and cognitive flexibility. Subsequent to this, each activity was reviewed by a trained cognitive neuroscientist (the author), in order to ensure that it did in fact target the appropriate cognitive domain in an (age-appropriately) challenging way. The purpose of this curriculum is to ensure that the classes maximize students' engagement and current abilities. Importantly, lessons were not designed to teach the children specific cognitive skills, but rather to engage these skills through music; the primary goal of each class remained effective music training.

Each Smart Start teacher ($N = 10$) was given training to help understand and incorporate this curriculum into their lesson planning. This training consisted of 10–30 hours of theoretical and practical work, plus an extra 6–10 hours of in-class training. The teachers, all of whom were experienced music teachers, created their own classes, each with its own structure, but were required to follow the principles of the new curriculum, and were encouraged to use specific activities from the curriculum books. The musical goals of these classes included awareness of basic musical concepts (such as beat, rhythm, pitch height, form, etc.), the development of the singing voice, greater comfort in moving to music, and experience with pitch and non-pitched percussion instruments. Older classes extended these goals to include imitation and adaptation of musical excepts and use of rudimentary music notation. These goals were supported by a variety of activities, including group and solo singing, movement activities, listening activities, activities with instruments, and musical games, as well as by free play time for independent, non-guided musical exploration.

TESTS AND MEASUREMENTS

Children in this study participated in a battery of five tests. Three of these were commercially available tests designed to test language-related abilities: the Peabody Picture Vocabulary Test, Fourth Edition (PPVT; Dunn & Dunn, 2007), the Comprehensive Test of Phonological Processing, Second Edition (CTOPP; Wagner, Torgesen, Rashotte, & Pearson, 2012), and a speech discrimination in noise (SIN) test derived from the Auditory Skills Assessment (ASA; Geffner & Goldman, 2010). The other two were musical tests, designed in house: a tapping synchronization-continuation task and a singing task.

The PPVT is a commonly used measurement of receptive vocabulary ability. It is designed for use among participants age 2.5 years and up, and generally takes 6–10 minutes to complete. Participants hear a spoken word and point to the picture (among a group of four) that best depicts that word. The results of the test are standardized, to take account of developmental changes. Average performance at each six-month age interval is normalized to 100, with a standard deviation of 15. This normalization is the result of testing with a large, nationally representative (US) sample, accounting for gender, socio-economic status, ethnicity, and geographic region. In the 3–6 age range, each six-month cohort was normed with a sample of 100–125 children; test-retest reliability in these age ranges was measured at .91–.94, indicating a highly reliable measurement. An independent test of an earlier version of the PPVT with a slightly older sample showed good stability over a longer term (11 months), with test-retest correlation at .84 and no significant differences between the two administrations (Bracken & Murray, 1984). This indicates that children's normalized scores do not tend to change significantly in retesting absent interventions.

The CTOPP is a multi-component test designed to measure different aspects of phonological abilities. Participants engage in tasks including phoneme isolation, non-word repetition, rapid letter naming, and others; the full test typically takes about 20–25 minutes to complete. The participants in our study were guided through the nine subtests of the CTOPP, which are administered orally by the experimenter and answered orally by the child. In the test as administered to children under age 6, raw scores are combined and standardized to yield four primary outcomes: Phonological Awareness, Phonological Memory, Rapid Symbolic Naming ability, and Rapid Non-Symbolic Naming ability, each of which are scaled to account for developmental changes. As in the PPVT, average performance at each age interval is normalized to 100, with a standard deviation of 15. This normalization also comes from a large representative (US) sample, accounting for gender, socio-economic status, ethnicity, and geographic region. In the 3–6 age range, each 1 year cohort was normed with a sample of 119–136 children; test-retest reliability for composite scores in these age ranges was measured at .80–.93, again indicating that children's normalized scores do not tend to change significantly in retesting.

To assess more basic perceptual issues in speech perception, we administered a modified version of the Speech Discrimination in Noise subtest, taken from the ASA. This test worked similarly to the PPVT, with children hearing a spoken word and choosing the matching picture from a set of four. Here, the words were chosen to be commonly known words, but were presented in noise (multi-talker babble). In addition, the alternate choices often differed by only one phoneme (e.g., a picture of a fan as an alternate choice for "van"). This test included ten total items. Although designed for children aged 3.5–7 years old, our pilot tests indicated that this subtest was too easy for most children. To improve the test, we amplified the dB level of the multi-talker background babble to create three different ambient noise levels. Children were initially tested at the loudest ambient noise level. If a child missed more than three of first five items, they would restart on the middle level, and again on the softest (original) level if still unable to successfully complete the task. The total number of correct answers in the final level and the ending level were recorded as scores; these were not standardized across age groups. The test typically took between 3–5 minutes to complete.

In the singing task, the child was asked to sing the alphabet song. This was chosen because it is widely known (even among children who do not yet know their alphabet per se) and has a limited range, but is not explicitly taught in the Smart Start classes. If children did not know the alphabet song, they were asked to sing "Mary Had a Little Lamb" or "Baa Baa Black Sheep" (which have nearly identical melodic structures). Children who did not know any of these songs were allowed to sing any other song they wished (including but not limited to "Happy Birthday" and the main theme from "Star Wars"); these, however were excluded from further analysis. Children were prompted by the experimenter playing the first seven notes of the song on the piano in the key of C, but were not required to sing in that key. This task lasted a total of 2–3 minutes.

Finally, in the tapping synchronization-continuation task, the child heard a metronomic beat, and was asked to tap on the table along with the beat, and then continue after the pacing beat finished. The stimulus file played a metronomic click with no volume variation and an inter-onset interval (IOI) of 500 ms (120 bps). There were 16 total clicks, with the final click 7.5 seconds after the first. They were asked to start tapping with the track as soon as they could, and keep tapping on the table in the continuation phase until cued to stop by the experimenter, who would signal this after 16 subsequent taps (generally, 8 further seconds). Children were given one practice trial, so that they would not be surprised when the click track ended. This task also lasted a total of 2–3 minutes.

PROCEDURE

All tests were administered by trained research assistants in a one-on-one basis in a private room. Visual stimuli were presented either on a laptop computer or through booklets. Auditory stimuli were presented over Sony MDR-G45LP or LilGadgets 151120 Connect+ headphones and two M-Audio Studiophile AV 40 external speakers (tapping synchronization-continuation task only). Tapping and singing responses were recorded with a Sennheiser MKE-2 microphone, placed on a desk near the child. The entire procedure took about 45–60 minutes, and was administered in one sitting. The same battery of tests was administered in the same order (PPVT, SIN, CTOPP, singing, tapping) during both pre and post sessions. Prior to or during their child's pre-test session, parents also filled out a questionnaire about their family's musical and language background, and the short version of the Inventory of Children's Individual Differences to assess their child's personality based on the big 5 personality traits (Openness, Conscientiousness, Extraversion, Agreeableness, Neuroticism); this was done to check whether individual differences in personality might mediate links between cognitive and musical abilities, as suggested by Corrigall, Schellenberg, and Misura (2013).

