19. MUSIC AND LANGUAGE: MILESTONES OF DEVELOPMENT

Anthony Brandt, Rice University
Molly Gebrian, University of Arizona
L. Robert Slevc, University of Maryland

This is a draft of a chapter that has been accepted for publication by Oxford University Press in the forthcoming book *The Oxford Handbook of Language and Music* edited by Daniela Sammler, due for publication in 2023.

ABSTRACT

The human brain continues to mature throughout childhood, making our species particularly susceptible to experience. Given the diversity of music and language around the globe, how these are acquired during childhood is revealing about the feedback loop between our biological predispositions and exposure. Evidence suggests that children begin as generalists and become specialists, with music and language deeply entangled in infancy and modularity emerging over time. In addition, development proceeds along parallel tracks, with comparable cognitive milestones. Although there is a tendency to celebrate our precociousness, it may be that we should really extol the slow and protected aspects of development: our unfledged entry into the world affords us the extended time necessary to internalize these products of culture. The present chapter begins by exploring the variety of music and languages around the world. It then tracks developmental milestones from birth throughout childhood, examines linked developmental disorders, and closes with a discussion of open questions and future directions.

Keywords: music, language, language acquisition, childhood development, musical development, music cognition, emergent modularity, singing, speaking, infant development

19.1 INTRODUCTION

In order for our large heads to make it through the birth canal, our species relies on an evolutionary trade-off: even babies carried full term are effectively born prematurely. For an idea of how we compare to other animals, foals are able to walk within two hours of birth, lion cubs within about two weeks. It takes human infants many months to take their first steps. As a result, “For humans at birth, the brain is remarkably unfinished, and interaction with the world is necessary to complete it” (Eagleman, 2020 p. 20). Our brain’s coming of age extends well into childhood: unlike our animal brethren, we have to keep a watchful eye on rambunctious teenagers, whose prefrontal cortex is still maturing.

As our brains develop, the shaping of their structure and connectivity is dependent on signals both from within and without (Tierney and Nelson, 2009). Yet to what degree? To what extent are we genetically pre-programmed, to what extent adaptable to experience? Are certain abilities biologically privileged? How do people raised in different backgrounds compare?

Here the acquisitions of music and language offer windows into the developing brain. Both means of communication have been found in some form in all known human cultures, and are acquired in the course of typical childhood development. There is significant overlap between them: researchers have found shared links in structural processing in language and music, (Asano, Boeckx, & Seifert, 2021; Patel, 2003; Sammler et al., 2013; Slevc et al., 2009; Koelsch, 2012; Yu et al., 2017), error detection (Jentschke and Koelsch, 2009; Fitzroy and Sanders, 2013), cognitive control (Slevc and Okada, 2015), as well as a host of advantages for musicians in learning languages (Musacchia et al., 2007; Dittinger et al., 2016; Magne et al., 2004; Milovanov et al., 2008) and for speakers of tone languages in sensitivity to music (Bidelman, Hutka & Moreno, 2013). Yet language and music are also clearly distinguished. For starters, language is referential, enabling us to share our observations, recollections, plans, thoughts, and feelings. Meanwhile, music is more ambiguous: indeed, Cross reasons that music's "floating intentionality" offers a complementary tool to language's specificity that is particularly advantageous in situations of social uncertainty" (Cross, 2008). In adult brains, processing is only partially entangled: there are crucial hubs for language in Broca's area and Wernicke's area; meanwhile much musical processing is found in complementary regions (Fedorenko et al., 2011; Norman-Haignere et al., 2015; Ogg et al., 2019; Rogalsky et al., 2011; Norman-Haignere et al, 2022). Music and language are both similar and different enough to raise provocative questions about their biological foundations and susceptibility to experience.

Analyzing the developmental trajectory of music and language thus promises to offer key insights into what degree brains are specialized at birth or become so as we mature, how much we are shaped by exposure, and to what degree the neurological changes happening inside us are reflected in outward behavioral milestones as children begin to sing, clap, and speak. The present chapter will track developmental milestones from birth throughout childhood and will examine developmental disorders of music and language. Delving into these questions begins with a study of how these means of communication operate around the globe.

19.2 MUSIC AND LANGUAGES AROUND THE WORLD

How much does acquiring one's native culture differ from place to place? How varied is the adult world? When it comes to music, the answer is: a lot. For instance, traditional Western tonality distinguishes between consonant combinations of tones that are "pleasing" and "stable," versus dissonant ones that are "frictional" and "unstable." However, in Bulgarian diaphonic folk music, the choir's singing is often characterized by lines moving in parallel seconds – a stream of dissonances unacceptable in traditional Western tonality (Seskauskaité, 2004). In the polyphonic music of the San people and Aka pygmies in Africa, the superposition of patterns produces pan-scalar sonorities, in which combinations of notes from the source scale sound freely together, without regard for Western notions of consonance and dissonance. In traditional Western music, no matter how many dissonances might occur, the music will eventually resolve to a final consonance; in contrast, a Banda-Linda trumpet song might end with all of the trumpets – each tuned to a different pitch – sounding together in a giant cluster (Agawu, 2016). Balinese gamelan is even more idiosyncratic: Balinese flutes are often in completely wayward tunings with regards to rest of the ensemble; and Balinese singers can change mode depending on the character or the scene, independent of the instruments with which they are performing (McPhee, 1966).

