



# Music and early language acquisition

Anthony Brandt<sup>1\*</sup>, Molly Gebrian<sup>1</sup> and L. Robert Slevc<sup>2</sup>

<sup>1</sup> Shepherd School of Music, Rice University, Houston, TX, USA

<sup>2</sup> Psychology, Language and Music Cognition Lab, University of Maryland, College Park, MD, USA

## Edited by:

Pascal Belin, University of Glasgow, UK

## Reviewed by:

Sascha Frühholz, University of

Geneva, Switzerland

Daniela Sammler, Max Planck

Institute for Human Cognitive and Brain Sciences, Germany

## \*Correspondence:

Anthony Brandt, Shepherd School of Music, Rice University, P.O. Box 1892, Houston, TX 77251, USA.  
e-mail: [abrandt@rice.edu](mailto:abrandt@rice.edu)

Language is typically viewed as fundamental to human intelligence. Music, while recognized as a human universal, is often treated as an ancillary ability – one dependent on or derivative of language. In contrast, we argue that it is more productive from a developmental perspective to describe spoken language as a special type of music. A review of existing studies presents a compelling case that musical hearing and ability is essential to language acquisition. In addition, we challenge the prevailing view that music cognition matures more slowly than language and is more difficult; instead, we argue that music learning matches the speed and effort of language acquisition. We conclude that music merits a central place in our understanding of human development.

**Keywords:** music, language, language acquisition, childhood development, musical development, music cognition, definition of music, emergent modularity

## INTRODUCTION

Just as infants yearn to walk, they have an accelerated drive for language: by age three or four, a child has essentially become competent in his or her native language. While linguistic abilities will continue to be refined, all of the requisite skills for the processing and performing of speech have been acquired (Kuhl, 2004).

Music is recognized as a universal feature of human cognition: every healthy human is born with the ability to appreciate it. However, music's role in human development is often viewed as ancillary and slower to mature. Wilson argues that “whereas language acquisition in children is fast and largely autonomous, music is acquired more slowly and depends on substantial teaching and practice.” As a result, he surmises that music appears “to be derived from language” (Wilson, 2012, p. 283). At its most extreme, Pinker (1997) has described music as “auditory cheesecake, an exquisite confection” without any biological utility.

In this paper, we present a contrasting view: spoken language is introduced to the child as a vocal performance, and children attend to its musical features first. Without the ability to hear musically, it would be impossible to learn to speak. In addition, we question the view that music is acquired more slowly than language (Wilson, 2012) and demonstrate that language and music are deeply entangled in early life and develop along parallel tracks. Rather than describing music as a “universal language,” we find it more productive from a developmental perspective to describe language as a special type of music in which referential discourse is bootstrapped onto a musical framework.

Several factors have clouded an understanding of the entanglement of music and language, especially in the very young. First, overly restrictive definitions of music often impose adult assumptions onto newborns. Second, music and language are often treated as largely independent systems whose convergence is dependent on factors such as musical training. Third, while language skill is typically measured against the adult population at large, musical skill is often measured against the expertise of professional musicians, leading to mismatched expectations that make music

learning seem more arduous and time-consuming. In this paper, we address these issues and show evidence for a deep early entanglement of music and language and for development along largely similar lines.

## DEFINING MUSIC

By adulthood, we all have well-developed ideas about music informed by our culture and individual taste. However, though we all feel we know what music is, it has proven remarkably hard to define. Cross and Morley (2008) cite two dictionary definitions of music: “the art of combining sounds of voices or instruments so as to achieve beauty of form and expression of emotion” and “the art or science of arranging sounds in notes and rhythms to give a desired pattern or effect.” They go on to state: “For contemporary musicologists and ethnomusicologists, these definitions are seriously unsatisfactory.” After reviewing other definitions, they conclude: “All these notions of music reveal themselves to be ideological constructs rooted in the workings of broader socioeconomic and political forces, which change.” (Cross and Morley, 2008, pp. 6–7).

Operating without a clear, generalized definition of music has made scientific conclusions difficult to evaluate, as results cannot be standardized and conflicting data is harder to resolve. Creating such a definition is therefore our starting point for investigating the connection between music and early language acquisition.

A comprehensive scientific definition of music must take into account the following:

## MUSIC VARIES ACROSS CULTURES

The world's indigenous musical traditions are remarkably diverse and often contradict each other in both overt and subtle ways. The discrimination of consonance and dissonance has been cited as a human universal, with dissonance treated as displeasing (Fritz et al., 2009). However, Markoff (1975, p. 135) points out: “The parallel seconds, so widespread in Bulgarian polyphonic folk-singing, may on first hearing impress the listener as being extremely

dissonant. Bulgarian folksingers, however, consider such interval combinations as representing a beauty which is likened to the ‘sound of ringing bells.’” Playing in tune is something Westerners frequently take for granted: the beating created by out of tune notes is considered unpleasant. However, Javanese gamelan ensembles are deliberately de-tuned by small intervals to create beating; notes in perfect accord would be considered “wan and lifeless” (Tenzer, 1991, p. 33). Western musicians often emphasize purity of tone; noise characteristics are considered clumsy. In contrast, Japanese shakuhachi players highlight the noise qualities of their instrument: the sounds of breath and attack transients are considered deeply expressive (Tokita and Hughes, 2008).

Entrainment to a steady pulse is frequently cited as a universal feature of indigenous ensemble music (Cross, 2012). However, Mongolian khoomi throat singers chant in groups without a steady beat (Tuva, 1990). Japanese gagaku contains unpulsed sections and unmeasured pauses (Kyoto Imperial Court Music Orchestra, 1993; Tokita and Hughes, 2008). In heterophonic music, each voice or instrument embellishes a shared melody in its own way, creating a freely pulsed web. The heterophonic performance of psalms in the Scottish Hebrides is said to represent each singer’s “relation to God on a personal basis” (Knudsen, 1968, p. 10). There are many other examples of heterophonic music in traditions as diverse as Ukrainian, Arabic, and South American folk music (Lieneff, 1958; Aretz, 1967; Racy, 2004).

Cultural diversity is true even of basic musical attributes such as how frequencies are classified. In Western music, frequency is mapped onto space: pitches are “high” and “low” and go “up” and “down.” However, in Bali, pitches are “small” and “large”; to the Sayá people of the Amazon “young” and “old”; and to the Shona people of Zimbabwe “crocodile” (for low frequencies) and “those who follow crocodiles” (for high ones; Zbikowski, 2008; Eitan and Timmers, 2010).

Harwood (1976, p. 528) concludes: “Contemporary ethnomusicological research yields an unequivocal response to the question of whether musical structure is similar across cultures. The answer is. . . that similarities are rare and unsystematic.”

#### **MUSICAL PRACTICE VARIES OVER TIME, EVEN WITHIN THE SAME TRADITION**

If this article had been written 700 years ago, creating music with harmonic progressions would have been considered unorthodox: the practice was confined to a few locations in Western Europe (Cohen, 2001). Today harmony is so commonplace that ancient melodies originally performed alone or with a drone are now often “retrofitted” with chord progressions. As a result, ubiquity or popularity in a particular time or place is not a reliable metric. Because the arts are open-ended, a definition also has to allow for as yet unimagined possibilities.

#### **MUSIC IS OFTEN VERY AMBIGUOUS, EVEN ON AN EMOTIONAL LEVEL**

Music is often described as a “language of emotions” (McGilchrist, 2009; Panksepp, 2009; Juslin and Sloboda, 2010). To many, music’s expressivity – unconstrained by literal meaning – is what makes it a “universal language” (Bernstein, 1976; Cross, 2005).

In order for this to be true, emotional readings should translate broadly across cultures. In Western music, one of the strongest

examples of well established emotional attribution is the contrast between the major and minor modes: the major mode is associated with positive affects such as joy, triumph, and tranquility; the minor mode is associated with negative affects such as grief and anger. However, these emotional associations are culturally determined (see, e.g., Dalla Bella et al., 2001): for instance, the Jewish folk song “Hava Nagila” is in the minor mode, but the lyrics are “Let’s rejoice! Let’s rejoice and be happy!” (Rossi, 1998). Indeed, the major/minor expressive contrast was not even well established in Western Europe until late in the eighteenth century: Orfeo’s plaintive aria “Che faro senza Euridice” (“How will I fare without Euridice”) from Gluck’s opera *Orfeo and Euridice* (1752) is written in the “happy” key of C-Major (Gluck, 1992/1752). Although it remains possible that some emotionally relevant aspects of music (perhaps based on more general acoustic cues to emotion) are, in fact, cultural universals (Juslin and Laukka, 2003; Juslin and Västfjäll, 2008; cf. Bryant and Barrett, 2007), we cannot assume that this is the case.

Listening to music is very subjective; as a result, emotional responses are inconsistent and subject to revision. In 1868, a New England critic wrote about an orchestral performance: “It opened with eight bars of a commonplace theme, very much like Yankee Doodle. . . I regret to say that [what followed] appeared to be made up of the strange, the ludicrous, the abrupt, the ferocious, and the screechy, with the slightest possible admixture, here and there, of an intelligible melody” (Slonimsky, 1965, p. 52). The work? The choral Finale of Beethoven’s *Ninth Symphony*, now celebrated as one of the Western canon’s most emotionally gripping works.

Ascertaining a creator’s intent is speculative. Although Leonard Bernstein felt confident his feelings matched those of Beethoven, he acknowledged “We’ll never know and we can’t phone him up.” (Bernstein, 1976, p. 138). The Soviet composer Dmitri Shostakovich was able to fool the authorities with musical tributes deemed to be sincere that Shostakovich privately declared to be bitterly ironic. Commentators still debate whether certain of his works are patriotic or subversive (Fay, 1980). Music’s ambiguity can actually be an advantage in group interactions, enabling it “to be efficacious for individuals and for groups in contexts where language would be unproductive or impotent, precisely because of the need for language to be interpreted unambiguously” (Cross and Morley, 2008, p. 10). While we often *invest* music with emotion and connect deeply to it for that reason, there is too much inconsistency and uncertainty in both personal and cultural views to describe music as a “language of emotions.”

#### **ANY SOUND CAN BE TREATED MUSICALLY**

We often think of music as being performed by voices and melodic instruments. However, the palette of instrumental sounds extends all the way from the sine wave purity of a Western flute to the white noise of a maraca or cymbal crash. While melody is certainly a central feature of music in cultures throughout the world, it is not a prominent feature of many African and Asian drumming traditions, jazz drum solos, or in the extensive body of Western unpitched percussion works. Aboriginal didgeridoos produce different pitches when played by different players; performances rely on rhythm and timbre rather than melody to create musical interest (Tarnopolsky et al., 2005).

Nor is music limited to conventional instruments: the Vanuatu people of the Banks Islands perform “water drumming” by beating rhythmic patterns on waves (von Hornbostel, 1933). Percussionists’ instrumental battery typically includes brake drums, clay pots, and chimes made of shells, glass, and metal. George Antheil incorporated airplane propellers in his *Ballet mécanique*. The advent of recorded media gave rise to *musique concrète*, in which industrial and natural sounds were used as musical material by such composers as Pierre Schaeffer and Pierre Henry.

In order to satisfy all of the above requirements, we propose the following definition: music is creative play with sound; it arises when sound meets human imagination. The term “music” also implies a value placed on the acoustic parameters of envelope, frequency, and spectrum<sup>1</sup> irrespective of any referential function. Musical content is created by the behavior and patterns of these parameters; it can apply to any activity involved with the production and human perception of sound. Any experience – from the strumming of a harp to the blowing of the wind – that involves the cognition of these basic attributes of sound is potentially *musical*. All of the historical features of music – whether they are steady pulse, recognizable melodies, familiar instruments, or even its treatment as an art-form – are higher order phenomena that are flexible, mutable, and culturally mediated.

Our definition puts no limitations on how sound is organized. There is no acoustic imperative that one sound *has* to lead to another. Having played a particular chord on the piano, there is nothing acoustically inevitable about what follows: I can play another chord, turn on a blender, begin chanting. As a result, how sounds are assembled and linked is extremely flexible. Music often promotes the illusion of flow through such higher order features as steady pulse, repeating patterns, imitation, and stepwise motion. However, flow can be manipulated and broken. In the Finale of his *Ninth Symphony*, Beethoven incorporates flashbacks to earlier movements. In his *Symphonies of Winds*, Igor Stravinsky alternates between contrasting passages as if cross-cutting in a film. The composer Karlheinz Stockhausen wrote a group of pieces in “moment form,” in which individual gestures are treated as independent and can be rearranged by the performer (Kramer, 1978).

Finally, we do not require that music exist within a temporal frame, with a clear beginning and end. In many indigenous cultures, musical behavior is woven into everyday life and not treated as a concert experience (Cross, 2012). In 1966, composer LaMonte Young and his partner Marian Zazeela created a “Dream House” in lower Manhattan. The house contained two oscillators that gradually went in and out of tune. For 4 years, Young and Zazeela lived in the house, inviting guests to make repeat visits. Jem Finer’s

“Longplayer” is even more extreme: it is a 1000 year composition for Tibetan singing bowls. Thus far, Young’s dream house and Finer’s “Longplayer” are idiosyncratic. But with the widespread use of electronic media, it is not hard to imagine a future in which sound installations may exist all over the world.

Defining music as “creative play with sound” is both rigorous and inclusive, embracing the full range of musical expression across time and cultures. Icelandic folk song, whose vocal lines follow contours but not precise pitches, Balinese gamelan music, with its often speeding and slowing of pulse, and the open form pieces of Earle Brown, in which no two performances are alike, would all be recognized as music. Any more restrictive a definition risks being contradicted. McAllester writes: “Any student of man must know that somewhere, someone is doing something that he calls music but nobody else would give it that name. That one exception would be enough to eliminate the possibility of a real universal.” (McAllester, 1971, 379).

It is a central human impulse to develop every one of our biological capacities – often beyond its original function. We move – so we run, jump, and dance. We grasp – so we paint, hammer, and slice. We breathe – into flutes, molten glass, and balloons. Music is the natural outcome of a species that takes every facet of its behavior and explores, amplifies, and extends it: it is an on-going conversation between our biological infrastructure and the plasticity of our imaginations. An elemental definition of music that applies broadly across geography, cultures, and eras is vital because it highlights the dynamism of this creative process. Our abilities to engage in and appreciate “creative play with sound” and to consider sounds irrespective of referential function lie at the heart of early language acquisition.