DESIGN AND ANALYSIS

Standardized scores were used as primary outcome measurements for the PPVT and the CTOPP, one from the PPVT and four (described above) from the CTOPP. These are each designed to have a mean of 100 and a standard deviation of 15, across the general population, for each age (calculated to the month for PPVT, and to six month intervals for the CTOPP). The SIN score was calculated as the number of correct responses, plus ten for each lower level avoided (this assumes that lower levels would be answered correctly; the range of possible scores is 0–30). This score is not standardized.

The age standardizations of the PPVT and CTOPP account for the expected changes from normal aging, and the test-retest reliabilities suggest that a cohort of children will tend to retain the same mean standardized score over the time between our pre and post testing sessions (though their raw, non-standardized scores will tend to increase). Thus, no correlation with age is expected for the PPVT and CTOPP scores, whereas it may be expected for the SIN score and the musical measurements, which are not age-standardized in this way.

The singing task was evaluated for pitch accuracy only, from children who sang the alphabet song or a song with the same melody. We used Melodyne (Celemony Software GmbH, Munich Germany) to evaluate the pitch of each sung note, over the first phrase of the song (sixteen notes, up to "p"). Because of normal blending between syllables in this song (which is characteristic of both children and adults, as in "elemeno" for "L-M-N-O"), we treated repeated pitches as one note, yielding eight total notes. The seven intervals formed by this were evaluated and compared to the ideal melodic intervals, which yields an error measurement for each sung interval. Our primary measurement of pitch accuracy was the mean absolute value of each interval error (so that positive and negative errors do not cancel each other out). As supplementary measures, we also calculated the total number of mistuned intervals (defined as greater than 50 cents of error; see Hutchins & Peretz, 2012; Pfordresher & Mantell, 2014) and the total number of contour errors (a sung interval in the wrong direction; i.e., moving up in pitch where the melody should be moving down, or vice versa). Note that these measurements are independent of key; if the child chose to sing in a different key, these measurements do not consider that an error. In instances where the pitch of the opening phrase could not be evaluated (typically because the child sang along with the piano), the final phrase was used instead, as it has an identical melodic structure. The same phrase was used in both pre and post measurements, whenever possible.

To evaluate the tapping task, we applied Audacity's (Audacity Team, Version 2.0.4) automatic sound finder algorithm to locate tapping onsets from the recordings. These tapping onsets were adjusted by hand when necessary. We used the onsets to calculate the IOI between each tap. We used the IOIs to calculate the mean IOI error during the synchronization and during the continuation phases as our primary measurements, and the standard deviation of the IOI and the mean asynchrony (during the synchronization phase only) as our supplementary measurements. To account for discrepancies in the total number of taps in each phase, the time to entrain to the rhythm, and tempo drift in the continuation phase, we used only the eight IOIs immediately before the end of the synchronization phase, and eight in the continuation phase. The three IOIs immediately following the end of the metronome were not used, as pilot testing indicated that some children were disrupted by this change, especially amongst the younger age groups.

Children who did not complete a task in either pre or post sessions were only omitted from analyses of that task, and not others; most children completed all tasks. Four children did not complete the tapping task, and 15 did not complete the singing task. In addition, 36 children were outside of the standardized age range for the CTOPP at either pre or post testing sessions; we extrapolated from CTOPP scoring principles to apply standardization to age ranges not covered by the test, and this showed no major differences in outcomes compared with the exclusion of these data.

Results

Our primary outcomes were evaluated with paired samples *t*-tests, comparing pre vs. post sessions. For those measurements that were not standardized, a follow-up analysis included age group as an independent variable. We applied a Bonferroni correction to the paired samples *t*-tests to account for the nine independent observations,[1] lowering the critical threshold to .006.

Table 1 shows the results for the nine main variables of interest. There were significant effects of Session in PPVT scores, Rapid Symbolic Naming, Rapid Non-Symbolic Naming, and mean singing error. All of these effects showed significantly better performance in the post session than in the pre session. The test scores of the children in our sample indicated improvements in vocabulary ability, improvements in the ability to name letters, numbers, pictures, and colors, and improvement in the ability to sing in tune. Importantly, all of these improvements but singing ability are from age-standardized measurements, indicating that the improvements are relative to age-normed abilities. A follow-up mixed measures ANOVA on the singing error measurements included Session as a within-subjects variable and Age Group as a between-subjects variable. This analysis found a significant main effect of Session, $F(1, 71) = 11.39$, $p = .001$, $\eta^2_p = .14$, but no main effect of age group and no interaction. This result (shown in Figure 1) indicates that normal development is likely not

[1] Measurements such as mean interval error and number of mistuned intervals, for example, are not independent of each other, but both show interesting views of the child's performance.

TABLE 1. *Means (Standard Deviations) and Statistical Analyses for the Nine Main Measurements, Comparing Pre and Post Tests*

Task	DV	Mean at Pre (*SD*)	Mean at Post (*SD*)	*t*	*df*	*p*	Cohen's *d*
PPVT	Vocabulary Score	109.4 (14.4)	113.0 (13.3)	3.55	89	**.001**	0.38
CTOPP	Phonological Awareness Score	112.2 (10.5)	111.4 (13.2)	−0.73	89	.47	0.08
CTOPP	Phonological Memory Score	114.7 (14.5)	116.2 (14.3)	−0.93	89	.36	0.10
CTOPP	Rapid Symbolic Naming Score	101.6 (17.4)	108.1 (12.6)	3.82	79	< **.001**	0.44
CTOPP	Rapid Non-Symbolic Naming Score	93.0 (16.6)	98.3 (18.5)	3.05	88	**.003**	0.32
ASA	Speech in Noise Score	21.9 (6.8)	21.9 (6.3)	0.07	89	.95	0.01
Singing	Mean Interval Error (cents)	91.7 (67.8)	63.6 (48.7)	−3.21	74	**.002**	0.38
Tapping	Rate Error- Synchronization (ms)	35 (39.0)	29 (38.0)	−1.01	84	.32	0.11
Tapping	Rate Error- Continuation (ms)	62 (54.0)	56 (51.0)	−0.72	84	.47	0.08

Note: Boldface indicates a significant difference between pre and post.