There is also considerable variation in tuning tolerance. Western classical music puts a premium on playing in tune. Meanwhile, in Central Africa's version of the pentatonic scale, a scalar step may sometimes be as much as a half-step off by Western standards of tuning (Arom, 1991): even within a single African village, it may be hard to find "different ensembles of instruments capable of playing together, precisely because each ensemble is tuned to a different system" (Tracey, 1958, 16). Commenting on this, Agawu remarks that there is always "the possibility that our 'thirds' will seem closer to flattened fourths on Tuesday or our perfect fifths will be indistinguishable from tritones on Sunday" (Agawu, 2016, 268). In Javanese *campursari*, a modern hybrid of Eastern and Western instruments, incompatible tunings are often juxtaposed. Confronted with these heterodox tunings, audience members "listen to the keyboard and other non-gamelan instruments at the same time, but *not* to their not-being-in-tune. They try to ignore the wholeness of the whole, because the whole lacks wholeness" (Mrazek, 1999, 67).

Siberian nomads may be the furthest removed from Western practice. In contrast to the pitch-centricity of the West, Siberian tribes "perceive music in terms of changing timbral colors" (Nikolsky, 2020, 133). Asked to imitate the nursery rhyme "Mary Had a Little Lamb," a Khomu jaw harp musician will "translate" the melody into timbral inflections (Nikolsky et al., 2020).

When pitch is relevant at all, it is treated as an accessory . Nikolsky notes that "timbral music does not observe 'wrong notes': informants are puzzled by questions about musical mistakes, as they believe that any expression is 'right'" (Nikolsky et al., 2020, 10). When a Yakut singer is asked to repeat their song, "they reproduce only the melodic contour and the rhythm – the exact intervals between the adjacent notes of the same tune change" (Alekseyev, 1976, 148). If asked about the pitch differences, performers usually become surprised and deny any difference, reaffirming that the music is exactly 'the same'" (Nikolsky, 2020, 148). Likewise, Beliayeva-Ekzempliarskaya (1925) and Antonisha (1939) found that children and adolescents who "lacked exposure to classical music could not detect a harmonic mismatch when a well-known melody was performed against accompaniment in a wrong key" (Nikolsky et al. 2020, 4).

Nikolsky notes that nomadic music is a solitary repertoire (Nikolsky et al., 2020). He suggests that, while pitch-centric music is optimal for group participation, timbral variations – much like speech—become indistinguishable if more than one person is playing. For those who view community bonding as primarily responsible for the genesis of music-making (Savage et al. 2020), these nomads offer a cautionary note: in their ancient culture, music is used for self-regulation and intimate connection (Nikolsky et al., 2020).

Given this diversity, what unites the world's repertoires – from a Beethoven symphony, to Javanese gamelan and a Khomus jaw harp solo – is "creative play with sound," in which there is attention to the acoustic parameters of sound such as pitch, rhythm, timbre, and volume irrespective of any referential meaning (Brandt, Gebrian & Slevc, 2012). Humans have ears and imagination; out of those, an astonishing range of music is made.

Still, the degree to which music cognition is biologically constrained is hotly debated. One way of studying this is to look for musical behaviors and interpretations shared across cultures. Accessing a large database of indigenous music, Mehr et al. (2019) found that "what should be universal about music is not specific melodies or rhythms but clusters of correlated behaviors, such as slow soothing lullabies sung by a mother to a child" (Mehr et al., 2019, 366). Cross similarly highlights lullabies: "Perhaps the best evidence for the universality of music has been found in the universally musical qualities of caregiver-infant vocal interactions" (Cross, 2009, 178). But even there, there are outliers: lullabies are generally melodically simple, repetitive, and gradually get softer; however, the BaYaka Pygmies of Central Africa will respond to a restive baby by "yodelling even louder while rhythmically patting the baby's back" (Trehub et al., 2015).