## THE MUSIC OF SPEECH

Language is commonly defined as a symbolic medium for communication, with a lexicon of meanings and syntax for organizing its propositions<sup>2</sup>. We don’t just speak to be heard, we speak to be *understood* – to make declarations of love, order a meal, and ask for directions. But while speech is symbolic, sound is the bearer of its message.

Depending on how one listens, the same stimuli can be perceived as language or music. When one repeatedly listens to the same looped recording of speech, it can begin to sound like singing (Deutsch et al., 2011; Tierney et al., in press): as attention to meaning is satiated, the melodic features of prosodic inflection come to the fore. Conversely, sine wave speech, which tracks the formant frequencies of a spoken utterance without other acoustic attributes of natural speech, sounds like whistles to naïve listeners. However, when subjects are primed to listen for speech, the clips are clearly intelligible (Remez et al., 1981).

Within many cultures, there are gray areas between music and speech. The Ewe tribe in West Africa use talking drums to communicate between villages (Gleick, 2011) while “speakers” of

<sup>1</sup>A sound’s envelope consists of an attack, sustain, and decay, measured as changes in amplitude over time. The envelope creates the basic unit of musical rhythm: the sound’s starting point and its duration. Frequency denotes the oscillations of a sound wave, measured in cycles per second. A sound’s spectrum describes the shape of a sound wave. The sine wave is the basic shape. Deformations create “overtones.” A “harmonic” spectrum involves a periodic fundamental frequency and overtones or partials in simple arithmetic ratios to the fundamental. In human hearing, harmonic sounds produce a stable sense of pitch. Enharmonic sounds include *any* sound that is *not* periodic and in which the overtones are not in simple arithmetic ratios. Depending on the complexity and lay-out of the partials, enharmonic sounds can produce varying degrees of pitch perception, from the “blurry” to the “noisy.”

<sup>2</sup>The Oxford English Dictionary defines language as: “the method of human communication, either spoken or written, consisting of the use of words in a structured and conventional way” (Language, 2012a). Merriam-Webster defines it as “the words, their pronunciation, and the methods of combining them used and understood by a community” (Language, 2012b).

Silbo Gomero use whistles to converse (Carreiras et al., 2005). In Cambodia, secular singing is typically accompanied by a fixed metrical pulse. Buddhist practice argues against music for spiritual practice, so the religious chants, which are highly melodic, are nevertheless treated as speech, to be performed without a rhythmic accompaniment (Sam, 1998). Poetry, with its attention to such sonic features such as rhyming, assonance, alliteration, and metric design, is widely regarded as hovering between music and speech. Indeed, epic poems are often sung: the Finnish Kalevala is frequently performed as a “singing match” between two voices (Siikala, 2000).

As adults, we process “canonical” speech and music differently: for example, speech and music show opposite patterns of hemispheric dominance, with speech processing relying more on the left hemisphere and music relying more on the right (e.g., Callan et al., 2006; Schön et al., 2010). Nevertheless, the neural regions underlying speech and music perception show significant overlap even in adults, with both types of stimuli recruiting a bilateral frontal-temporal network (Griffiths et al., 1999; Merrill et al., 2012). Furthermore, some differences between regions responsive to speech and song in adults is to be expected: over development, our brains become far more specialized in many domains (e.g., Durston et al., 2006; Scott et al., 2007). Although there is little work comparing neural responses to speech and music in infants, there is evidence that newborns show largely overlapping activation to infant directed speech and to instrumental music (Kotilahti et al., 2010) suggesting that processing differences in adult brains may have emerged gradually over the course of development.

It has been suggested that speech and music may have intrinsic differences in low-level auditory characteristics that require different types of aural processing: for instance, some have proposed that speech includes very rapidly changing temporal features whereas music is made up primarily of pitch features varying over a longer time window (e.g., Zatorre et al., 2002). However, speech and music turn out to be closely related in this regard. Perception of temporal changes on the order of 25–50 ms is crucial for the extraction of segmental and phonemic information from the speech signal (Tallal and Piercy, 1973; Rosen, 1992; Telkemeyer et al., 2009). Perception within this small time window is also crucial for instrument recognition. No musical instrument begins with a stable frequency: there is always an onset of noise, caused by the initial impulse that sets the sound in motion. This burst of noise is crucial for timbre perception (Hall, 1991). As a result, the same temporal acuity is required to process both speech and musical timbre (Shepard, 1980; Hukin and Darwin, 1995; Robinson and Patterson, 1995). This is true whether many instruments are playing or just one: Stepanek and Otcenasek (2005) demonstrate remarkable variety in the sounds of a violin, based on register, articulation, and fingering. Thus, both the perception of musical timbres and phonemes rely on rapid temporal processing.

In addition, languages vary in the extent to which they rely on these rapid phonemic cues: some African dialects incorporate as many as 150 separate phonemes, while others, such as Hawaiian, use fewer than 20 (Maddieson, 1984). Similar to speakers of Silbo Gomero, the Pirahã people of the Amazon can converse without phonemes with a humming language; using “intonation, timing, syllable patterns, and stress” and whistling with “no apparent limits

as to the quantity, complexity, or kind of information transmitted.” Although whistling and humming languages are rare, they are an important reminder that language performance is not confined to timbral control (Everett, 1985, 413–414).

Even in languages with rich phonemic inventories that would presumably rely heavily on timbral processing, only a small set of speech sounds actually require resolution on very rapid timescales (McGettigan and Scott, 2012). Instead, the primitives of speech perception might be on a longer timescale corresponding roughly to syllables (e.g., Morillon et al., 2010). Of course, music also relies on analysis over longer time windows: in an instrument such as a flute or piano, the noisy onset resolves into a sustained pitch – the basis of musical melody. This pitched sustain, which takes longer for the ear to measure, is an important part of speech perception as well. This is most clear in tone languages (the most widely spoken being Mandarin Chinese), where pitch is lexically contrastive. In the African language of Kele, the phrase “*alamhaka boili*” has two very different meanings depending on its pitch inflection: it can either mean “He watched the river-bank” or “He boiled his mother-in-law” (Gleick, 2011, 23).

Even in non-tone languages, pitch is an important feature of speech performance. Accented syllables help to parse streams of speech into individual words (e.g., Cutler and Norris, 1988). Pitch inflection is also a primary feature of prosody, which conveys semantic structure and emotional affect. In English, declarative sentences generally end with a drop in pitch, whereas questions end with a rise. Prosody also influences meaning through variations in emphasis. There’s a famous joke in which an out-of-work actor is told he’s been hired as a sub for a Shakespeare performance. He has one line: “Hark! I hear the cannon roar.” He spends the afternoon rehearsing it: “Hark! I hear the cannon roar.” “Hark! I hear the cannon roar.” “Hark! I hear the *cannon* roar.” Variations in pitch and rhythm create his different line readings. Finally, he dresses in costume, is pushed on-stage and greeted with a loud explosion, to which he exclaims, “What the heck was that?”

Thus, both music and speech require aural resolution at similar time-scales. From a musical perspective, speech is a concert of phonemes and syllables, melodically inflected by prosody.

The congruence between speech and music at an atomic perceptual level became more evident with the advent of recorded media and electronic acoustic analysis. As a result, its effects have had a particularly notable impact on twentieth century music. Luciano Berio’s song cycle *Circles* employs a vast battery of percussion, which he uses to mimic various consonant sounds in the text. In one famous passage, the sibilants in the singer’s text are imitated by maracas and other shaken and rattling percussion instruments. The effect is of the text resonating among the percussion, which sustain and amplify the timbre of the words.

In Alvin Lucier’s *I Am Sitting in a Room*. . . for electronic tape, a recording of the composer reading a prepared text is broadcast through loudspeakers and rerecorded. As this process is looped, the natural resonance frequency of the room is enhanced and many of the recognizable features of the composer’s speech degrade. What is left at the end is the resonance frequency of the room pulsing with the stresses and cadence of Lucier’s speech. The lexical and syntactic features of language are stripped away, leaving behind the rhythmic residue of speech.

Other examples draw on the melodic aspects of prosody. In his album *Artist in Residence*, the jazz composer Jason Moran begins one track with an excerpt from a lecture by the artist Adrian Piper. Moran then repeats the clip, this time shadowing it with a piano solo matched to the rhythm and contour of Piper's delivery. Finally, he replays the piano solo on its own and develops it into an extended improvisation. Thus, Piper's lecture provides the "melody" for the jazz solo (Moran, 2006).

The Berio, Lucier, and Moran works aren't merely settings of text. The music is *drawn out* of the text: there is a direct translation or transformation of acoustic features of speech into a purely musical form.

Other composers and performers have made music out of phonemes treated for their musical value alone. Scat-singing is a type of jazz improvisation in which wordless singing, nonsense syllables, and the occasional fragment of speech animate a vocal line, often to humorous effect. By marrying dexterous phonemic patterns with fluid and rhythmic musical lines, scat-singers highlight the human voice as a virtuosic and colorful musical instrument.

In his seminal work *Aventures*, Gyorgy Ligeti invents a glossary of nonsense syllables that serve the work's intricate musical structure. Three singers perform their imaginary discourse with exaggerated prosody and a full battery of non-linguistic vocal sounds, including breathing, laughing, sighing, burping, and crying. What results is an intense portrait of human vocal communication devoid of referential meaning.

These creative examples make explicit the musicality of speech. Speech is sound. Its acoustic attributes – pitch, rhythm, and timbre – can serve strictly musical purposes.

Just as composers have made music out of speech, so too does every human voice. As adults, we learn to tone down the features of speech that do not contribute to meaning. In contrast, infants rely on a complete battery of musical information to learn speech: timbre, pitch, dynamic stress, and rhythm. There is no evidence that timbral information alone would be enough to acquire language; in fact, speech perception can be relatively successful even in the *absence* of timbral cues (Shannon et al., 1995). As the succeeding section will show, the comprehensive nature of the infant's aural attention is a great asset in acquiring language: the infant's attention to *all* of the musical features of speech provide a richer context for language induction.

## MUSIC AND EARLY LANGUAGE ACQUISITION

In order to function in a community, basic speech has to be mastered by everyone. It needs to be understood even when delivered quickly and it needs to be capable of being performed even in moments of stress. All of these factors contribute to the design of this unique form of vocal performance.

But there is another critical feature of language: it needs to be learned by children. Many linguists and anthropologists emphasize that language as a symbolic system of expression is constrained by children's ability to learn. Deacon writes: "The structure of a language is under intense selection pressure because in its reproduction from generation to generation, it must pass through a narrow bottleneck: children's minds." (Deacon, 1997, 110) Language is a compromise between what adults need to say and children's ability to process and perform what they hear. And,

crucially, what infants hear is, by the broad definition above, a form of music.

### NEWBORNS' SENSITIVITY TO MUSICAL SOUNDS

Newborn infants' extensive abilities in different aspects of speech perception have often been cited as evidence that language is innate (e.g., Vouloumanos and Werker, 2007). However, these abilities are dependent on their discrimination of the *sounds* of language, the most musical aspects of speech. We argue not that language has a privileged status in the newborn brain, but rather that *music* has a privileged status that enables us to acquire not only the musical conventions of our native culture, but also enables us to learn our native language. *Without the ability to hear musically, we would be unable to learn language.* Infants are famously able to discriminate the phonemes of all languages (Eimas et al., 1971; Werker and Tees, 1984; Dehaene-Lambertz and Dehaene, 1994), an ability that is evidence of sensitivity to timbre, as discussed above. Although newborns' ability to discriminate different instrumental timbres has not yet been tested, infants are able to use timbre to segregate sound sequences into separate perceptual streams (McAdams and Bertoncini, 1997). If phonemic contrasts and instrumental timbral contrasts rely on the overlapping perceptual mechanisms in infants, one would expect similarly precocious abilities in instrumental timbre discrimination among newborns.

In addition to timbre, newborns are sensitive to the rhythmic components of language and can distinguish between languages based on their rhythmic characteristics (whether or not the contrast includes their native language; Nazzi et al., 1998). Newborns have a preference for their native language as well (Moon et al., 1993), however this has only been explored using languages from two different rhythmic classes. Because the ability to discriminate between two languages of the *same* rhythmic class (e.g., English and German) does not appear until 4 months of age (Nazzi et al., 1998; Gervain and Mehler, 2010), newborns may show a preference for *any* language belonging to the same rhythmic class as their native language. If so, then newborns may not prefer their native language *per se*, but rather the rhythmic characteristics of that language (cf. Friederici et al., 2007). Indeed, infants' early attention to rhythm (e.g., Ramus and Mehler, 1999; Ramus et al., 1999) suggest that they are absorbing the sonic structure of their native language – its rhythms of stresses, its phonemic character – much in the same way that we listen to music.

Newborns can also discriminate a variety of other linguistic characteristics based on the musical aspects of language. For example, infants can distinguish the characteristic prosody (or melody) of their native language from others (Friederici, 2006). In fact, infants show electrophysiological evidence for discrimination of affective prosody even in the first few days of life (Cheng et al., 2012). Another piece of evidence that melodic abilities are important for language development comes from infant cries: the melodic complexity of crying increases over the first few months of life (Wermke and Mende, 2009), and infants who do not show such increasing melodic complexity also show poorer language performance 2 years later (Wermke et al., 2007). Infants can also discriminate individual words with different patterns of lexical stress (Sansavini et al., 1997), can detect acoustic cues that signal word boundaries (Christophe et al., 1994), can distinguish

function from content words based on differing acoustic characteristics (Shi et al., 1999), and show sensitivity to prosodic boundaries in sentences (Pannekamp et al., 2006). Interestingly, word segmentation is, at first, based largely on rhythmic (stress) information, and only later do infants demonstrate sensitivity to other non-stress-based cues (Jusczyk et al., 1999). These findings suggest that these discrimination abilities may explain how infants solve the bootstrapping problem – i.e., how to connect the sounds to meaning. Put another way, infants use the musical aspects of language (rhythm, timbral contrast, melodic contour) as a scaffolding for the later development of semantic and syntactic aspects of language. Infants are not just listening for affective cues nor are they focused exclusively on meaning: they are listening for how their language is composed.