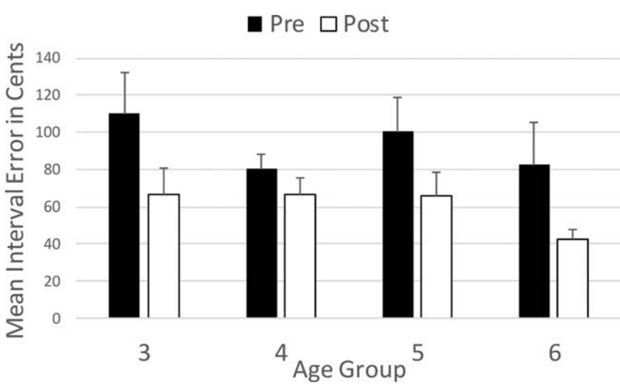

FIGURE 1. The mean interval error in cents for the singing task at pre and post tests for children in each age group, with standard error bars. Higher values indicate less accurate singing.

responsible for the improvement in singing abilities. Another set of supplementary analyses used mixed measures ANOVAs comparing PPVT and CTOPP scores of children who spoke a primary language other than English at home to those who did not, across both sessions. Children whose primary home language was English scored higher on the PPVT than those whose primary home language was not English, $F(1, 88) = 18.58$, $p < .001$, $\eta^2_p = .17$. However, no other main effects of home language were found, nor was any interaction found between home language and Session.

Supplementary measurements of production data showed largely the same patterns as the main results. None of the tapping measurements showed significance, indicating that the lack of improvement in tapping accuracy was not compensated for by improvements in variability or asynchrony. In the singing measurements, a significant improvement was found in the number of mistuned intervals, $t(74) = 2.83$, $p = .006$, Cohen's $d = 0.33$, indicating that the decreases in interval error likely led to perceptible improvements in tuning. However, no main effects were found in the number of contour errors. This is likely due to the overall low prevalence of contours errors in the first place.

As a metric for reliability, we also examined the correlations between pre and post measurements for each of our primary measurements. All measurements from the PPVT, CTOPP and the SIN measurement were highly reliable, with significant pre-post correlations. The singing accuracy data proved somewhat reliable, with a (non-significant) correlation of $r(75) = .184$ (supplemented by a marginally significant correlation of $r(75) = .194$, $p < .10$, for the number of mistuned intervals measurement). The tapping data proved less reliable, with a pre-post correlation of $r(85) = .011$ for the tapping accuracy with the metronome and $r(85) = .17$ for tapping accuracy after the metronome.

To examine the relationships between the different measured variables, we ran a set of correlations between the nine dependent measures, as well as age and the five personality measurements. To capture pre and post together, we used their mean value [(pre + post) / 2]. While the personality measurements showed strong intercorrelations, they did not prove to correlate significantly with the other variables (with the sole exception being a significant correlation between extraversion and the tapping accuracy after the metronome, $r(82) = .25$, $p = .025$; thus, these were omitted from the table. The correlation matrix is shown in Table 2.

In this sample, the age of the participant was correlated with performance on the SIN and on both primary tapping measurements. These correlations, in conjunction with the lack of correlations between pre and post measurements, indicate that age, rather than experience or training, was the primary driver of ability on these two tasks. SIN performance and both tapping measurements were also correlated with each other. Many of the components of the CTOPP were intercorrelated, and were correlated with the PPVT as well. This is not surprising, as they each measure different components of

TABLE 2. *Correlation Matrix for the Nine Main Measurements and Age*

	Age	PPVT	Phonological Awareness	Phonological Memory	Rapid Symbolic Naming	Rapid Non-Symbolic Naming	Speech in Noise	Singing Error	Tapping Error (Synch)	Tapping Error (Cont)
Age	–	.05 *ns*	−.02 *ns*	.19 *ns*	.09 *ns*	**.25** $\mathbf{p = .02}$	**.52** $\mathbf{p < .001}$	−.08 *ns*	**−.34** $\mathbf{p = .001}$	**−.26** $\mathbf{p = .02}$
PPVT		–	**.37** $\mathbf{p = .001}$	**.48** $\mathbf{p < .001}$	.15 *ns*	.14 *ns*	**.34** $\mathbf{p = .001}$	.07 *ns*	−.19 *ns*	−.01 *ns*
Phonological Awareness			–	**.46** $\mathbf{p < .001}$	**.47** $\mathbf{p < .001}$	.17 *ns*	.10 *ns*	**−.24** $\mathbf{p = .04}$	−.13 *ns*	−.02 *ns*
Phonological Memory				–	**.39** $\mathbf{p < .001}$	.16 *ns*	**.47** $\mathbf{p < .001}$	−.18 *ns*	−.20 *ns*	**−.27** $\mathbf{p = .01}$
Rapid Symbolic Naming					–	**.59** $\mathbf{p < .001}$	**.24** $\mathbf{p = .03}$	**−.41** $\mathbf{p < .001}$	−.20 *ns*	−.19 *ns*
Rapid Non-Symbolic Naming						–	.19 *ns*	−.10 *ns*	.00 *ns*	−.18 *ns*
Speech in Noise							–	−.02 *ns*	**−.30** $\mathbf{p = .005}$	**−.33** $\mathbf{p = .002}$
Singing Error								–	.22 *ns*	.04 *ns*
Tapping Error (Synch)									–	**.42** $\mathbf{p < .001}$
Tapping Error (Cont)										–

Note: The measurements used are an average of pre and post. Boldface indicates a significant correlation.

linguistic abilities. More interestingly, singing ability was correlated with two components of the CTOPP: Phonological Awareness and Rapid Symbolic Naming, which may indicate some overlap between singing and speech abilities.

A follow-up correlation analysis was run with the same variables, but partialed out the effects of age. This pattern of significance remained the same as in the previous analysis, with the exception that the correlation between tapping error in the synchronization phase (but not the continuation phase) is no longer significant when age is partialled out, $r(82) = -.23$, $p = .03$.