Musicality involves an ineffable mixture of biological primitives, cultural conditioning, and personal experience that is particularly challenging to quantify. Cross-modal associations illustrate this delicate balancing act. They are clearly culturally informed: for instance, in the West, pitches are "high" and "low;" but for Farsi speakers in Persia they are "thin" and "thick" (Dolscheid et al., 2013), for the Kreung of Cambodia they are "tight" and "loose" (Parkinson et al., 2012), for the Balinese and Javanese they are "small" and "large," for the Suya of the Amazon they are "young" and "old," and for the Bashi of Africa, they are "weak" and "strong" (Eitan and Timmons, 2010). As Eitan and Timmons note, these metaphors are more than "convenient figures of speech" (405): they play a fundamental role in how music is internalized. Thus, in the West, we speak of "soaring melodies," "sinking bass lines," and "roving harmonies."

Yet, as startling as these differences are, the ways pitch is represented share an underlying correspondence of "less" and "more" of some variable (height, size, age, etc.) and, in numerous cases, are clearly stimulated by the acoustic fact that larger sound generators tend to produce lower frequencies. Indeed, Walker et al. (2010) found that 3- to 4-month-old preverbal infants looked longer at stimuli in which higher sounds were synchronized with higher and brighter images than when those correspondences were inverted. Researchers have also found that, at least in some cases, listeners enculturated in one metaphor framework could be prompted to temporarily "jump ship" to another one when coordinated with visual cues: for instance, Dutch subjects could be primed to associate pitches with thickness like native speakers of Farsi do; and the Kreung could be primed to associate them with height as we do in the West (Dolscheid et al., 2014; Parkinson et al., 2012). Thus, there may be innate predispositions that still allow for considerable flexibility.

Thanks to the fecundity of human imagination, music fulfills many different functions: entertainment, ritual, healing, soothing the young. For the Kaluli of New Guinea, it is a way of communing with the dead (Cross, 2003). The flute performances of the Mafa of Cameroon are displays of stamina and fitness (Fritz, 2009). Siberian nomads use timbre-based personal songs as a way of proclaiming their identity (Nikolsky et al., 2020); indeed, the disintegration of a personal song – characterized by "intense timbral modulations…raving-like, occasionally shrieking, moaning, and clapping in metric disarray" (8) is a sign of arctic hysteria, a psychological disorder that afflicts nomads as a result of the cold and darkness. Stone (1982) and Cross (2003) point to cultures in which music is so bound up with dance and other social exchanges that there is not a separate word for it. Stone remarks that "The isolation of musical sound from other arts proves a Western abstraction" (7) and calls for caution when studying the music of other cultures.

Taking all of this into account, musicality is highly susceptible to exposure. That exposure may begin in the womb, but we are not born with a fully mature music cognition network. Rather, like software engineers assembling packages of code into longer algorithms, the brain pulls together the necessary processing resources based on the stimuli to which it is exposed. For one's native music, that happens throughout childhood (Hannon & Trainor 2007) and differs from place to place and across the span of history.

What about language? Children with normal hearing acquire language through speech, and speech is a vocal performance, a "concert of phonemes and syllables, melodically inflected by prosody" (Brandt, Gebrian & Slevc, 2012). Those spoken improvisations also come in a remarkable variety: each language creates its own sound-world, distinguished by idiosyncratic patterns of timbre, rhythm, and pitch. First off, languages differ in their phonemic inventories: the "click" language !Xóo, spoken by a small community in Botswana and Namibia, has a staggering one hundred and thirty consonants, twenty-eight vowels, and three different tones; meanwhile, Aita Rotokas, spoken on the Pacific island of New Guinea, has only nine consonants and five vowels (Robinson, 2006). Rhythmically, languages such as English and German are stress-timed, meaning that stressed syllables are lengthened, whereas French, Turkish, and Mandarin Chinese are syllable-timed, with syllables spoken in roughly equal duration. Languages also make use of pitch in different ways. In Southeast Asian languages such as Mandarin, pitch contours are determinants of meaning. Meanwhile, register tone languages such as Serbo-Croatian, Swedish, and Yoruba rely on pitch oscillations, typically between two or three tones, to distinguish different words. Overall, scholars estimate that between 60-70% of the world's languages incorporate pitch inflections into their spoken vocabulary (Best, 2019). Thanks to its four register tones, as well as all possible combinations of paired tones and two falling three-tone sequences, the Chatino language of Mexico may be the world's most complex tonal language (Pride, 1963). A smaller subset, including English, French, and German, are described as non-tonal: pitch is only used for punctuation, emphasis, and affect.

Underneath this variety in the music of speech, Chomsky (1965) and others have theorized that all languages share a "universal grammar:" they are made up of atomic units which can be unendingly recombined according to syntactic rules; and they are characterized by recursion, the capacity for embedded structure. Chomsky has argued that this fundamental linguistic consistency indicates that, thanks to an evolutionary leap, humans are born with a "language faculty," and that children acquire language easily in spite of a poverty of stimulus because they are innately prepared for it.