### REFINEMENT OF SOUND PERCEPTION OVER DEVELOPMENT

Gradually, infants' abilities become more refined and culture-specific. At 6 months of age, infants can still discriminate all the phonemic contrasts of the world's languages (Cheour et al., 1998; Rivera-Gaxiola et al., 2005), although they show evidence of being attuned to the vowel sounds of their native language over other languages (Kuhl et al., 1992). Similarly, infants at this age do not show a perceptual bias for the music of their native culture: while Western adults more readily detect changes in melodies made up of pitches from the Western major/minor scale system than in melodies using Javanese scales, infants detect changes equally well in both scale systems (Lynch et al., 1990). This is also seen in the perception of musical meter: Western music overwhelmingly uses simple meters where the underlying beat pattern (regardless of the specific rhythm) is symmetrical and regular. A march goes along in groups of two, a waltz in groups of three and the two never mix (imagine trying to waltz to a piece of music constantly changing between 1-2-3, 1-2-3, and 1-2, 1-2 – you would trip over your partner's feet). Non-Western cultures more commonly use this type of metrical mixing, or complex meters (e.g., Reinhard et al., 2001; Petrov et al., 2011). While Western adults have a harder time detecting changes in complex meters than in simple meters, 6 month old infants again can detect changes equally well (Hannon and Trehub, 2005a; note, though, that infants do develop a *preference* for the meter of their own culture earlier, even while they can accurately discriminate the meter of other cultures; Soley and Hannon, 2010).

Between 6 and 12 months of age, infants' linguistic and musical perception begins to become more specific to their native culture (Figure 1). This occurs earlier for vowel sounds than for consonants: 4–6 month old infants discriminate between non-native vowel contrasts, but 6–8 month old infants do not (Polka and Werker, 1994). In contrast, 6–8 month old infants still readily discriminate between non-native consonants and it is not until 10–12 months of age that most infants lose sensitivity to non-native contrasts (Werker and Tees, 1984)<sup>3</sup>. In addition to these changes in phonemic perception, by 9 months, infants have become especially

sensitive to the stress pattern of their native language (see Jusczyk, 2000, for a review).

The same progression is seen in the perception of pitch and rhythm. By 12 months of age, Western infants show superior detection of mistuned notes in melodies from Western scales versus Javanese scales, just as adults do (Lynch and Eilers, 1992). This is in direct contrast to the 6-month old infants discussed earlier who showed no bias for their native scale structure. Likewise, 12-month old Western infants no longer easily process rhythm in complex meters in contrast to the 6-month old infants discussed above (Hannon and Trehub, 2005b). This perceptual narrowing, specific to an infant's cultural experience, seems to be a domain-general phenomenon across perceptual modalities and is not specific to either music or language (Pascalis et al., 2002; Scott et al., 2007).

What kind of neurophysiological changes underlie this gradual specialization for the speech and musical sounds of one's native culture? Presumably the brain becomes more specialized over development, reflecting a gradual emergence of adult-like networks (cf. Johnson, 2011). For example, "temporal voice areas" (Belin et al., 2000) in the superior temporal sulcus develop selectivity for voice identity and emotional prosody between 4 and 7 months of age (Grossmann et al., 2010; Blasi et al., 2011), corresponding nicely with the increasing sensitivity to culture-specific aspects of speech and music at that age.

As infants gradually become more sensitive to both the musical and linguistic sounds of their culture (and less sensitive to the characteristic sounds from other cultures), they also begin to lay the foundation for processing meaning and syntax. For instance, English speaking infants at 7.5 months show a preference for stress-initial words, which is the predominant stress pattern in English (Jusczyk et al., 1999). Eight-month old infants have become sensitive to the word order conventions of their native language, largely through their use of word frequency, and prosodic information (Weissenborn et al., 1996; Gervain et al., 2008; Nespor et al., 2008; Hochmann et al., 2010). Again, infants are first attuned to the musical aspects of language (stress patterns, prosody).

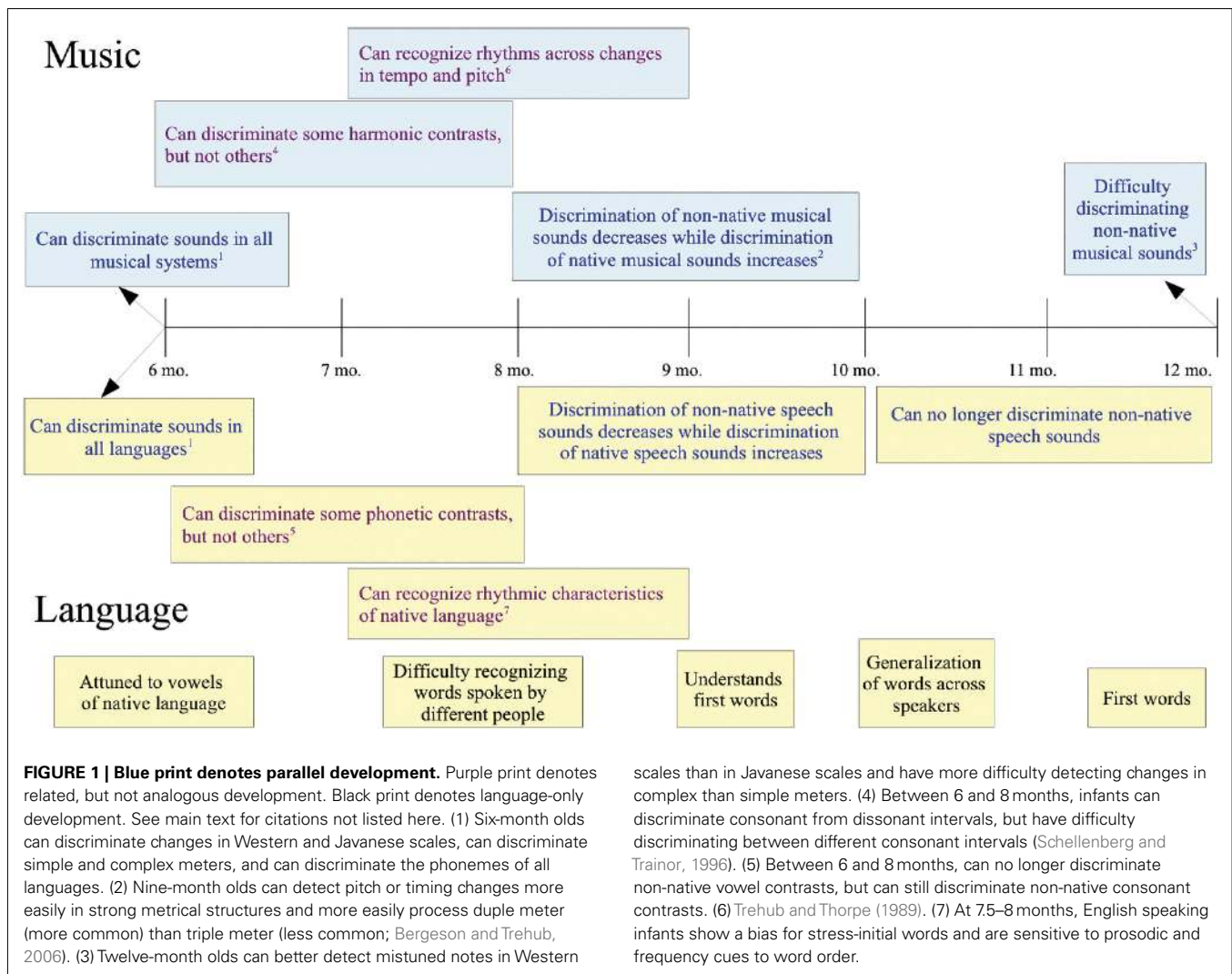
All the aspects of language that an infant can perceive at birth and all of those aspects that are learned during the first year of life are *musical* by the definition of music that we are advocating. The aspects of language that differ the most from music come later: *the further removed a feature of language is from music, the later it is learned*. At around 9 months, infants show evidence of understanding their first words (Friederici, 2006). Once infants discover that words have referential meaning, semantic, and syntactic development takes over. Infants typically begin to talk between 11 and 13 months, experience a vocabulary growth spurt between 18 and 24 months, and reach the high point of their syntactic learning between 18 and 36 months (Friederici, 2006; Kuhl, 2010). From this point on, music and language likely proceed on relatively separate, but parallel, tracks as the musical aspects of language become secondary to its referential and discursive functions.

### SHARED LEARNING MECHANISMS

Infants learn the musical information of speech both by being spoken and sung to directly and by "overhearing" other language and music. Although all speech has musical aspects (see above), speech that is directed to infants is typically characterized by an

<sup>3</sup>Though note that speech sounds that do not exist in the native language, such as Zulu click contrasts for English speakers, remain easily distinguishable by adults (Best et al., 1988).





even greater degree of musicality. This infant directed speech, or *motherese*, is relatively high pitched, slow, and rhythmic, with a larger pitch range and more exaggerated melodic contours than typical adult directed speech (Fernald, 1989). The use of *motherese* appears to be a cultural universal (Fernald, 1992; Falk, 2004) and intentions expressed in infant directed speech can be understood even across very different languages and cultures (Bryant and Barrett, 2007). Infants show a strong preference for infant directed speech (Fernald, 1985; Werker and McLeod, 1989), which seems to reflect the musical aspects of *motherese* as this preference remains (in 4 month old infants) even when the speech samples are filtered to remove lexical content while preserving the prosody (Fernald, 1985; see also Fernald and Kuhl, 1987).

Parents not only speak to their children in musical ways; singing also takes on a specific set of characteristics when directed to children (Trehub and Trainor, 1998). As with *motherese*, characteristics of infant directed song are shared across cultures and can be recognized cross-culturally (Trehub et al., 1993). Even infants of deaf parents (who presumably have heard relatively little

infant directed speech or song) prefer this style of singing to adult directed music (Masataka, 1999).

Why do people speak and sing to infants in this especially musical way? It may be that infant directed speech and singing serves as an aid for language learning by capturing and engaging attention and communicating affective information (Fernald, 1989) and later by enhancing the important patterns in language (such as vowel categories and word divisions; e.g., Kuhl et al., 1997). Others have argued that the role of infant directed speech is not for learning *per se*, but rather serves as a vehicle for emotional communication (Trainor et al., 2000). Of course, infant directed speech might serve all of these purposes, and its role in language acquisition might change over the course of development: first playing an attentional and affective role, and later directing attention to linguistically relevant information (Fernald, 1992). Whatever its purpose, it is clear that a child's first direct exposure to verbal communication frequently has heightened prosody, with the music, and meaning of speech all bound up together.

Human children are not just taught language directly: they learn it through immersion. Even at birth, infants have had considerable opportunities to learn: a fetus starts responding to sound at about the third trimester, and this time until about 6 months of age is a critical period of auditory perceptual development (Birnholtz and Benacerraf, 1983; Graven and Browne, 2008). The sorts of sound that reach prenatal human ears differ from those conducted through air: high-frequency sounds are strongly attenuated, so prenatal stimulation is dominated by low-frequency energy. A fetus can thus detect vowels and musical pitches, but can perceive little of the auditory characteristics that will identify consonants or overtones (Gerhardt and Abrams, 2000).

Note, then, that the features of speech that a fetus is exposed to (and, thus, that an infant is most familiar with) are the more musical features of speech: low-frequency vowel sounds, pitch, and rhythm. Familiarity with these features may explain infants' sensitivity to various aspects of sound at birth: newborns show a preference for their mothers' voice (Mehler et al., 1978), for the sounds of their native language (Mehler et al., 1988; Moon et al., 1993), and can even recognize specific sound stimuli that they hear regularly *in utero*, be it a specific spoken passage (DeCasper and Spence, 1986) or a specific piece of instrumental music (James et al., 2002).

Learning from mere exposure continues after birth as well. Incidental learning of this kind is likely crucial for many aspects of development, including learning the sound structure of the language and music of one's culture. When referring to language acquisition, this process is typically referred to as *statistical learning* (or sometimes *implicit learning*; cf. Perruchet and Pacton, 2006). Statistical learning refers to the (largely implicit) acquisition of structure in the environment, possibly reflecting relatively simple Hebbian learning processes in neural structures. This type of mechanism has seen considerable research in the realm of language development (see Romberg and Saffran, 2010, for a recent review) and has been considered as a mechanism of musical development as well (McMullen and Saffran, 2004; Hannon, 2010).

Statistical learning appears to be a very general phenomenon. It has been demonstrated not only for speech segmentation (Saffran et al., 1996; Mattys and Jusczyk, 2001) and phonetic category learning (Maye et al., 2002), but also for learning of patterns in tone sequences (Saffran et al., 1999), timbre sequences (Tillmann and McAdams, 2004), and even visual and tactile sequences (e.g., Conway and Christiansen, 2005). Interestingly, it seems that musical patterns (as defined here) are perhaps the most amenable to statistical learning: while adult participants can extract statistical regularities from visual and tactile sequences, they do so less well than with auditory sequences (Conway and Christiansen, 2005). Furthermore, a period of auditory deprivation in congenitally deaf children leads to impairments not only in auditory learning, but also in *visual* sequencing abilities (Conway et al., 2011), suggesting that musical learning may be critical not only to learn sound based statistics, but for the learning of sequential and temporal patterns in general (cf. the *auditory scaffolding hypothesis*; Conway et al., 2009). This account is not without its challenges; for example, recent evidence suggests that (adult) congenital amusics show impaired statistical learning of musical tone sequences despite normal statistical learning of speech sounds (Peretz et al.,

2012). However it is likely that statistical learning of musical information – whether the learned patterns eventually are used in the service of speech or music – is a critical part of auditory development.

### LINKED DEVELOPMENTAL DEFICITS

Additional support for the idea that musical hearing is critical to language acquisition and ability comes from studies of children with language disorders and language delays. These children not only show difficulties with the musical aspects of language, but – very tellingly – they show impairments in music processing, too. Although the initial entanglement of music and language gradually unravels over the course of development, the fact that underlying deficits in musical hearing are associated with a variety of language impairments argues for the idea that although music and language grow apart, they are never truly separate in the brain.

Dyslexia, in particular, has been associated with more general auditory processing deficits. It is outside the scope of this paper to discuss these findings and theory at length (see Hämäläinen et al., *in press*, for a recent review), but it is worth noting some of the relevant research. One proposal is that dyslexia results from an underlying problem with rapid temporal processing (Tallal and Piercy, 1973), specifically of the quickly changing formant transitions that distinguish one consonant from another. Treatment programs that use exaggerated versions of these contrasts as well as musical stimuli (e.g., pitch glides) appear to improve reading ability by way of improvements in rapid temporal acuity (Merzenich et al., 1996; Tallal and Gaab, 2006; Gaab et al., 2007), although many studies of these treatments have not been well controlled (McArthur, 2009). Nevertheless, it is clear that dyslexia is associated with rapid temporal processing deficits, and given that phonemic distinctions are akin to the perception and discrimination of instrumental timbres, one would also expect dyslexics to have trouble distinguishing different instruments. Although little attention has been paid to timbre perception in dyslexics (or to timbre perception in general), there is some evidence that dyslexic children do show significantly impaired perception of timbre (Overy, 2000; Overy et al., 2003).