Discussion

Overall, the data show a clear pattern of improvement in the PPVT, CTOPP, and singing tasks, and no effects in the SIN or tapping tasks. Comparing pre to post measurements, we found improvements in measurements of receptive vocabulary, rapid symbolic and non-symbolic naming abilities, and singing accuracy. Each of these, with the exception of singing ability, are normalized measurements taken from age-standardized tests, making it unlikely that they reflect improvement solely through normal aging. Furthermore, the lack of correlation with age in these tasks (with the exception of the Rapid Non-Symbolic Naming task) underlines the lack of role that age seems to play here. Rather, this improvement seems to be associated with the intervening seven months of music classes. Although the test design does not prove a causal relationship between music classes and this improvement, it does provide evidence that is consistent with this hypothesis. It is also worth noting that the families participating in this study are generally of higher-than-average socio-economic status. While this may make the design slightly more conservative in nature due to the fact that higher socio-economic status is likely to lead to higher language abilities, and thus less room for improvement between sessions, there is also the greater likelihood that these children may have participated in other extracurricular activities, which may also promote development of these abilities.

The singing data provides another clue that there may be a link between music classes and improvements in language ability. While it is not an age-standardized test like the CTOPP or PPVT, the results from the singing test showed that normal development or aging is not primarily responsible for the decrease in singing error; in fact, for our sample, there was no difference in singing ability associated with age. Unlike the language tasks, singing is a skill that is explicitly practiced in all of the Smart Start classes; children both listen to singing and produce it in group and solo contexts in most classes. Thus it is not entirely surprising that children taking music classes would have improved in this ability. As the test song was not explicitly practiced in the music classes, this likely reflects a general improvement in pitch accuracy in singing, rather than learning of a particular song. However, in most studies of music and language in young children, explicit musical ability, and in particular increases of musical ability, are not measured. Our results showed strong improvements in singing ability across the board; the 33% reduction in singing errors puts the children well above the singing abilities of many adults (Hutchins & Peretz, 2012; Pfordresher & Brown, 2007). This indicates that the music classes succeeded in their primary goal of improving musical abilities, and may be taken as corroborative evidence that training, especially at a young age, can improve singing abilities in the general population.

Furthermore, the overall singing error was correlated with phonological awareness scores and rapid symbolic naming ability, indicating a possible source of the transfer between music and language abilities. The link between singing and language abilities has been seen in adult musicians (see, for example, Christiner & Reiterer, 2013, 2015; Ludke, Ferreira, & Overy, 2014); by measuring musical abilities in children, this study shows evidence for a developmental link between the two in early childhood. The close overlap between the mechanisms used to produce singing and speech may facilitate the transfer between the two abilities. In particular, singing requires the production of phonetic information with precise timing, and the improvement in vocal-motor control required to support singing (see Hutchins & Moreno, 2013; Hutchins & Peretz, 2012, for an extended discussion) may carry over to linguistic tasks as well (as well as vice-versa). The speeded production required by in the rapid symbolic naming task may be particularly associated with this motor-control aspect of singing. Based on these data, further studies of the link between these two specific abilities are warranted, and would help to clarify the overall relationship between speech and language.

While it is possible that the use of the alphabet song as the test song could be posited as a reason for the correlation we found, we consider this to be unlikely for three major reasons. First, although rapid symbolic naming involves naming written letters, it is equally composed of naming written numbers, which are not part of the song. Second, the part of the alphabet song which was analyzed also included many excerpts from the final phrase ("Now I know my ABCs . . . "), which is not a recital of letters. Finally, and most importantly, error in the singing task had nothing to do with letter naming accuracy, but was based on pitch alone. Lyric errors were not coded, and it was not even necessary to know the words to sing the pitches accurately. While most children did know the song relatively well, many produced letters out of order or omitted letters, and some even sang "Twinkle Twinkle Little Star" in its place; none of these were considered to be errors, and no visual cues to the lyrics were presented during the singing test. Thus, we conclude that this result is more reflective of a general transfer between music and phonological skills, rather than an artifact of the chosen song.

Previous research has shown associations between music lessons and phonological abilities, and a meta-analysis of 13 studies has demonstrated a significant, if small, effect of music lessons on phonological awareness (Gordon, Fehd, & McCandliss, 2015). Our results show an improvement in some aspects of phonological ability, but no effect in phonological awareness. Two forces may be at play here. First, the children participating in this study already score well above the standardized average in the CTOPP's Phonological Awareness and Phonological Memory (but not in Rapid Symbolic or Non-Symbolic Naming) scales, even before the music classes, leaving less room for improvement in these categories; this is likely because our participant families are often of a higher socio-economic status than the average. Second, Gordon, Fehd, and McCandliss's (2015) meta-analysis indicated that approximately 40 hours of music instruction were needed to show an effect on phonological awareness scores. The children in this study took a total of 28 hours of lessons. Thus is may be that even this year of study was not a long enough time frame to find noticeable effects on this particular skill. Part of our motivation to continue this as an ongoing longitudinal study is to examine the time course of this potential skill development in depth. With more data, we will be able to pull apart several of these effects.

TAPPING

The other surprising non-effect in this study is the lack of an improvement in the tapping task. Tapping is a fundamental musical activity, and the children in our study practiced tapping, patting, and other beat-keeping activities as part of each class. In addition, singing, the other major musical skill we measured, showed a significant improvement, making it unlikely that the courses were simply ineffective at teaching music. Tapping, and more generally, rhythmic ability, has been closely linked to many of the language-related improvements shown in other studies (Huss et al., 2011; Moritz et al., 2013; Woodruff Carr et al., 2014), and has been shown to be deficient in dyslexics (Huss et al., 2011; Thomson et al., 2006; Thomson & Goswami, 2008; Wolff, 2002). Furthermore, the temporal processing required for coordinated tapping seems to be specifically related to rapid automatized naming tasks such as the CTOPP's rapid symbolic and non-symbolic naming tasks (Tallal, Miller, & Fitch, 1995; Wolf, Bowers, & Biddle, 2000). The lack of tapping improvement could be related to non-effects in other areas as well as the lack of a correlation between tapping measures and our linguistic measures, even where they might otherwise be expected. Rather, what we find is a relationship between singing and these measures, which may be indicative of relatively more focus on this aspect of musicality in the classes. As discussed earlier, singing also requires precise timing of motor control, and this may be a better proxy for this factor in this quite young cohort.

In this case, however, the problem seems to lie with the task used to measure tapping. In this study, children were asked to synchronize with a metronome, and continue at the same pace for several seconds afterwards. The results from this task proved to be unreliable in the sense that there was no correlation between pre and post session measurements in our synchronization measurements; this trend was evident with several different task metrics, making it unlikely that the measurement itself was to blame. The variance from this task may have overwhelmed any possible measurable effects. This variance may have been caused by children's relative unfamiliarity with this type of task, compared to something like singing. Tapping to a metronome may in fact be more difficult than tapping to a song, due to the relative paucity of rhythmic information. In addition, the correlation between age and tapping ability, in addition to development of rhythmic abilities, may also indicate a confusion with the instructions; this was also evidenced by the number of children who failed to continue tapping in the continuation phase, after the metronome had stopped.

Finally, there is evidence that children have a preferred tapping rate. This is measured by the spontaneous tapping rate, without an external pacing stimulus. For children in this age range, the preferred tapping rate tends to be about 400 ms IOI (Drake, Jones, & Baruch, 2000; Provasi & Bobin-Bègue, 2003). Our pacing stimulus was somewhat slower than this, at 500 ms IOI (120 beats per second), which is closer to the preferred tapping rate of children with music training (Drake et al., 2000). If children in this study are strongly influenced by their preferred tapping rate, this may overwhelm the effect of any ability to synchronize.

One way to remedy this problem is to present multiple tempi, both faster and slower than the preferred tempo. This would allow us to measure not only the overall tapping error, but also the effect of the pacing signal. Some children may not be accurate at any tempo, but will nevertheless be influenced by the pacing signal, which may be a precursor to more general tapping accuracy. Given that dyslexics and children with SLI, too, can show interactions between their rhythmic deficits and the pacing tempo used (Corriveau & Goswami, 2009; Wolff, Michel, Ovrut, & Drake, 1990), this may also shed more light on the connection between rhythmic and phonological abilities in childhood. We plan to implement this change in future testing sessions.

Another possible problem may come from the relatively sparse musical information present in the metronomic pacing signal—the children have typically trained beat-keeping behaviors in a more robust musical environment, and may not yet be able to transfer this ability to the task as presented. A comparison between the children's abilities in sparse and rich musical contexts may help to shed light on this possible explanation as well.

Overall, the results from this first year of our longitudinal study are in line with expectations. We show an improvement in some musical skills and some language-related skills over the course of a year of musical study, and a correlation between the two abilities. In the coming years, we will continue to monitor the same cohort of children as they progress through their classes, in order to learn more about the time course of musical and language ability development.

Author Note

I would like to thank the Marilyn Thompson Centre, which contributed funding for the development of the course curriculum. I would also like to thank the faculty who taught the courses, and all the participants of the study. Special thanks to Catherine West for overseeing the courses, Angela Elster for her work in course development, and Safia Khalil, Amanda Chan, Ritu Sharma, and the Research Centre staff for coordinating and administering testing sessions, as well as to Dominique Vuvan for reviewing an earlier draft of this paper. I would also like to thank the anonymous reviewers for their helpful comments and suggestions.

Correspondence concerning this article should be addressed to Sean Hutchins, The Royal Conservatory of Music, 273 Bloor St. W, Toronto ON, M5S 1W2, Canada. E-mail: Sean.Hutchins@rcmusic.ca

References

Anderson, S., Skoe, E., Chandrasekaran, B., Zecker, S., & Kraus, N. (2010). Brainstem correlates of speech-in-noise perception in children. *Hearing Research*, *270*(1–2), 151–157. https://doi.org/10.1016/j.heares.2010.08.001

Anvari, S. H., Trainor, L. J., Woodside, J., & Levy, B. A. (2002). Relations among musical skills, phonological processing, and early reading ability in preschool children. *Journal of Experimental Child Psychology*, *83*(2), 111–130. https://doi.org/10.1016/S0022-0965(02)00124-8

Baharloo, S., Johnston, P. A., Service, S. K., Gitschier, J., & Freimer, N. B. (1998). Absolute pitch: An approach for identification of genetic and nongenetic components. *The American Journal of Human Genetics*, *62*(2), 224–231. https://doi.org/10.1086/301704

Bailey, J. A., & Penhune, V. B. (2010). Rhythm synchronization performance and auditory working memory in early- and late-trained musicians. *Experimental Brain Research*, *204*(1), 91–101. https://doi.org/10.1007/s00221-010-2299-y

Bailey, J. A., Zatorre, R. J., & Penhune, V. B. (2013). Early musical training is linked to gray matter structure in the ventral premotor cortex and auditory–motor rhythm synchronization performance. *Journal of Cognitive Neuroscience*, *26*(4), 755–767. https://doi.org/10.1162/jocn_a_00527

Bermudez, P., Lerch, J. P., Evans, A. C., & Zatorre, R. J. (2009). Neuroanatomical correlates of musicianship as revealed by cortical thickness and voxel-based morphometry. *Cerebral Cortex*, *19*(7), 1583–1596. https://doi.org/10.1093/cercor/bhn196

Bhide, A., Power, A., & Goswami, U. (2013). A rhythmic musical intervention for poor readers: A comparison of efficacy with a letter-based intervention. *Mind, Brain, and Education*, *7*(2), 113–123. https://doi.org/10.1111/mbe.12016

Bialystok, E., & DePape, A.-M. (2009). Musical expertise, bilingualism, and executive functioning. *Journal of Experimental Psychology: Human Perception and Performance*, *35*(2), 565–574. https://doi.org/10.1037/a0012735

Bidelman, G. M., Gandour, J. T., & Krishnan, A. (2011). Musicians and tone-language speakers share enhanced brainstem encoding but not perceptual benefits for musical pitch. *Brain and Cognition*, *77*(1), 1–10. https://doi.org/10.1016/j.bandc.2011.07.006

Bidelman, G. M., Hutka, S., & Moreno, S. (2013). Tone language speakers and musicians share enhanced perceptual and cognitive abilities for musical pitch: Evidence for bidirectionality between the domains of language and music. *PLOS ONE*, *8*(4), e60676. https://doi.org/10.1371/journal.pone.0060676

Bidelman, G. M., & Krishnan, A. (2010). Effects of reverberation on brainstem representation of speech in musicians and non-musicians. *Brain Research*, *1355*, 112–125. https://doi.org/10.1016/j.brainres.2010.07.100

Bidelman, G. M., Krishnan, A., & Gandour, J. T. (2011). Enhanced brainstem encoding predicts musicians' perceptual advantages with pitch. *European Journal of Neuroscience*, *33*(3), 530–538. https://doi.org/10.1111/j.1460-9568.2010.07527.x

Bishop-Liebler, P., Welch, G., Huss, M., Thomson, J. M., & Goswami, U. (2014). Auditory temporal processing skills in musicians with dyslexia. *Dyslexia*, *20*(3), 261–279. https://doi.org/10.1002/dys.1479

Bracken, B. A., & Murray, A. M. (1984). Stability and predictive validity of the PPVT-R over an eleven month interval. *Educational and Psychological Research*, *4*(1), 41–44.