Dyslexic children may also be less sensitive to the amplitude modulations in speech (and other sounds) than normally developing children (Goswami et al., 2002). Indeed, dyslexic children's perception of rise times and perceptual centers (the moment when a sound is perceived to occur) is impaired compared to typically developing children across a variety of language backgrounds (Goswami et al., 2002; Muneaux et al., 2004; Surányi et al., 2009). Interestingly, precocious readers show greater sensitivity to rise time than control subjects, and this sensitivity relates to reading progress (Goswami et al., 2002). Because sensitivity to rise time and perceptual centers is essentially sensitivity to the rhythm of language, dyslexic children are predicted to have difficulties with rhythmic tasks. In fact, dyslexic children do have trouble speaking in time with a metronome, tapping in time with a metronome, rhythm perception (saying whether two rhythms are the same or different), and tempo perception (Overy, 2000; Goswami et al., 2002; Huss et al., 2011).

Additional support for the idea that musical hearing is necessary for reading competency comes from longitudinal studies



of newborns showing that cortical responses to speech and non-speech stimuli at birth are significant predictors of later dyslexia and reading problems (Molfese, 2000; Leppänen et al., 2010) and from a variety of other findings that pitch processing and other abnormal patterns of sound processing predict later reading ability (Leppänen et al., 2010, 2011). There is thus considerable evidence for auditory processing deficits in dyslexia, suggesting that developing competence in reading requires competence in musical hearing. Without an accurate perception of the musical elements of language, learning to read is very difficult, if not impossible.

Another language impairment that may reflect an underlying problem with musical hearing is Specific Language Impairment (SLI). SLI is a failure of normal language development despite normal intelligence and learning environment and an absence of hearing or emotional problems. As in dyslexia, infants that go on to develop SLI have difficulty with rapid temporal processing (Benasich and Tallal, 2002) and discrimination of rise time contrasts (Corriveau et al., 2007). Children later diagnosed with SLI also seem to have a reduced sensitivity to the duration of sounds, which can be seen as young as 2 months of age (Friedrich et al., 2004; Corriveau et al., 2007). The prototypical deficit in SLI is with syntactic processing, which extends to the processing of musical syntax as well (Jentschke et al., 2008). Finally, SLI is associated with impaired statistical learning of both speech stimuli and of non-linguistic tone sequences (Evans et al., 2009). These data suggest that many language learning deficits might be better understood as deficits in the processing complex auditory input (i.e., music). This broader definition may not only be more accurate, but may also help researchers and clinicians develop and advocate for more varied types of intervention.

#### PARALLELS IN MUSIC AND LANGUAGE DEVELOPMENT BEYOND THE FIRST YEAR

The relationship between music and language continues past the first year of life (Figure 2). However, one challenge with comparing language and music development in later childhood is that, while speech ability is measured against the general population, musical ability is often implicitly measured against the virtuosity and expertise of professional musicians. This has contributed to the perception that, whereas language is an innate skill, music is a “gift” and much slower to mature. Becoming a pianist or violinist does depend on a great deal of teaching and practice, but this is the acquisition of a very specialized physical skill. Meanwhile, acquiring the musical conventions of your culture is no more demanding than mastering your native language (Bigand and Poulin-Charronnat, 2006).

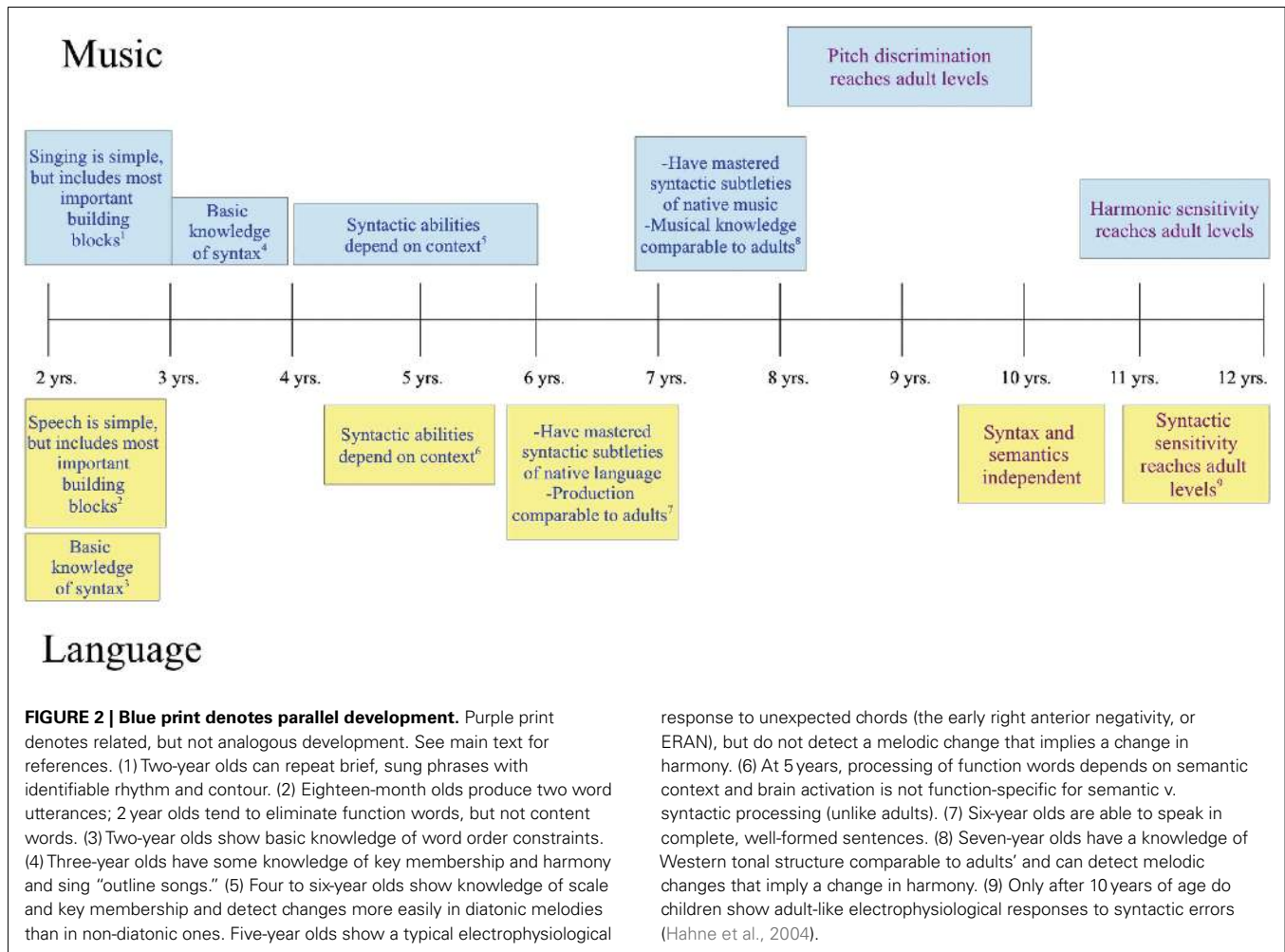
When considering general musical ability (rather than formal musical training), it seems that musical and linguistic development continues on parallel tracks after the first year of life. Between 2–3 years of age, toddlers gain competence with the syntax of their native language (e.g., Höhle et al., 2001) and with the syntax of their culture’s music (in Western music, knowledge of key membership and harmony; Corrigan and Trainor, 2009). This is not complete syntactic competence, however: at age 5, semantics and syntax are still interdependent for children (Friederici, 1983; Brauer and Friederici, 2007), although by age 6, children appear to have mastered the basic syntax of their native language

(Scott, 2004; Nuñez et al., 2011). Similarly, knowledge of musical key membership seems to have developed by about 5 years of age (Trehub et al., 1986; Trainor and Corrigan, 2010). However, while 5 year old children show adult-like electrophysiological responses to incorrect chords (Koelsch et al., 2003), they fail to detect a change in a melody that implies a different harmony (Trainor and Trehub, 1994). This requires a more nuanced understanding of harmony that does not fully develop until age 7 (Trainor and Trehub, 1994) when their knowledge of their native tonal structure is comparable to an adult’s (Speer and Meeks, 1985; McMullen and Saffran, 2004). Similarly, syntactic learning of more complex linguistic constructions continues through age 10 (Friederici, 1983). Children’s pitch discrimination abilities also reach adult levels between 8–10 years of age (Werner and Marean, 1996), and by age 12, their sensitivity to implied harmonies reaches adult levels (Costa-Giomi, 2003).

The similarities of these timelines are remarkable, especially given that all of these studies investigated children from Western cultures, which typically prioritize language learning in school curricula while placing less emphasis on music. Indeed, children given music lessons reach musical developmental milestones sooner than children without music lessons (for a review, see Trainor and Corrigan, 2010). Given that musical development keeps pace with linguistic development even in the face of limited musical instruction, it becomes even clearer that music acquisition is neither especially slow nor effortful.

The evidence discussed so far focuses on music perception, but what about the ability to *make* music? Although Western culture draws a separation between musicians and non-musicians, this has not always been the case and is certainly not true for many non-Western societies where singing (and dancing) are as integral to the community and to the culture as speaking. Because of the relative unimportance of music and singing in Western culture, however, little attention has been paid to the development of singing abilities. Nevertheless, the few studies that have been done suggest that the development of singing ability follows the same trajectory across cultures, with interesting parallels to aspects of language development. For instance, 2 year olds can repeat brief phrases that have an identifiable rhythmic and melodic contour (Dowling, 1999), which is about the same time that young children start to produce short linguistic utterances (Friederici, 2006; Gervain and Mehler, 2010). Three-year olds will mix elements of songs from their culture with their own idiosyncratic vocal improvisations, singing “outline” songs that follow the general contour of culture-specific melodies (Moog, 1976; Davidson, 1994; Hargreaves, 1996), which may be akin to the tendency for 2–3 year olds to eliminate function words, but not the content words in speech (Gerken et al., 1990). Singing ability continues to improve until about 11 years of age (e.g., Howard et al., 1994; Welch, 2002), though ability improves faster and to a greater extent in cultures that emphasize singing (Kreutzer, 2001; Welch, 2009).

Trehub and Trainor (1993) have reasoned that, if the evidence were to show that musical development exhibits “a developmental pace characteristic of innate guided learning, this would revive interest in its biological significance” (317). Music acquisition does, remarkably, keep pace with linguistic development, even in Western cultures where it is not on an equal educational footing



with language. The parallels are even closer in cultures that emphasize music (cf. Kreutzer, 2001). If musical development appears to be slower and more effortful than language acquisition, it seems to be largely a product of culture, not biology.

In sum, infants’ learning of sound structure is based strongly on the musical aspects of sound. This is true in infant directed auditory input, where musical features are exaggerated in both speech and song, and in incidental statistical learning, which relies strongly on musical features like rhythm (especially early in development). These types of exposure are not independent – in fact, statistical learning can be enhanced in the context of infant directed speech (Thiessen et al., 2005) – but the development of both music and speech rely heavily on the musical aspects of children’s environments.

Why is music a part of every human culture? Because language is initially transmitted to children through speech, music cognition may play a strong adaptive function, enabling children’s linguistic skills to mature more rapidly. Arguments for innate language ability often appeal to the “poverty of stimulus” problem (Chomsky, 1980): language is too complex for children to learn based on positive evidence alone. Along with social cues such as facial expressions and physical gestures, the musical features of language

may help surmount the “poverty of stimulus” and provide a richer context for language induction. From a developmental perspective, the progression is clear: first we play with sounds; then we play with meanings and syntax. It is our innate musical intelligence that makes us capable of mastering speech. Music as an art-form may develop from this initial entanglement: it may enable us to continue to explore and exploit features of music cognition that language does not prioritize.

### MUSICIANS AND LATER LANGUAGE ACQUISITION

If, as we propose, music cognition plays a strong role in early language acquisition, we would expect that musical training would correlate with improvements in language learning later in life. In fact, musical training and expertise confer many linguistically relevant advantages (for recent reviews, see Kraus and Chandrasekaran, 2010; Besson et al., 2011; Strait and Kraus, 2011; Slevc, 2012). These include advantages in “low-level” sound processing: musicians show a more faithful brainstem representation of pitch (measured using the frequency-following response or FFR) than non-musicians (Bidelman et al., 2011) presumably as a result of feedback pathways from the cortex to the brainstem (see Kraus and Chandrasekaran, 2010, for a review). This is not only true

for musical stimuli, but also for spoken syllables in one's native language (Musacchia et al., 2007) and in a foreign tone language (Wong et al., 2007). Importantly, these low-level enhancements have practical advantages; for example, musicians are better able to perceive speech-in-noise than non-musicians (Parbery-Clark et al., 2009; Bidelman and Krishnan, 2010). This enhancement extends even to older adults, where musical training seems to protect against the typical decline in the ability to perceive speech in noisy environments (Parbery-Clark et al., 2011).

Musical training also leads to advantages in the processing of prosody: musicians show greater sensitivity than non-musicians to emotional prosodic cues (Thompson et al., 2004; Lima and Castro, 2011) and better detection of subtle prosodic variations at the end of utterances in both their native and in a foreign language (Schön et al., 2004; Marques et al., 2007). Musical training is also associated with better discrimination of subtle timing contrasts in both native and foreign speech (Marie et al., 2011; Sadakata and Sekiyama, 2011). These advantages, too, have practical advantages, for example, in the ability to perceive and learn second language sound structures (Slevc and Miyake, 2006; Lee and Hung, 2008; Delogu et al., 2010).

Linguistic benefits of musical training are not confined to adult musicians: children taking music lessons also show linguistic enhancements relative to their non-musician peers. Like adults, they are better at detecting subtle prosodic variations at the end of utterances (Magne et al., 2006). They also show enhanced passive and active syllable processing, especially voice onset time, a critical ability in distinguishing consonants (Chobert et al., 2011), and show advantages in reading development and phonological awareness (e.g., Lamb and Gregory, 1993; Anvari et al., 2002; Forgeard et al., 2008). Music lessons in children can even enhance pre-linguistic communicative development (i.e., communicative gesturing; Gerry et al., 2012).