Chan, A. S., Ho, Y.-C., & Cheung, M.-C. (1998). Music training improves verbal memory. *Nature*, *396*(6707), 128–128. https://doi.org/10.1038/24075

Chobert, J., François, C., Velay, J.-L., & Besson, M. (2014). Twelve months of active musical training in 8- to 10-year-old children enhances the preattentive processing of syllabic duration and voice onset time. *Cerebral Cortex*, *24*(4), 956–967. https://doi.org/10.1093/cercor/bhs377

Chobert, J., Marie, C., François, C., Schön, D., & Besson, M. (2011). Enhanced passive and active processing of syllables in musician children. *Journal of Cognitive Neuroscience*, *23*(12), 3874–3887. https://doi.org/10.1162/jocn_a_00088

Christiner, M., & Reiterer, S. M. (2013). Song and speech: Examining the link between singing talent and speech imitation ability. *Frontiers in Psychology*, *4*. https://doi.org/10.3389/fpsyg.2013.00874

Christiner, M., & Reiterer, S. M. (2015). A Mozart is not a Pavarotti: Singers outperform instrumentalists on foreign accent imitation. *Frontiers in Human Neuroscience*, *9*. https://doi.org/10.3389/fnhum.2015.00482

Corrigall, K. A., Schellenberg, E. G., & Misura, N. M. (2013). Music training, cognition, and personality. *Frontiers in Psychology*, *4*. https://doi.org/10.3389/fpsyg.2013.00222

Corriveau, K. H., & Goswami, U. (2009). Rhythmic motor entrainment in children with speech and language impairments: Tapping to the beat. *Cortex*, *45*(1), 119–130. https://doi.org/10.1016/j.cortex.2007.09.008

Degé, F., Kubicek, C., & Schwarzer, G. (2011). Music lessons and intelligence: A relation mediated by executive functions. *Music Perception*, *29*, 195–201. https://doi.org/10.1525/mp.2011.29.2.195

Degé, F., & Schwarzer, G. (2011). The effect of a music program on phonological awareness in preschoolers. *Frontiers in Psychology*, *2*. https://doi.org/10.3389/fpsyg.2011.00124

Drake, C., Jones, M. R., & Baruch, C. (2000). The development of rhythmic attending in auditory sequences: Attunement, referent period, focal attending. *Cognition*, *77*(3), 251–288. https://doi.org/10.1016/S0010-0277(00)00106-2

Dunn, L. M., & Dunn, D. M. (2007). *PPVT-4: Peabody picture vocabulary test*. San Antonio, TX: Pearson Assessments.

Elbert, T., Pantev, C., Wienbruch, C., Rockstroh, B., & Taub, E. (1995). Increased cortical representation of the fingers of the left hand in string players. *Science*, *270*(5234), 305–307. https://doi.org/10.1126/science.270.5234.305

Forgeard, M., Winner, E., Norton, A., & Schlaug, G. (2008). Practicing a musical instrument in childhood is associated with enhanced verbal ability and nonverbal reasoning. *PLoS ONE*, *3*(10). https://doi.org/10.1371/journal.pone.0003566

Foster, N. E. V., & Zatorre, R. J. (2010). Cortical structure predicts success in performing musical transformation judgments. *NeuroImage*, *53*(1), 26–36. https://doi.org/10.1016/j.neuroimage.2010.06.042

Gaab, N., Tallal, P., Kim, H., Lakshminarayanan, K., Archie, J. J., Glover, G. H., & Gabrieli, J. D. E. (2005). Neural correlates of rapid spectrotemporal processing in musicians and nonmusicians. *Annals of the New York Academy of Sciences*, *1060*, 82–88. https://doi.org/10.1196/annals.1360.040

Gaser, C., & Schlaug, G. (2003). Brain structures differ between musicians and non-musicians. *Journal of Neuroscience*, *23*(27), 9240–9245.

Geffner, D., & Goldman, R. (2010). *Auditory skills assessment*. San Antonio, TX: Pearson Education.

George, E. M., & Coch, D. (2011). Music training and working memory: An ERP study. *Neuropsychologia*, *49*(5), 1083–1094. https://doi.org/10.1016/j.neuropsychologia.2011.02.001

Gordon, R. L., Fehd, H. M., & McCandliss, B. D. (2015). Does music training enhance literacy skills? A meta-analysis. *Frontiers in Psychology*, *6*. https://doi.org/10.3389/fpsyg.2015.01777

Gordon, R. L., Shivers, C. M., Wieland, E. A., Kotz, S. A., Yoder, P. J., & Devin McAuley, J. (2015). Musical rhythm discrimination explains individual differences in grammar skills in children. *Developmental Science*, *18*(4), 635–644. https://doi.org/10.1111/desc.12230

Goswami, U. (2011). A temporal sampling framework for developmental dyslexia. *Trends in Cognitive Sciences*, *15*(1), 3–10. https://doi.org/10.1016/j.tics.2010.10.001

Goswami, U., Huss, M., Mead, N., Fosker, T., & Verney, J. P. (2013). Perception of patterns of musical beat distribution in phonological developmental dyslexia: Significant longitudinal relations with word reading and reading comprehension. *Cortex*, *49*(5), 1363–1376. https://doi.org/10.1016/j.cortex.2012.05.005

Habib, M., Lardy, C., Desiles, T., Commeiras, C., Chobert, J., & Besson, M. (2016). Music and dyslexia: A new musical training method to improve reading and related disorders. *Frontiers in Psychology*, *7*. https://doi.org/10.3389/fpsyg.2016.00026

Hanna-Pladdy, B., & MacKay, A. (2011). The relation between instrumental musical activity and cognitive aging. *Neuropsychology*, *25*(3), 378–386. https://doi.org/10.1037/a0021895

Herrera, L., Lorenzo, O., Defior, S., Fernandez-Smith, G., & Costa-Giomi, E. (2011). Effects of phonological and musical training on the reading readiness of native- and foreign-Spanish-speaking children. *Psychology of Music*, *39*(1), 68–81. https://doi.org/10.1177/0305735610361995

Ho, Y.-C., Cheung, M.-C., & Chan, A. S. (2003). Music training improves verbal but not visual memory: Cross-sectional and longitudinal explorations in children. *Neuropsychology*, *17*(3), 439–450. https://doi.org/10.1037/0894-4105.17.3.439

Huss, M., Verney, J. P., Fosker, T., Mead, N., & Goswami, U. (2011). Music, rhythm, rise time perception and developmental dyslexia: Perception of musical meter predicts reading and phonology. *Cortex*, *47*(6), 674–689. https://doi.org/10.1016/j.cortex.2010.07.010