One might argue that these advantages reflect innate differences instead of being an effect of musical training itself. While there is relatively little longitudinal data thus far, there is evidence that differences in brain anatomy associated with musicians (e.g., the size of the corpus callosum) can already be seen in the brains of children who have taken 30 months of music lessons, despite showing no differences prior to starting music lessons ( Schlaug et al., 2009a). Similarly, longitudinal studies of children assigned to music or painting lessons show that musical training benefits reading abilities (after only 6 months of lessons) and speech segmentation (after 2 years of lessons), as evidenced by both behavioral and electrophysiological measures (Moreno et al., 2009; François et al., in press).

Patel's (2011) OPERA hypothesis proposes that these benefits of musical training result from overlapping language/music networks, the fact that music involves precise auditory processing, emotional engagement, repetition (i.e., practice), and high attentional demands. Evidence that attending to the musical features of speech is an effective language learning strategy may have implications for adult learning and recovery as well. For example, a focus on musical aspects of speech may improve second language acquisition (cf. Slevc and Miyake, 2006) and musically based therapy may effectively treat developmental and acquired language deficits (see, e.g., Schlaug et al., 2009b).

Although these findings cannot establish a direct link between music and language learning in infants, they are an anticipated outcome of our hypothesis. In addition, the extensive volume of work enhances the view that music and language share many similar properties – something we might expect infants to observe, especially before they are attuned to speech's referential meaning.

## CHALLENGES AND CAVEATS

It is, of course, somewhat controversial to claim that speech is processed as a special form of music. Many have claimed that speech and music are separable modular systems (e.g., Peretz and Coltheart, 2003). Such a separation has even been claimed to be innate, given evidence that infants show left hemispheric lateralization for speech perception and right hemispheric lateralization for frequency perception (Dehaene-Lambertz et al., 2002, 2010). However, these hemispheric asymmetries may reflect cortical specialization for more general auditory properties rather than specificity for speech or music *per se* (e.g., a left hemisphere specialization for rapid temporal processing; Zatorre et al., 2002; Hickok and Poeppel, 2007). Although the hemispheric division of labor is likely not straightforward (Telkemeyer et al., 2009; McGettigan and Scott, 2012), the insight remains that hemispheric differences likely reflect processing asymmetries in aspects of auditory processing rather than specializations for speech or music.

Furthermore, other work does not support this early specialization for language, showing either no lateralization for speech stimuli (e.g., Dehaene-Lambertz, 2000; Kotilahti et al., 2010), or even *right* hemispheric lateralization for speech (Perani et al., 2011) that closely parallels activation to music in an earlier study (Perani et al., 2010). These findings suggest that hemispheric specialization emerges over the course of development. In further support for this idea, early damage to the right hemisphere tends to lead to more severe later language problems than early damage to the left hemisphere (Bates et al., 1997).

A second powerful argument for the neural separability of music and language comes from the dissociation of musical and linguistic abilities sometimes seen in brain damaged patients. Musical deficits can occur without linguistic deficits; in particular, amusics have great difficulty with pitch processing yet typically seem to have normal language abilities (e.g., Ayotte et al., 2002; Peretz, 2006). However, recent work suggests that amusics do in fact have difficulty with aspects of prosodic perception (Liu et al., 2010) and with aspects of phonological and phonemic awareness (Jones et al., 2009). It may seem surprising that these impairments go unnoticed among amusics, however this likely results from the multiple cues to meaning in spoken language, which allow for successful processing of conversational speech even with impaired pitch processing.

Linguistic deficits can accompany preserved musical processing as well; for example, the Russian composer Vissarion Shebalin continued to compose after a series of strokes left him with profound language deficits (Luria et al., 1965; see also Basso and Capitani, 1985; Tzortzis et al., 2000). Such cases provide strong evidence for some degree of music/language separability, however, they

may reflect damage to abilities that have become specialized and neurally separated over development (cf. Karmiloff-Smith, 1995). Supporting this claim, all reported cases of preserved musical processing accompanying linguistic deficits involve professional musicians, who one might expect to show a relatively higher degree of specialization (Tzortzis et al., 2000). These cases also reflect a variety of deficits; cases of preserved musical *sound processing* in the presence of linguistic *sound processing* deficits are elusive at best (e.g., Mendez, 2001; Slevc et al., 2011). While there is little data on language deficits without musical deficits in non-musicians, some evidence does suggest that aphasia in non-musicians may also be accompanied by deficits in aspects of pitch and harmonic processing (Frances et al., 1973; Tallal and Piercy, 1973; Patel et al., 2008).

One might also object to the thesis that language acquisition is inherently musical based on a wide range of evidence that language acquisition is inherently a *social* process. For example, the effectiveness of the sorts of infant directed communication discussed above may be in part musical, but clearly one of the main reasons for the capture of attention is that infant directed communication is highly socially and emotionally expressive (Trehub, 2003). Music is well suited for this sort of communication, however, it is not *only* music that can have these effects; sign language speakers also use infant directed *sign language*, and both deaf and hearing infants prefer this “sign motherese” to adult directed sign language (Masataka, 1996).

The fact that deaf children are able to learn sign languages (at least when receiving appropriate input) may seem especially problematic for the view that musical perception underlies language learning. However it may be that the very aspects of music that are advocated here – especially its nature as a flexible and constantly evolving form of expression – make it well suited to adapt even to different modalities. In particular, the rhythmic and expressive nature of gesture and sign babbling (e.g., Petitto and Marentette, 1991; McClave, 1994) might be a sort of visual parallel to the music of speech.

There are many other questions that need to be addressed before the nature of music and language and its entanglement in the brain (infant or adult) can be satisfactorily resolved. As noted earlier, there are almost no studies of infant timbre processing, nor has much work investigated timbre processing in dyslexia, SLI, or other language disorders. Testing timbre discrimination, especially of instrumental attacks using both native and non-native instrumental timbres, would be informative: if it were shown that phonemic processing was innately separate from the processing of musical timbre, it would raise substantial questions about our claims. Likewise, there is only a small (but growing) body of work on the existence of subtle linguistic deficits in amusia and of subtle (or not so subtle) musical deficits in aphasia. Research on musical development between 12–24 months of age is scarce, perhaps simply because infants of that age are difficult to test; closing this gap would contribute to our understanding of the co-development of music and language. Finally, more research into music and linguistic processing in non-Western cultures is needed. Until research addresses language and music from a broad cross-cultural context, any claims must be circumscribed within a specific cultural context.

## CONCLUSION

A child’s first words are eagerly awaited not only as a cognitive milestone, but as a bond with the adult world – one that heralds the full measure of human thought and expression. But that is not how language cognition begins. For the first year of life, babies hear language as an intentional and often repetitive vocal performance. They listen to it not only for its emotional content but also for its rhythmic and phonemic patterns and consistencies. As Newham says: “. . . whereas the verbal infant will later organize such sounds according to the rules of the dictionary, the baby, not yet familiar with such a scheme, arranges them according to an intuitive, creative, and innate sense of pitch, melody, and rhythm in a fashion directly akin to the composition of music.” (Newham, 1995–1996, p. 67). As sounds are mapped onto meaning, language’s referential function increasingly commands the child’s attention because of its social importance. However, during the first year of life, a different type of listening prevails. Music as an art-form may be a way of prolonging this earlier period, when we encountered the world as a concert and sentences were merely sounds<sup>4</sup>.

So while music and language may be cognitively and neurally distinct in adults, we suggest that language is simply a subset of music from a child’s view. By this account, music, and language are examples of emergent modularity (see, e.g., Karmiloff-Smith, 1995; Johnson, 2010; Johnson, 2011) arising from a common cognitive root.

Is it appropriate to call this common root “musical”? Throughout the world, normally hearing infants are taught language through speech. Both music and speech involve “creative play with sound” and require an attention to acoustic features. The primary difference between them is that the speech is referential. However, as the cited studies have demonstrated, that is a distinction that infants are not yet capable of making in the first year of life.

Would it be better to call the infant’s listening skills “proto-musical”? If a mother sings a lullaby to her child on the first day of life, no one would expect the child to understand the lyrics; but we might reasonably deduce that the child recognizes the mother’s repetition and soothing prosody – and falls asleep accordingly. If we define music as “creative play with sound,” then the evidence suggests that musical engagement is a great way of describing of what infants are doing. As long as the definition of music is the elemental one we are advocating, this terminology does not impose adult categories on the young. As we develop, our musical intelligence is refined based on cultural norms and individual taste and our cognition of music and language become more modular. However, we never lose our innate capacity to treat *any* sound imaginatively. Other than culturally specific features, it is hard to understand what would separate “proto-music” from “music;” any dividing line would likely not translate across cultures and risks sowing confusion.

Music is often described as a universal language but it is neither: musical universals across eras and cultures have been stubbornly

<sup>4</sup>Note the similarity to claims that the evolution of music and language are closely linked (Brown, 1999; Mithen, 2006; Panksepp, 2009): these authors suggest that language may have developed from more primitive vocalizations – we “sang” before we spoke. It thus may be the case that, with respect to the development of music and language, ontogeny recapitulates phylogeny (cf. Gould, 1977).



difficult to find; you cannot order a soda or use the future tense without vocabulary and syntax. However, it may *feel* like a universal language because, for normally developing humans, it underlies the way that we acquire language: as “creative play with sound,” it directs our attention to and amplifies the features of speech that we were paying attention to *before* we were listening for referential meaning. Human creativity, aural abilities, and a desire to communicate underlie both music and language. Listening to music may give us insights into how language sounds to us before we understand it – and how we experience our world before we have words.

## REFERENCES

- Anvari, S. H., Trainor, L. J., Woodside, J., and Levy, B. A. (2002). Relations among musical skills, phonological processing, and early reading ability in preschool children. *J. Exp. Child Psychol.* 83, 111–130.
- Aretz, I. (1967). The polyphonic chant in South America. *Int. Council Tradition. Music* 19, 49–53.
- Ayotte, J., Peretz, I., and Hyde, K. (2002). Congenital amusia: a group study of adults afflicted with a music-specific disorder. *Brain* 125, 238–251.
- Basso, A., and Capitani, E. (1985). Spared musical abilities in a conductor with global aphasia and ideomotor apraxia. *J. Neurol. Sci.* 48, 407–412.
- Bates, E., Thal, D., Trauner, D., Fenson, J., Aram, D., Eisele, J., and Nass, R. (1997). From first words to grammar in children with focal brain injury. *Dev. Neuropsychol.* 13, 275–343.
- Belin, P., Zatorre, R. J., Lafaille, P., Ahad, P., and Pike, B. (2000). Voice-selective areas in human auditory cortex. *Nature* 403, 309–312.
- Benasich, A. A., and Tallal, P. (2002). Infant discrimination of rapid auditory cues predicts later language impairment. *Behav. Brain Res.* 136, 31–49.
- Bergeson, T. R., and Trehub, S. E. (2006). Infants’ perception of rhythmic patterns. *Music Percept.* 23, 345–360.
- Bernstein, L. (1976). *The Unanswered Question: Six Talks at Harvard*. Cambridge, MA: Harvard University Press.
- Besson, M., Chobert, J., and Marie, C. (2011). Transfer of training between music and speech: common processing, attention, and memory. *Front. Psychol.* 2:94. doi:10.3389/fpsyg.2011.00094
- Best, C. T., McRoberts, G. W., and Sitohle, N. M. (1988). Examination of the perceptual re-organization for speech contrasts: Zulu click discrimination by English-speaking adults and infants. *J. Exp. Psychol. Hum. Percept. Perform.* 14, 345–360.
- Bidelman, G. M., and Krishnan, A. (2010). Effects of reverberation on brainstem representation of speech in musicians and non-musicians. *Brain Res.* 1355, 112–125.
- Bidelman, G. M., Krishnan, A., and Gandour, J. T. (2011). Enhanced brainstem encoding predicts musicians’ perceptual advantages with pitch. *Eur. J. Neurosci.* 33, 530–538.
- Bigand, E., and Poulin-Charronnat, B. (2006). Are we “experienced listeners”? A review of the musical capacities that do not depend on formal musical training. *Cognition* 100, 100–130.
- Birnholtz, J., and Benacerraf, B. (1983). The development of human fetal hearing. *Science* 222, 516–518.
- Blasi, A., Mercure, E., Lloyd-Fox, S., Thomson, A., Brammer, M., Sauter, D., Deeley, Q., Barker, G. J., Renvall, V., Deoni, S., Gasston, D., Williams, S. C. R., Johnson, M., Simmons, A., and Murphy, D. G. M. (2011). Early specialization for voice and emotion processing in the infant brain. *Curr. Biol.* 21, 1220–1224.
- Brauer, J., and Friederici, A. D. (2007). Functional neural networks of semantic and syntactic processes in the developing brain. *J. Cogn. Neurosci.* 19, 1609–1623.
- Brown, S. (1999). “The ‘musilanguage’ model of music evolution,” in *The Origins of Music*, eds N. L. Wallin, B. Merker, and S. Brown (Cambridge: The MIT Press), 271–301.
- Bryant, G. A., and Barrett, H. C. (2007). Recognizing intentions in infant-directed speech: evidence for universals. *Psychol. Sci.* 18, 746–751.
- Callan, D. E., Tsytarev, V., Hanakawa, T., Callan, A. M., Katsuhara, M., Fukuyama, H., and Turner, R. (2006). Song and speech: brain regions involved with perception and covert production. *Neuroimage* 31, 1327–1342.
- Carreiras, M., Lopez, J., Rivero, F., and Corina, D. (2005). Linguistic perception: neural processing of a whistled language. *Nature* 433, 31–32.
- Cheng, Y., Lee, S.-Y., Chen, H.-Y., Wang, P.-Y., and Decety, J. (2012). Voice and emotion processing in the human neonatal brain. *J. Cogn. Neurosci.* 24, 1411–1419.
- Cheour, M., Ceponiene, R., Lehtokoski, A., Luuk, A., Allik, J., Alho, K., and Näätänen, R. (1998). Development of language-specific phoneme representations in the infant brain. *Nat. Neurosci.* 1, 351–353.
- Chobert, J., Marie, C., François, C., Schön, D., and Besson, M. (2011). Enhanced passive and active processing of syllables in musician children. *J. Cogn. Neurosci.* 23, 3874–3887.
- Chomsky, N. (1980). *Rules and Representations*. New York, NY: Columbia University Press.
- Christophe, A., Dupoux, E., Bertoncini, J., and Mehler, J. (1994). Do infants perceive word boundaries? An empirical study of the bootstrapping of lexical acquisition. *J. Acoust. Soc. Am.* 95, 1570–1580.
- Cohen, D. E. (2001). The imperfect in search of perfection. *Music Theory Spectrum* 23, 139–169.
- Conway, C. M., and Christiansen, M. H. (2005). Modality-constrained statistical learning of tactile, visual, and auditory sequences. *J. Exp. Psychol. Learn. Mem. Cogn.* 31, 24–39.
- Conway, C. M., Pisoni, D. B., Anaya, E. M., Karpicke, J., and Henning, S. C. (2011). Implicit sequence learning in deaf children with cochlear implants. *Dev. Sci.* 14, 69–82.
- Conway, C. M., Pisoni, D. B., and Kronenberger, W. G. (2009). The importance of sound for cognitive sequencing abilities: the auditory scaffolding hypothesis. *Curr. Dir. Psychol. Sci.* 18, 275–279.
- Corrigall, K. A., and Trainor, L. J. (2009). Effects of musical training on key and harmony perception. *Ann. N. Y. Acad. Sci.* 1169, 164–168.
- Corriveau, K., Pasquini, E., and Goswami, U. (2007). Basic auditory processing skills and specific language impairment: a new look at an old hypothesis. *J. Speech Lang. Hear. Res.* 50, 647–666.
- Costa-Giomi, E. (2003). Young children’s harmonic perception. *Ann. N. Y. Acad. Sci.* 999, 477–484.
- Cross, I. (2005). “Music and meaning, ambiguity and evolution,” in *Musical Communication*, eds D. Miell, R. MacDonald, and D. Hargreaves (Oxford: Oxford University Press), 27–43.
- Cross, I. (2012). “Music as a social and cognitive process,” in *Language and Music as Cognitive Systems*, eds P. Rebuschat, M. Rohrmeier, J. A. Hawkins, and I. Cross (Oxford: Oxford University Press), 315–328.
- Cross, I., and Morley, I. (2008). “The evolution of music: theories, definitions and the nature of the evidence,” in *Communicative Musicality*, eds S. N. Malloch and C. Trevarthen (Oxford: Oxford University Press), 61–82.
- Cutler, A., and Norris, D. (1988). The role of strong syllables in segmentation for lexical access. *J. Exp. Psychol. Hum. Percept. Perform.* 14, 113–121.
- Dalla Bella, S., Peretz, I., Rousseau, L., and Gosselin, N. (2001). A developmental study of the affective value of tempo and mode in music. *Cognition* 80, B1–B10.
- Davidson, L. (1994). “Song singing by young and old: a developmental approach to music,” in *Musical Perceptions*, eds R. Aiello and J. Sloboda (New York: Oxford University Press), 99–130.
- Deacon, T. W. (1997). *The Symbolic Species: The Coevolution of Language and the Brain*. New York, NY: W. W. Norton and Company, Inc.
- DeCasper, J., and Spence, M. (1986). Prenatal maternal speech influences newborns’ perception of speech sounds. *Infant Behav. Dev.* 9, 133–150.
- Dehaene-Lambertz, G. (2000). Cerebral specialization for speech and non-speech stimuli in infants. *J. Cogn. Neurosci.* 12, 449–460.
- Dehaene-Lambertz, G., and Dehaene, S. (1994). Speed and cerebral correlates of syllable discrimination in infants. *Nature* 370, 1–4.