Hutchins, S., & Moreno, S. (2013). The Linked Dual Representation model of vocal perception and production. *Frontiers in Psychology*, *4*. https://doi.org/10.3389/fpsyg.2013.00825

Hutchins, S., & Peretz, I. (2012). A frog in your throat or in your ear? Searching for the causes of poor singing. *Journal of Experimental Psychology: General*, *141*(1), 76–97. https://doi.org/10.1037/a0025064

Hutchins, S., Roquet, C., & Peretz, I. (2012). The vocal generosity effect: How bad can your singing be? *Music Perception*, *30*, 147–159. https://doi.org/10.1525/mp.2012.30.2.147

Kishon-Rabin, L., Amir, O., Vexler, Y., & Zaltz, Y. (2011). Pitch discrimination: Are professional musicians better than non-musicians? *Journal of Basic and Clinical Physiology and Pharmacology*, *12*(2), 125–144. https://doi.org/10.1515/JBCPP.2001.12.2.125

Koelsch, S., Fritz, T., Schulze, K., Alsop, D., & Schlaug, G. (2005). Adults and children processing music: An fMRI study. *NeuroImage*, *25*(4), 1068–1076. https://doi.org/10.1016/j.neuroimage.2004.12.050

Kraus, N., Hornickel, J., Strait, D. L., Slater, J., & Thompson, E. (2014). Engagement in community music classes sparks neuroplasticity and language development in children from disadvantaged backgrounds. *Frontiers in Psychology*, *5*. https://doi.org/10.3389/fpsyg.2014.01403

Kraus, N., Slater, J., Thompson, E. C., Hornickel, J., Strait, D. L., Nicol, T., & White-Schwoch, T. (2014). Auditory learning through active engagement with sound: Biological impact of community music lessons in at-risk children. *Frontiers in Neuroscience*, *8*. https://doi.org/10.3389/fnins.2014.00351

Lamb, S. J., & Gregory, A. H. (1993). The relationship between music and reading in beginning readers. *Educational Psychology*, *13*(1), 19–27. https://doi.org/10.1080/0144341930130103

Lee, K. M., Skoe, E., Kraus, N., & Ashley, R. (2009). Selective subcortical enhancement of musical intervals in musicians. *Journal of Neuroscience*, *29*(18), 5832–5840. https://doi.org/10.1523/JNEUROSCI.6133-08.2009

Loui, P., Kroog, K., Zuk, J., Winner, E., & Schlaug, G. (2011). Relating pitch awareness to phonemic awareness in children: Implications for tone-deafness and dyslexia. *Frontiers in Psychology*, *2*. https://doi.org/10.3389/fpsyg.2011.00111

Ludke, K. M., Ferreira, F., & Overy, K. (2014). Singing can facilitate foreign language learning. *Memory and Cognition*, *42*(1), 41–52. https://doi.org/10.3758/s13421-013-0342-5

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. *Cognitive Psychology*, *41*(1), 49–100. https://doi.org/10.1006/cogp.1999.0734

Moradzadeh, L., Blumenthal, G., & Wiseheart, M. (2015). Musical training, bilingualism, and executive function: A closer look at task switching and dual-task performance. *Cognitive Science*, *39*(5), 992–1020. https://doi.org/10.1111/cogs.12183

Moreno, S., Bialystok, E., Barac, R., Schellenberg, E. G., Cepeda, N. J., & Chau, T. (2011). Short-term music training enhances verbal intelligence and executive function. *Psychological Science*, *22*(11), 1425–1433. https://doi.org/10.1177/0956797611416999

Moritz, C., Yampolsky, S., Papadelis, G., Thomson, J., & Wolf, M. (2013). Links between early rhythm skills, musical training, and phonological awareness. *Reading and Writing*, *26*(5), 739–769. https://doi.org/10.1007/s11145-012-9389-0

Oechslin, M. S., Ville, D. V. D., Lazeyras, F., Hauert, C.-A., & James, C. E. (2013). Degree of musical expertise modulates higher order brain functioning. *Cerebral Cortex*, *23*(9), 2213–2224. https://doi.org/10.1093/cercor/bhs206

Overy, K. (2003). Dyslexia and music. *Annals of the New York Academy of Sciences*, *999*(1), 497–505. https://doi.org/10.1196/annals.1284.060

Pallesen, K. J., Brattico, E., Bailey, C. J., Korvenoja, A., Koivisto, J., Gjedde, A., & Carlson, S. (2010). Cognitive control in auditory working memory is enhanced in musicians. *PLOS ONE*, *5*(6), e11120. https://doi.org/10.1371/journal.pone.0011120

Parbery-Clark, A., Skoe, E., & Kraus, N. (2009). Musical experience limits the degradative effects of background noise on the neural processing of sound. *Journal of Neuroscience*, *29*(45), 14100–14107. https://doi.org/10.1523/JNEUROSCI.3256-09.2009

Parbery-Clark, A., Skoe, E., Lam, C., & Kraus, N. (2009). Musician enhancement for speech-in-noise. *Ear and Hearing*, *30*(6), 653–661. https://doi.org/10.1097/AUD.0b013e3181b412e9

Parbery-Clark, A., Strait, D. L., Anderson, S., Hittner, E., & Kraus, N. (2011). Musical experience and the aging auditory system: Implications for cognitive abilities and hearing speech in noise. *PLOS ONE*, *6*(5), e18082. https://doi.org/10.1371/journal.pone.0018082

Patel, A. D. (2011). Why would musical training benefit the neural encoding of speech? The OPERA hypothesis. *Auditory Cognitive Neuroscience*, *2*, 142. https://doi.org/10.3389/fpsyg.2011.00142

Patel, A. D. (2014). Can nonlinguistic musical training change the way the brain processes speech? The expanded OPERA hypothesis. *Hearing Research*, *308*, 98–108. https://doi.org/10.1016/j.heares.2013.08.011

Patscheke, H., Degé, F., & Schwarzer, G. (2016). The effects of training in music and phonological skills on phonological awareness in 4- to 6-year-old children of immigrant families. *Frontiers in Psychology*, *7*. https://doi.org/10.3389/fpsyg.2016.01647

Pfordresher, P. Q., & Brown, S. (2007). Poor-pitch singing in the absence of "tone deafness." *Music Perception*, *25*, 95–115. https://doi.org/10.1525/mp.2007.25.2.95

Pfordresher, P. Q., & Mantell, J. T. (2014). Singing with yourself: Evidence for an inverse modeling account of poor-pitch singing. *Cognitive Psychology*, *70*, 31–57. https://doi.org/10.1016/j.cogpsych.2013.12.005

Provasi, J., & Bobin-Bègue, A. (2003). Spontaneous motor tempo and rhythmical synchronisation in 2½- and 4-year-old children. *International Journal of Behavioral Development*, *27*(3), 220–231. https://doi.org/10.1080/01650250244000290

Schellenberg, E. G. (2004). Music lessons enhance IQ. *Psychological Science*, *15*(8), 511–514. https://doi.org/10.1111/j.0956-7976.2004.00711.x

Schellenberg, E. G. (2006). Long-term positive associations between music lessons and IQ. *Journal of Educational Psychology*, *98*(2), 457–468.