## AUTHORS’ NOTE

The authors wish thank the following for their helpful discussions and advice: Rich Baraniuk, Jonathan Berger, C. Sidney Burrus, J. Todd Frazier, Robert Gjerdingen, Erin Hannon, David Huron, Nina Kraus, Stephen McAdams, Anirruddh Patel, Elizabeth Redcay, Jason Reitman, Kay Kaufman Shelemay, JP Slavinsky, Trent Walker, and Lawrence Zbikowski. We would also like to wish to express our gratitude to the following at Rice University for their support of our project: Caroline Levander, Vice Provost for Interdisciplinary Initiatives, Jan Odegard, Executive Director of the Ken Kennedy Institute, and Robert Yekovich, Dean of the Shepherd School of Music.



- Dehaene-Lambertz, G., Dehaene, S., and Hertz-Pannier, L. (2002). Functional neuroimaging of speech perception in infants. *Science* 298, 2013–2015.
- Dehaene-Lambertz, G., Montavont, A., Jobert, A., Alliol, L., Dubois, J., Hertz-Pannier, L., and Dehaene, S. (2010). Language or music, mother or Mozart? Structural and environmental influences on infants' language networks. *Brain Lang.* 114, 53–65.
- Delogu, F., Lampis, G., and Belardinelli, M. O. (2010). From melody to lexical tone: musical ability enhances specific aspects of foreign language perception. *Eur. J. Cogn. Psychol.* 22, 46–61.
- Deutsch, D., Henthorn, T., and Lapidis, R. (2011). Illusory transformation from speech to song. *J. Acoust. Soc. Am.* 129, 2245–2252.
- Dowling, W. J. (1999). "The development of music perception and cognition," in *The Psychology of Music*, 2nd Edn, ed. D. Deutsch (London: Academic Press), 603–625.
- Durston, S., Davidson, M. C., Tottenham, N., Galvan, A., Spicer, J., Fossella, J. A., and Casey, B. J. (2006). A shift from diffuse to focal cortical activity with development. *Dev. Sci.* 9, 1–8.
- Eimas, P., Siqueland, E., Jusczyk, P., and Vigorito, J. (1971). Speech perception in infants. *Science* 171, 202–206.
- Eitan, Z., and Timmers, R. (2010). Beethoven's last piano sonata and those who follow crocodiles: cross-domain mappings of auditory pitch in a musical context. *Cognition* 114, 405–422.
- Evans, J., Saffran, J., and Robe-Torres, K. (2009). Statistical learning in children with specific language impairment. *J. Speech Lang. Hear. Res.* 52, 321–335.
- Everett, D. L. (1985). "Syllable weight, sloppy phonemes, and channels in Pirahã discourse," in *Proceedings of the Eleventh Annual Meeting of the Berkeley Linguistics Society*, Berkeley, 408–416.
- Falk, D. (2004). Prelinguistic evolution in early hominins: whence motherese? *Behav. Brain Sci.* 27, 491–503; discussion 503–583.
- Fay, L. E. (1980). Shostakovich versus Volkov: whose testimony? *Russ. Rev.* 39, 484–493.
- Fernald, A. (1985). Four-month-old infants prefer to listen to motherese. *Infant Behav. Dev.* 8, 181–195.
- Fernald, A. (1989). Intonation and communicative intent in mothers' speech to infants: is the melody the message? *Child Dev.* 60, 1497–1510.
- Fernald, A. (1992). "Human maternal vocalizations to infants as biologically relevant signals: an evolutionary perspective," in *The Adapted Mind: Evolutionary Psychology and the Generation of Culture*, eds J. H. Barkow, L. Cosmides, and J. Tooby (New York, NY: Oxford University Press), 391–428.
- Fernald, A., and Kuhl, P. (1987). Acoustic determinants of infant preference for motherese speech. *Infant Behav. Dev.* 10, 279–293.
- Forgeard, M., Schlaug, G., Norton, A., Rosam, C., Iyengar, U., and Winner, E. (2008). The relation between music and phonological processing in normal-reading children and children with dyslexia. *Music Percept.* 25, 383–390.
- Frances, R., Lhermitte, E., and Verdy, M. (1973). Le deficit musical des aphasiques. *Int. Rev. Appl. Psychol.* 22, 117–135.
- François, C., Chobert, J., Besson, M., and Schön, D. (in press). Music training for the development of speech segmentation. *Cereb. Cortex* 1–6. doi:10.1093/cercor/bhs180
- Friederici, A. D. (1983). Children's sensitivity to function words during sentence comprehension. *Linguistics* 21, 717–739.
- Friederici, A. D. (2006). The neural basis of language development and its impairment. *Neuron* 52, 941–952.
- Friederici, A. D., Friedrich, M., and Christophe, A. (2007). Brain responses in 4-month-old infants are already language specific. *Curr. Biol.* 17, 1208–1211.
- Friedrich, M., Weber, C., and Friederici, A. D. (2004). Electrophysiological evidence for delayed mismatch response in infants at-risk for specific language impairment. *Psychophysiology* 41, 772–782.
- Fritz, T., Jentschke, S., Gosselin, N., Sammler, D., Peretz, I., Turner, R., Friederici, A. D., and Koelsch, S. (2009). Universal recognition of three basic emotions in music. *Curr. Biol.* 19, 573–576.
- Gaab, N., Gabrieli, J. D. E., Deutsch, G. K., Tallal, P., and Temple, E. (2007). Neural correlates of rapid auditory processing are disrupted in children with developmental dyslexia and ameliorated with training: an fMRI study. *Restor. Neurol. Neurosci.* 25, 295–310.
- Gerhardt, K. J., and Abrams, R. M. (2000). Fetal exposures to sound and vibroacoustic stimulation. *J. Perinatol.* 20(8 Pt 2), S21–S30.
- Gerken, L., Landau, B., and Remez, R. (1990). Function morphemes in young children's speech perception and production. *Dev. Psychol.* 26, 204–216.
- Gerry, D., Unrau, A., and Trainor, L. J. (2012). Active music classes in infancy enhance musical, communicative and social development. *Dev. Sci.* 15, 398–407.
- Gervain, J., and Mehler, J. (2010). Speech perception and language acquisition in the first year of life. *Annu. Rev. Psychol.* 61, 191–218.
- Gervain, J., Nespor, M., Mazuka, R., Horie, R., and Mehler, J. (2008). Bootstrapping word order in prelexical infants: a Japanese-Italian cross-linguistic study. *Cogn. Psychol.* 57, 56–74.
- Gleick, J. (2011). *The Information: A History, a Theory, a Flood*. New York, NY: Pantheon.
- Gluck, C. W. R. V. G. (1992/1752). *Orfeo ed Euridice*. Mineola, NY: Dover.
- Goswami, U., Thomson, J., Richardson, U., Stainthorp, R., Hughes, D., Rosen, S., and Scott, S. K. (2002). Amplitude envelope onsets and developmental dyslexia: a new hypothesis. *Proc. Natl. Acad. Sci. U.S.A.* 99, 10911–10916.
- Gould, S. J. (1977). *Ontogeny and Phylogeny*. Cambridge, MA: The Belknap Press.
- Graven, S. N., and Browne, J. V. (2008). Auditory development in the fetus and infant. *Newborn Infant. Nurs. Rev.* 8, 187–193.
- Griffiths, T. D., Johnsrude, I., Dean, J. L., and Green, G. G. (1999). A common neural substrate for the analysis of pitch and duration pattern in segmented sound? *Neuroreport* 10, 3825–3830.
- Grossmann, T., Oberecker, R., Koch, S. P., and Friederici, A. D. (2010). The developmental origins of voice processing in the human brain. *Neuron* 65, 852–858.
- Hahne, A., Eckstein, K., and Friederici, A. D. (2004). Brain signatures of syntactic and semantic processes during children's language development. *J. Cogn. Neurosci.* 16, 1302–1318.
- Hall, D. (1991). *Musical Acoustics*, 2nd Edn. Pacific Grove, CA: Brooks/Cole Publishers.
- Hämäläinen, J. A., Salminen, H. K., and Leppänen, P. H. (in press). Basic auditory processing deficits in dyslexia: systematic review of the behavioral and event-related potential/field evidence. *J. Learn. Disabil.*
- Hannon, E. E. (2010). "Musical enculturation: How young listeners construct musical knowledge through perceptual experience," in *Neocognitivism: The New Science of Cognitive Development*, ed. S. P. Johnson (New York, NY: Oxford University Press), 132–158.
- Hannon, E. E., and Trehub, S. E. (2005a). Metrical categories in infancy and adulthood. *Psychol. Sci.* 16, 48–55.
- Hannon, E. E., and Trehub, S. E. (2005b). Tuning in to musical rhythms: infants learn more readily than adults. *Proc. Natl. Acad. Sci. U.S.A.* 102, 12639–12643.
- Hargreaves, D. J. (1996). "The development of artistic and musical competence," in *Musical Beginnings*, eds I. Deliège and J. Sloboda (Oxford: Oxford University Press), 145–170.
- Harwood, D. (1976). Universals in music: a perspective from cognitive psychology. *Ethnomusicology* 20, 521–533.
- Hickok, G., and Poeppel, D. (2007). The cortical organization of speech processing. *Nat. Rev. Neurosci.* 8, 393–402.
- Hochmann, J. R., Endress, A. D., and Mehler, J. (2010). Word frequency as a cue for identifying function words in infancy. *Cognition* 115, 444–457.
- Höhle, B., Weissenborn, J., Schmitz, M., and Ischebeck, A. (2001). "Discovering word order regularities: the role of prosodic information for early parameter setting," in *Approaches to Bootstrapping: Phonological, Lexical, Syntactic and Neurophysiological Aspects of Early Language Acquisition*, eds J. Weissenborn and B. Höhle (Amsterdam: John Benjamins Publishing Company), 249–265.
- Howard, D. M., Angus, J. A., and Welch, G. F. (1994). Singing pitching accuracy from years 3 to 6 in a primary school. *Proc. Inst. Acoust.* 16, 223–230.
- Hukin, R. W., and Darwin, C. J. (1995). Comparison of the effect of onset asynchrony on auditory grouping in pitch matching and vowel identification. *Percept. Psychophys.* 57, 191–196.
- Huss, M., Verney, J. P., Fosker, T., Mead, N., and Goswami, U. (2011). Music, rhythm, rise time perception and developmental dyslexia: perception of musical meter predicts reading and phonology. *Cortex* 47, 674–689.
- James, D. K., Spencer, C. J., and Stepis, B. W. (2002). Fetal learning: a prospective randomized controlled study. *Ultrasound Obstet. Gynecol.* 20, 431–438.
- Jentschke, S., Koelsch, S., Sallat, S., and Friederici, A. D. (2008). Children with specific language impairment also show impairment of music-syntactic processing. *J. Cogn. Neurosci.* 20, 1940–1951.
- Johnson, M. H. (2011). Interactive specialization: a domain-general

- framework for human functional brain development? *Dev. Cogn. Neurosci.* 1, 7–21.
- Johnson, S. (2010). *Neoconstructivism: The New Science of Cognitive Development*. New York: Oxford University Press.
- Jones, J. L., Lucker, J., Zalewski, C., Brewer, C., and Drayna, D. (2009). Phonological processing in adults with deficits in musical pitch recognition. *J. Commun. Disord.* 42, 226–234.
- Jusczyk, P. (2000). *The Discovery of Spoken Language*. Cambridge, MA: MIT Press.
- Jusczyk, P. W., Hohne, E. A., and Bauman, A. (1999). Infants' sensitivity to allophonic cues to word segmentation. *Percept. Psychophys.* 61, 1465–1476.
- Juslin, P. N., and Laukka, P. (2003). Communication of emotions in vocal expression and music performance: different channels, same code? *Psychol. Bull.* 129, 770–814.
- Juslin, P. N., and Sloboda, J. A. (2010). *Handbook of Music and Emotion: Theory, Research, Applications*. Oxford: Oxford University Press.
- Juslin, P. N., and Västfjäll, D. (2008). Emotional responses to music: the need to consider underlying mechanisms. *Behav. Brain Sci.* 31, 559–575; discussion 575–621.
- Karmiloff-Smith, A. (1995). *Beyond Modularity: A Developmental Perspective on Cognitive Science*. Cambridge, MA: MIT Press.
- Knudsen, T. (1968). Ornamental hymn/psalm singing in Denmark, the Faroe Islands and the Hebrides. *DFS Inf.* 68, 5–22.
- Koelsch, S., Grossmann, T., Gunter, T. C., Hahne, A., Schröger, E., and Friederici, A. D. (2003). Children processing music: electric brain responses reveal musical competence and gender differences. *J. Cogn. Neurosci.* 15, 683–693.
- Kotilahti, K., Nissilä, I., Näsi, T., Lipiäinen, L., Noponen, T., Meriläinen, P., Huottilainen, M., and Fellman, V. (2010). Hemodynamic responses to speech and music in newborn infants. *Hum. Brain Mapp.* 31, 595–603.
- Kramer, J. D. (1978). Moment form in twentieth century music. *Musical Q.* 64, 177–194.
- Kraus, N., and Chandrasekaran, B. (2010). Music training for the development of auditory skills. *Nat. Rev. Neurosci.* 11, 599–605.
- Kreutzer, N. (2001). Acquisition among rural Shona-speaking Zimbabwean children from birth to 7 years. *J. Res. Music Educ.* 49, 198–211.
- Kuhl, P. K. (2004). Early language acquisition: cracking the speech code. *Nat. Rev. Neurosci.* 5, 831–843.
- Kuhl, P. K. (2010). Brain mechanisms in early language acquisition. *Neuron* 67, 713–727.
- Kuhl, P. K., Andruski, J. E., Chistovich, I., Chistovich, L. A., Kozhevnikova, E. V., Ryskina, V. L., Stolyarova, E. I., Sundberg, U., and Lacerda, F. (1997). Cross-language analysis of phonetic units in language addressed to infants. *Science* 277, 684–686.
- Kuhl, P. K., Williams, K., Lacerda, F., Stevens, K., and Lindblom, B. (1992). Linguistic experience alters phonetic perception in infants by 6 months of age. *Science* 255, 606–608.
- Kyoto Imperial Court Music Orchestra. (1993). *Gagaku: The Imperial Court Music of Japan*. Kyoto: Lyricord Discs Inc.
- Lamb, S. J., and Gregory, A. H. (1993). The relationship between music and reading in beginning readers. *Educ. Psychol.* 13, 19–27.
- Language. (2012a). *Oxforddictionaries.com*. Available at: <http://oxforddictionaries.com/definition/english/language> [accessed July 2012].
- Language. (2012b). *Merriam-Webster.com*. Available at: <http://www.merriam-webster.com/dictionary/language> [accessed July 2012].
- Lee, C.-Y., and Hung, T.-H. (2008). Identification of Mandarin tones by English-speaking musicians and nonmusicians. *J. Acoust. Soc. Am.* 124, 3235–3248.
- Leppänen, P. H. T., Hämäläinen, J. A., Guttorm, T. K., Eklund, K. M., Salminen, H., Tanskanen, A., Torppa, M., Puolakanaho, A., Richardson, U., Pennala, R., and Lyytinen, H. (2011). Infant brain responses associated with reading-related skills before school and at schoolage. *Clin. Neurophysiol.* 42, 35–41.
- Leppänen, P. H. T., Paavo, H. T., Hämäläinen, J. A., Salminen, H. K., Eklund, K. M., Guttorm, T. K., Lohvansuu, K., Puolakanaho, A., and Lyytinen, H. (2010). Newborn brain event-related potentials revealing atypical processing of sound frequency and the subsequent association with later literacy skills in children with familial dyslexia. *Cortex* 46, 1362–1376.
- Lieneff, E. (1958). *Folk Songs of the Ukraine*. Godfrey, IL: Monticello College Press.
- Lima, C. F., and Castro, S. L. (2011). Speaking to the trained ear: musical expertise enhances the recognition of emotions in speech prosody. *Emotion* 11, 1021–1031.
- Liu, F., Patel, A. D., Fourcin, A., and Stewart, L. (2010). Intonation processing in congenital amusia: discrimination, identification and imitation. *Brain* 133, 1682–1693.
- Luria, A. R., Tsvetkova, L. S., and Futer, D. S. (1965). Aphasia in a composer. *J. Neurol. Sci.* 2, 288–292.
- Lynch, M. P., and Eilers, R. E. (1992). A study of perceptual development for musical tuning. *Percept. Psychophys.* 52, 599–608.
- Lynch, M. P., Eilers, R. E., Kimbrough Oller, D., and Urbano, R. C. (1990). Innateness, experience, and music perception. *Psychol. Sci.* 1, 272–276.
- Maddieson, I. (1984). *Patterns of Sounds*. Cambridge, MA: Cambridge University Press.
- Magne, C., Schon, D., and Besson, M. (2006). Musician children detect pitch violations in both music and language better than nonmusician children: behavioral and electrophysiological approaches. *J. Cogn. Neurosci.* 18, 199–211.
- Marie, C., Magne, C., and Besson, M. (2011). Musicians and the metric structure of words. *J. Cogn. Neurosci.* 23, 294–305.
- Markoff, I. (1975). Two-part singing from the Razlog district of southwestern Bulgaria. *Yearbook Int. Folk Music Council* 7, 134–144.
- Marques, C., Moreno, S., Castro, S. L., and Besson, M. (2007). Musicians detect pitch violation in a foreign language better than non-musicians: behavioural and electrophysiological evidence. *J. Cogn. Neurosci.* 19, 1453–1463.
- Masataka, N. (1996). Perception of motherese in a signed language by 6-month-old deaf infants. *Dev. Psychol.* 32, 874–879.
- Masataka, N. (1999). Preference for infant-directed singing in 2-day-old hearing infants of deaf parents. *Dev. Psychol.* 35, 1001–1005.
- Mattys, S., and Jusczyk, P. W. (2001). Do infants segment words or recurring contiguous patterns? *J. Exp. Psychol. Hum. Percept. Perform.* 27, 644–655.
- Maye, J., Werker, J. F., and Gerken, L. (2002). Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition* 82, B101–B111.
- McAdams, S., and Bertoncini, J. (1997). Organization and discrimination of repeating sound sequences by newborn infants. *J. Acoust. Soc. Am.* 102, 2945–2953.
- McAllester, D. P. (1971). Some thoughts on “universals” in world music. *Ethnomusicology* 15, 379–380.
- McArthur, G. (2009). Auditory processing disorders: can they be treated? *Curr. Opin. Neurol.* 22, 137–143.
- McClave, E. (1994). Gestural beats: the rhythm hypothesis. *J. Psycholinguist. Res.* 23, 45–66.
- McGettigan, C., and Scott, S. K. (2012). Cortical asymmetries in speech perception: what's wrong, what's right and what's left? *Trends Cogn. Sci. (Regul. Ed.)* 16, 269–276.
- McGilchrist, I. (2009). *The Master and his Emissary: The Divided Brain and the Making of the Western World*. New Haven, CT: Yale University Press.
- McMullen, E., and Saffran, J. R. (2004). Music and language: a developmental comparison. *Music Percept.* 21, 289–311.
- Mehler, J., Bertoncini, J., Barriere, M., and Gerschenfeld, D. J. (1978). Infant recognition of mother's voice. *Perception* 7, 491–497.
- Mehler, J., Jusczyk, P., and Lamsertz, G. (1988). A precursor of language acquisition in young infants. *Cognition* 29, 143–178.
- Mendez, M. F. (2001). Generalized auditory agnosia with spared music recognition in a left-hander. Analysis of a case with a right temporal stroke. *Cortex* 37, 139–150.
- Merrill, J., Sammler, D., Bangert, M., Goldhahn, D., Lohmann, G., Turner, R., and Friederici, A. D. (2012). Perception of words and pitch patterns in song and speech. *Front. Psychol.* 3:76. doi:10.3389/fpsyg.2012.00076
- Merzenich, M., Jenkins, W., Johnston, P., Schreiner, C., Miller, S., and Tallal, P. (1996). Temporal processing deficits of language-learning impaired children ameliorated by training. *Science* 271, 77–81.
- Mithen, S. (2006). *The Singing Neanderthals: The Origins of Music, Language, Mind and Body*. Cambridge, MA: Harvard University Press.
- Molfese, D. L. (2000). Predicting dyslexia at 8 years of age using neonatal brain responses. *Brain Lang.* 72, 238–245.
- Moog, H. (1976). *The Musical Experience of the Pre-School Child*, trans. C. Clarke. London: Schott.
- Moon, C., Cooper, R. P., and Fifer, W. P. (1993). Two-day-olds prefer their native language. *Infant Behav. Dev.* 16, 495–500.
- Moran, J. (2006). *Artist-in-Residence*. New York: Blue Note Records.
- Moreno, S., Marques, C., Santos, A., Santos, M., Castro, S. L., and Besson, M. (2009). Musical training influences linguistic abilities in 8-year-old children: more evidence for brain plasticity. *Cereb. Cortex* 19, 712–723.