Schlaug, G., Jäncke, L., Huang, Y., Staiger, J. F., & Steinmetz, H. (1995). Increased corpus callosum size in musicians. *Neuropsychologia*, *33*(8), 1047–1055. https://doi.org/10.1016/0028-3932(95)00045-5

Slater, J., Skoe, E., Strait, D. L., O'Connell, S., Thompson, E., & Kraus, N. (2015). Music training improves speech-in-noise perception: Longitudinal evidence from a community-based music program. *Behavioural Brain Research*, *291*, 244–252. https://doi.org/10.1016/j.bbr.2015.05.026

Slater, J., Strait, D. L., Skoe, E., O'Connell, S., Thompson, E., & Kraus, N. (2014). Longitudinal effects of group music instruction on literacy skills in low-income children. *PLOS ONE*, *9*(11), e113383. https://doi.org/10.1371/journal.pone.0113383

Slevc, L. R., Davey, N. S., Buschkuehl, M., & Jaeggi, S. M. (2016). Tuning the mind: Exploring the connections between musical ability and executive functions. *Cognition*, *152*, 199–211. https://doi.org/10.1016/j.cognition.2016.03.017

Slevc, L. R., & Miyake, A. (2006). Individual differences in second-language proficiency does musical ability matter? *Psychological Science*, *17*(8), 675–681. https://doi.org/10.1111/j.1467-9280.2006.01765.x

Steele, C. J., Bailey, J. A., Zatorre, R. J., & Penhune, V. B. (2013). Early musical training and white-matter plasticity in the corpus callosum: Evidence for a sensitive period. *Journal of Neuroscience*, *33*(3), 1282–1290. https://doi.org/10.1523/JNEUROSCI.3578-12.2013

Strait, D. L., Kraus, N., Parbery-Clark, A., & Ashley, R. (2010). Musical experience shapes top-down auditory mechanisms: Evidence from masking and auditory attention performance. *Hearing Research*, *261*(1–2), 22–29. https://doi.org/10.1016/j.heares.2009.12.021

Swaminathan, S., & Gopinath, J. K. (2013). Music training and second-language English comprehension and vocabulary skills in Indian children. *Psychological Studies*, *58*(2), 164–170. https://doi.org/10.1007/s12646-013-0180-3

Takeuchi, A. H., & Hulse, S. H. (1993). Absolute pitch. *Psychological Bulletin*, *113*(2), 345–361. https://doi.org/10.1037/0033-2909.113.2.345

Tallal, P., & Gaab, N. (2006). Dynamic auditory processing, musical experience and language development. *Trends in Neurosciences*, *29*(7), 382–390. https://doi.org/10.1016/j.tins.2006.06.003

Tallal, P., Miller, S., & Fitch, R. H. (1995). Neurobiological basis of speech: A case for the preeminence of temporal processing. *The Irish Journal of Psychology*, *16*(3), 194–219. https://doi.org/10.1080/03033910.1995.10558057

Thompson, W. F., Schellenberg, E. G., & Husain, G. (2004). Decoding speech prosody: Do music lessons help? *Emotion*, *4*(1), 46–64. https://doi.org/10.1037/1528-3542.4.1.46

Thomson, J. M., Fryer, B., Maltby, J., & Goswami, U. (2006). Auditory and motor rhythm awareness in adults with dyslexia. *Journal of Research in Reading*, *29*(3), 334–348. https://doi.org/10.1111/j.1467-9817.2006.00312.x

Thomson, J. M., & Goswami, U. (2008). Rhythmic processing in children with developmental dyslexia: Auditory and motor rhythms link to reading and spelling. *Journal of Physiology-Paris*, *102*(1–3), 120–129. https://doi.org/10.1016/j.jphysparis.2008.03.007

Tsang, C. D., & Conrad, N. J. (2011). Music training and reading readiness. *Music Perception*, *29*, 157–163. https://doi.org/10.1525/mp.2011.29.2.157

Wagner, R., Torgesen, J., Rashotte, C., & Pearson, N. (2012). *Comprehensive test of phonological processing* (2nd ed.). Austin, TX: Pro-Ed.

Wolf, M., Bowers, P. G., & Biddle, K. (2000). Naming-speed processes, timing, and reading: A conceptual review. *Journal of Learning Disabilities*, *33*(4), 387–407. https://doi.org/10.1177/002221940003300409

Wolff, P. H. (2002). Timing precision and rhythm in developmental dyslexia. *Reading and Writing*, *15*(1–2), 179–206. https://doi.org/10.1023/A:1013880723925

Wolff, P. H., Michel, G. F., Ovrut, M., & Drake, C. (1990). Rate and timing precision of motor coordination in developmental dyslexia. *Developmental Psychology*, *26*(3), 349–359.

Woodruff Carr, K., White-Schwoch, T., Tierney, A. T., Strait, D. L., & Kraus, N. (2014). Beat synchronization predicts neural speech encoding and reading readiness in preschoolers. *Proceedings of the National Academy of Sciences*, *111*(40), 14559–14564. https://doi.org/10.1073/pnas.1406219111

Zendel, B. R., & Alain, C. (2012). Musicians experience less age-related decline in central auditory processing. *Psychology and Aging*, *27*(2), 410–417. https://doi.org/10.1037/a0024816

Zuk, J., Benjamin, C., Kenyon, A., & Gaab, N. (2014). Behavioral and neural correlates of executive functioning in musicians and non-musicians. *PLOS ONE*, *9*(6), e99868. https://doi.org/10.1371/journal.pone.0099868

Zuk, J., Bishop-Liebler, P., Ozernov-Palchik, O., Moore, E., Overy, K., Welch, G., & Gaab, N. (2017). Revisiting the "enigma" of musicians with dyslexia: Auditory sequencing and speech abilities. *Journal of Experimental Psychology: General*, *146*(4), 495–511. https://doi.org/10.1037/xge0000281