- Morillon, B., Lehongre, K., Frackowiak, R. S. S., Ducorps, A., Kleinschmidt, A., Poeppel, D., and Giraud, A.-L. (2010). Neurophysiological origin of human brain asymmetry for speech and language. *Proc. Natl. Acad. Sci. U.S.A.* 107, 18688–18693.
- Muneaux, M., Ziegler, J. C., Truc, C., Thomson, J., and Goswami, U. (2004). Deficits in beat perception and dyslexia: evidence from French. *Neuroreport* 15, 1255–1259.
- Musackha, G., Sams, M., Skoe, E., and Kraus, N. (2007). Musicians have enhanced subcortical auditory and audiovisual processing of speech and music. *Proc. Natl. Acad. Sci. U.S.A.* 104, 15894–15898.
- Nazzi, T., Bertoncini, J., and Mehler, J. (1998). Language discrimination by newborns: toward an understanding of the role of rhythm. *J. Exp. Psychol. Hum. Percept. Perform.* 24, 756–766.
- Nespor, M., Shukla, M., van deVijver, R., Avesani, C., Schraudolph, H., and Donati, C. (2008). Different phrasal prominence realization in VO and OV languages. *Lingue e Linguaggio*. 7, 1–28.
- Newham, P. (1995–1996). Making a song and dance: the musical voice of language. *J. Imag. Lang. Learn. Teaching* 111. Available at: <http://www.coreilimagination.com/Books.html> [accessed May 10, 2012].
- Núñez, S. C., Dapretto, M., Katzir, T., Starr, A., Bramen, J., Kan, E., Bookheimer, S., and Sowell, E. R. (2011). fMRI of syntactic processing in typically developing children: structural correlates in the inferior frontal gyrus. *Dev. Cogn. Neurosci.* 1, 313–323.
- Overy, K. (2000). Dyslexia, temporal processing and music: the potential of music as an early learning aid for dyslexic children. *Psychol. Music* 28, 218–229.
- Overy, K., Nicolson, R. I., Fawcett, A. J., and Clarke, E. F. (2003). Dyslexia and music: measuring musical timing skills. *Dyslexia* 9, 18–36.
- Panksepp, J. (2009). The emotional antecedents to the evolution of music and language. *Music. Sci.* 13, 229–259.
- Pannekamp, A., Weber, C., and Friederici, A. D. (2006). Prosodic processing at the sentence level in infants. *Neuroreport* 17, 675–678.
- Parbery-Clark, A., Skoe, E., and Kraus, N. (2009). Musical experience limits the degradative effects of background noise on the neural processing of sound. *J. Neurosci.* 29, 14100–14107.
- Parbery-Clark, A., Strait, D. L., Anderson, S., Hittner, E., and Kraus, N. (2011). Musical experience and the aging auditory system: implications for cognitive abilities and hearing speech in noise. *PLoS ONE* 6, e18082. doi:10.1371/journal.pone.0018082
- Pascalis, O., de Haan, M., and Nelson, C. A. (2002). Is face processing species-specific during the first year of life? *Science* 296, 1321–1323.
- Patel, A. D. (2011). Why would musical training benefit the neural encoding of speech? The OPERA hypothesis. *Front. Psychol.* 2:142. doi:10.3389/fpsyg.2011.00142
- Patel, A. D., Iversen, J. R., Wassenaar, M., and Hagoort, P. (2008). Musical syntactic processing in agrammatic Broca's aphasia. *Aphasiology* 22, 776–789.
- Perani, D., Saccuman, M. C., Scifo, P., Anwander, A., Spada, D., Baldoli, C., Poloniato, A., Lohmann, G., and Friederici, A. D. (2011). Neural language networks at birth. *Proc. Natl. Acad. Sci. U.S.A.* 108, 16056–16061.
- Perani, D., Saccumann, M. C., Scifo, P., Spada, D., Andreolli, G., Rovelli, R., Baldoli, C., and Koelsch, S. (2010). Functional specializations for music processing in the human newborn brain. *Proc. Natl. Acad. Sci. U.S.A.* 107, 4758–4763.
- Peretz, I. (2006). The nature of music from a biological perspective. *Cognition* 100, 1–32.
- Peretz, I., and Coltheart, M. (2003). Modularity of music processing. *Nat. Neurosci.* 6, 688–691.
- Peretz, I., Saffran, J., Schön, D., and Goswami, N. (2012). Statistical learning of speech, not music, in congenital amusia. *Ann. N. Y. Acad. Sci.* 1252, 361–366.
- Perruchet, P., and Pacton, S. (2006). Implicit learning and statistical learning: one phenomenon, two approaches. *Trends Cogn. Sci. (Regul. Ed.)* 10, 233–238.
- Petitto, L., and Marentette, P. (1991). Babbling in the manual mode: evidence for the ontogeny of language. *Science* 251, 1493–1496.
- Petrov, S., Manolova, M., and Buchanan, D. A. (2011). *Bulgaria. II. Traditional Music*. Grove Music Online. Available at: <http://www.oxfordmusiconline.com.ezproxy.rice.edu/subscriber/article/grove/music/04289> [accessed May 24, 2010].
- Pinker, S. (1997). *How the Mind Works*. New York, NY: W. W. Norton and Company, Inc.
- Polka, L., and Werker, J. F. (1994). Developmental changes in perception of nonnative vowel contrasts. *J. Exp. Psychol. Hum. Percept. Perform.* 20, 421–435.
- Racy, A. J. (2004). *Making Music in the Arab World: The Culture and Artistry of tarab*. Cambridge: Cambridge University Press.
- Ramus, F., and Mehler, J. (1999). Language identification with suprasegmental cues: study based on speech resynthesis. *J. Acoust. Soc. Am.* 105, 512–521.
- Ramus, F., Nespor, M., and Mehler, J. (1999). Correlates of linguistic rhythm in the speech signal. *Cognition* 73, 265–292.
- Reinhard, K., Stokes, M., and Reinhard, U. (2001). *Turkey. II. Folk Music*. Grove Music Online. Available at: <http://www.oxfordmusiconline.com/subscriber/article/grove/music/44912> [accessed May 24, 2012].
- Remez, R., Rubin, P., Pisoni, D., and Carrell, T. (1981). Speech perception without traditional speech cues. *Science* 212, 947–950.
- Rivera-Gaxiola, M., Klarman, L., García-Sierra, A., and Kuhl, P. K. (2005). Neural patterns to speech and vocabulary growth in American infants. *Neuroreport* 16, 495–498.
- Robinson, K., and Patterson, R. D. (1995). The duration required to identify the instrument, the octave, or the pitch chroma of a musical note. *Music Percept.* 13, 1–14.
- Romberg, A., and Saffran, J. R. (2010). Statistical learning and language acquisition. *Wiley Interdiscip. Rev. Cogn. Sci.* 1, 906–914.
- Rosen, S. (1992). Temporal information in speech: acoustic, auditory and linguistic aspects. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 336, 367–373.
- Rossi, B. (1998). *Jewish Folksongs*. Udine: Pizzicato Edizione Musicale.
- Sadakata, M., and Sekiyama, K. (2011). Enhanced perception of various linguistic features by musicians: a cross-linguistic study. *Acta Psychol. (Amst.)* 138, 1–10.
- Saffran, J. R., Aslin, R. N., and Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science* 274, 1926–1928.
- Saffran, J. R., Johnson, E. K., Aslin, R. N., and Newport, E. L. (1999). Statistical learning of tone sequences by human infants and adults. *Cognition* 70, 27–52.
- Sam, S.-A. (1998). “The Khmer people,” in *The Garland Encyclopedia of World Music*, eds T. E. Miller and S. Williams (New York: Garland Publishing), 151–217.
- Sansavini, A., Bertoncini, J., and Giovanelli, G. (1997). Newborns discriminate the rhythm of multisyllabic stressed words. *Dev. Psychol.* 33, 3–11.
- Schellenberg, E. G., and Trainor, L. J. (1996). Sensory consonance and the perceptual similarity of complex-tone harmonic intervals: tests of adult and infant listeners. *J. Acoust. Soc. Am.* 100, 3321–3328.
- Schlaug, G., Forgeard, M., Zhu, L., Norton, A., Norton, A., and Winner, E. (2009a). Training-induced neuroplasticity in young children. *Ann. N. Y. Acad. Sci.* 1169, 205–208.
- Schlaug, G., Marchina, S., and Norton, A. (2009b). Evidence for plasticity in white matter tracts of chronic aphasic patients undergoing intense intonation-based speech therapy. *Ann. N. Y. Acad. Sci.* 1169, 385–394.
- Schön, D., Gordon, R. L., Campagne, A., Magne, C., Astésano, C., Anton, J.-L., and Besson, M. (2010). Similar cerebral networks in language, music and song perception. *Neuroimage* 51, 450–461.
- Schön, D., Magne, C., and Besson, M. (2004). The music of speech: music facilitates pitch processing in language. *Psychophysiology* 41, 341–349.
- Scott, C. (2004). “Syntactic ability in children and adolescents with language and learning disabilities,” in *Language Development Across Childhood and Adolescence*, ed. R. A. Berman (Amsterdam: John Benjamins Publishing Company), 111–134.
- Scott, L. S., Pascalis, O., and Nelson, C. A. (2007). A domain-general theory of the development of perceptual discrimination. *Curr. Dir. Psychol. Sci.* 16, 197–201.
- Shannon, R. V., Zeng, F. G., Kamath, V., Wygonski, J., and Ekelid, M. (1995). Speech recognition with primarily temporal cues. *Science* 270, 303–304.
- Shepard, R. (1980). Multidimensional scaling, tree-fitting, and clustering. *Science* 210, 390–398.
- Shi, R., Werker, J. F., and Morgan, J. L. (1999). Newborn infants' sensitivity to perceptual cues to lexical and grammatical words. *Cognition* 72, B11–B21.
- Siikala, A.-L. (2000). Body, performance and agency in Kalevala rune-singing. *Oral Tradit.* 15, 255–278.
- Slevc, L. R. (2012). Language and music: sound, structure, and meaning. *Wiley Interdiscip. Rev. Cogn. Sci.* 3, 483–492.
- Slevc, L. R., Martin, R. C., Hamilton, A. C., and Joannisse, M. F. (2011). Speech perception, rapid temporal processing, and the left hemisphere: a case study of unilateral

- pure word deafness. *Neuropsychologia* 49, 216–230.
- Slevc, L. R., and Miyake, A. (2006). Individual differences in second language proficiency: does musical ability matter? *Psychol. Sci.* 17, 675–681.
- Slonimsky, N. (1965). *Lexicon of Musical Invective: Critical Assaults on Composers Since Beethoven's Time*. New York, NY: Coleman-Ross Co.
- Soley, G., and Hannon, E. E. (2010). Infants prefer the musical meter of their own culture: a cross-cultural comparison. *Dev. Psychol.* 46, 286–292.
- Speer, J. R., and Meeks, P. U. (1985). School children's perception of pitch in music. *Psychomusicology* 5, 49–56.
- Stepanek, J., and Otcenasek, Z. (2005). "Acoustical correlates of the main features of violin timbre perception," in *Proceedings of the Conference on Interdisciplinary Musicology*, Nara, 1–9.
- Strait, D., and Kraus, N. (2011). Playing music for a smarter ear: cognitive, perceptual and neurobiological evidence. *Music Percept.* 29, 133–146.
- Surányi, Z., Csépe, V., Richardson, U., Thomson, J. M., Honbolygó, F., and Goswami, U. (2009). Sensitivity to rhythmic parameters in dyslexic children: a comparison of Hungarian and English. *Reading Writing* 22, 41–56.
- Tallal, P., and Gaab, N. (2006). Dynamic auditory processing, musical experience and language development. *Trends Neurosci.* 29, 382–390.
- Tallal, P., and Piercy, M. (1973). Defects of non-verbal auditory perception in children with developmental aphasia. *Nature* 241, 468–469.
- Tarnopolsky, A., Fletcher, N., Hollenberg, L., Lange, B., Smith, J., and Wolfe, J. (2005). Acoustics: the vocal tract and the sound of a didgeridoo. *Nature* 436, 39.
- Telkemeyer, S., Rossi, S., Koch, S. P., Nierhaus, T., Steinbrink, J., Poepel, D., Obrig, H., and Wartenburger, I. (2009). Sensitivity of newborn auditory cortex to the temporal structure of sounds. *J. Neurosci.* 29, 14726–14733.
- Tenzer, M. (1991). *Balinese Music*. Seattle, WA: University of Washington Press.
- Thiessen, E. D., Hill, E. A., and Saffran, J. R. (2005). Infant-directed speech facilitates word segmentation. *Infancy* 7, 53–71.
- Thompson, W. F., Schellenberg, E. G., and Husain, G. (2004). Decoding speech prosody: do music lessons help? *Emotion* 4, 46–64.
- Tierney, A., Dick, F., Deutsch, D., and Sereno, M. (in press). Speech versus song: multiple pitch-sensitive areas revealed by a naturally occurring musical illusion. *Cereb. Cortex*.
- Tillmann, B., and McAdams, S. (2004). Implicit learning of musical timbre sequences: statistical regularities confronted with acoustical (dis)similarities. *J. Exp. Psychol. Learn Mem. Cogn.* 30, 1131–1142.
- Tokita, A. M., and Hughes, D. W. (eds). (2008). "Context and change in Japanese music," in *The Ashgate Research Companion to Japanese Music* (Aldershot: Ashgate Publishing Limited), 1–34.
- Trainor, L. J., Austin, C. M., and Desjardins, N. (2000). Is infant-directed speech a result of the vocal expression of emotion? *Psychol. Sci.* 11, 188–195.
- Trainor, L. J., and Corrigan, K. A. (2010). "Music acquisition and effects of musical experience," in *Music Perception*, Vol. 36, eds M. Riess Jones, R. R. Fay, and A. N. Popper (New York, NY: Springer), 89–127.
- Trainor, L. J., and Trehub, S. E. (1994). Key membership and implied harmony in Western tonal music: developmental perspectives. *Percept. Psychophys.* 56, 125–132.
- Trehub, S. E. (2003). The developmental origins of musicality. *Nat. Neurosci.* 6, 669–673.
- Trehub, S. E., Cohen, A., Thorpe, L., and Morrongiello, B. (1986). Development of the perception of musical relations: semitone and diatonic structure. *J. Exp. Psychol. Hum. Percept. Perform.* 12, 295–301.
- Trehub, S. E., and Thorpe, L. A. (1989). Infants' perception of rhythm: categorization of auditory sequences by temporal structure. *Can. J. Psychol.* 43, 217–229.
- Trehub, S. E., and Trainor, L. (1993). "Listening strategies in infancy: the roots of language and musical development," in *Cognitive Aspects of Human Audition*, eds S. McAdams and E. Bigand (London: Oxford University Press), 278–327.
- Trehub, S. E., and Trainor, L. (1998). "Singing to infants: babies and play songs," in *Advances in Infancy Research*, Vol. 12, eds L. P. Lipsitt, C. K. Rovee-Collier, and H. Hayne (Stamford, CT: Albex Publishing Corp), 43–77.
- Trehub, S. E., Unyk, A. M., and Trainor, L. J. (1993). Maternal singing in cross-cultural perspective. *Infant Behav. Dev.* 16, 285–295.
- Tuva. (1990). *Tuva: Voices from the Center of Asia*. Cambridge, MA: Smithsonian Folkways Recordings, distributed by Rounder Records.
- Tzortzis, C., Goldblum, M. C., Dang, M., Forette, E., and Boller, F. (2000). Absence of amusia and preserved naming of musical instruments in an aphasic composer. *Cortex* 36, 227–242.
- von Hornbostel, E. M. (1933). The ethnology of African sound-instruments. *Africa (Lond.)* 6, 277–311.
- Vouloumanos, A., and Werker, J. F. (2007). Listening to language at birth: evidence for a bias for speech in neonates. *Dev. Sci.* 10, 159–164.
- Weissenborn, J., Höhle, B., Kiefer, D., and Cavar, D. (1996). "Children's sensitivity to word-order violations in German: evidence for very early parameter-setting," in *Proceedings of the 22nd Annual Boston University Conference on Language Development*, Boston, MA, 756–777.
- Welch, G. F. (2002). "Early childhood musical development," in *The Arts in Children's Lives: Context, Culture and Curriculum*, eds L. Bresler and C. Thompson, (Dordrecht, NL: Kluwer), 113–128.
- Welch, G. F. (2009). Evidence of the development of vocal pitch matching ability in children. *Jap. J Music Educ. Res.* 21, 1–13.
- Werker, J., and McLeod, P. (1989). Infant preference for both male and female infant-directed talk: a developmental study of attentional and affective responsiveness. *Can. J. Psychol. Rev.* 43, 230–246.
- Werker, J. F., and Tees, R. (1984). Cross-language speech perception: evidence for perceptual reorganization during the first year of life. *Infant Behav. Dev.* 7, 49–63.
- Wermke, K., Leising, D., and Stellzig-Eisenhauer, A. (2007). Relation of melody complexity in infants' cries to language outcome in the second year of life: a longitudinal study. *Clin. Linguist. Phon.* 21, 961–973.
- Wermke, K., and Mende, W. (2009). Musical elements in human infants' cries: in the beginning is the melody. *Music. Sci.* 13, 151–175.
- Werner, L. A., and Marean, G. C. (1996). *Human Auditory Development*. Madison, WI: Brown Benchmark.
- Wilson, E. O. (2012). *The Social Conquest of the Earth*. New York, NY: Liveright Publishers.
- Wong, P. C. M., Skoe, E., Russo, N. M., Dees, T., and Kraus, N. (2007). Musical experience shapes human brainstem encoding of linguistic pitch patterns. *Nat. Neurosci.* 10, 420–422.
- Zatorre, R. J., Belin, P., and Penhune, V. B. (2002). Structure and function of auditory cortex: music and speech. *Trends Cogn. Sci. (Regul. Ed.)* 6, 37–46.
- Zbikowski, L. (2008). "Metaphor and music," in *The Cambridge Handbook of Metaphor and Thought*, ed. R. W. Jr. Gibbs (Cambridge: Cambridge University Press), 502–524.

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 15 June 2012; accepted: 15 August 2012; published online: 11 September 2012.

Citation: Brandt A, Gebrian M and Slevc LR (2012) Music and early language acquisition. *Front. Psychology* 3:327. doi: 10.3389/fpsyg.2012.00327

This article was submitted to *Frontiers in Auditory Cognitive Neuroscience*, a specialty of *Frontiers in Psychology*. Copyright © 2012 Brandt, Gebrian and Slevc. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in other forums, provided the original authors and source are credited and subject to any copyright notices concerning any third-party graphics etc.