However, much as Nikolsky has brought to light musical traditions far removed from Western culture, Wray and Grace (2007) have studied languages spoken in remote areas. They draw a contrast between outward facing societies that routinely welcome strangers and these insular communities. Some hunter-gatherer enclaves do not have a word for "word." In place of "rule-based compositionality," they speak in elliptical turns of phrase understood by intimates but opaque to outsiders. Everett has chronicled the idiosyncrasies of the Piraha people, who communicate "as much by singing, whistling and humming as they do by consonants and vowels" (Everett, 2005). Everett observed that Piraha speakers do not generalize beyond their immediate experience, and that their language lacks words for numbers and colors, makes scant use of pronouns, and shows no evidence of embedding.

Hermetic languages may be taught differently too. While we are accustomed to language instruction working from the bottom up – vocabulary first, worked into sentences – Laycock (1979) and Thurston (1987) report cases of remote languages being taught from the top down – that is, by turns of phrase. Laycock describes being instructed by a village elder in a Papuan village: "Only certain kinds of errors were corrected, and then by repeating the whole sentence, not dissecting it…No attempt was made to explain any of the morphology…or even to separate out individual words from sentences, except in the case of important nouns" (91) such as those for tobacco, fire, and water. Working with the same community, Thurston likewise observed that language instruction was geared almost entirely towards colloquialisms, with an eye towards their use in actual social situations.

Henrich (2020) has theorized that humanity has gradually progressed from kinship-based societies to more heterogeneous ones, giving birth to a modern mentality that puts a premium on self-reliance, trust in strangers, and literacy. Wray and Grace observe that our familiar rules-based languages are optimal for communities in which new arrivals need to quickly be brought up to speed. As with Piraha, they find recursion rare in the hermetic speech of tribal familiars, and argue that written language – a later development and one found only in a handful of the world's 7,000 languages – may be responsible for the prevalence of nested structures in language. Thus, Chomsky's universal grammar may be most fully realized in heterogeneous, outward facing societies, marking it as a cultural, rather than genetic, evolution.

Meanwhile, it is not clear to what degree children are attuned to the regularities of more rules-based languages. Esperanto is a synthetic language designed to be as regular as possible; its goal was to promote world peace by creating a true *lingua franca*. All verbs are conjugated the same way; adjectives such "felica" (happy) and "laca" (tired) all end in with the letter "a"; words are stressed on the penultimate syllable. It may be the most orderly language currently in use. Yet Bergen found that children raised on Esperanto did not deduce the generative features of the language and instead made often inexplicable errors (Bergen, 2001). Bergen found this surprising "because it seems to contradict bioprogram and other universalist predictions" (580). If the children were innately primed for a "universal grammar," why were they missing the obvious? For Wray & Grace, this supports the view that children "acquire language without recourse to full systematicity" (555) and, for processing efficiency, tend initially to pay more attention to phrases and sentences than individual words.

In sum, music and language are creative responses to social circumstances and, as such, are tailored to the cultures in which they arise. There are features that may be widespread and even commonplace, but it remains an open question to what degree it is possible to generalize. All of this has implications for development. Nature makes it possible for a child to be born in any community and acquire its native music and language. Children hit certain analogous milestones; beyond that, differences in upbringing, cultural priorities and expectations, and the nature of what they are learning must always be taken into account. For instance, consider that

> Tuvan children learn early to vocally imitate typical environmental sounds with amazing precision, adopting learned timbral distinctions for the creation of their own music (Levin and Suzukei, 2006, p. 85–7) – very much like Western children model their vocal improvisations upon commonly heard tunes (Bjørkvold, 1992) (Nikolsky et al., 2020, 3).

Given that language and music are human inventions, we wouldn't expect our biology to prescribe that we learn one faster than the other. Indeed, the preponderance of evidence indicates that music and language are deeply entangled in early childhood, and development proceeds along parallel tracks. Infants begin their lives listening to speech and music in very similar ways.

19.3 MUSIC AND LANGUAGE ACQUISITION

19.3.1 Language and Music Abilities in Infants

Despite being born with underdeveloped brains as compared to other species, the foundations of human aural cognition begin in utero (McMullen & Saffran, 2004). As a result, infants display astute perceptual abilities and are sensitive to a wide variety of linguistic and musical contrasts at birth. This precocious perceptual acuity in the language domain has been taken as evidence that language learning is an innate ability, but infants are sensitive to the *sounds* of language at birth, not meaning or syntax. The parameters of pitch, rhythm, and timbre in particular are of utmost importance in helping infants learn how the sounds of their native language map onto meaning. Although these building blocks are used differently in language versus music (as understood by adults), the progression of infants' sensitivity to these parameters is strikingly matched over the course of development.

Using timbre to discriminate between sounds plays a critical role in the earliest stages of aural development. Infants are famously able to discriminate between the phonemes of any language up until around 6 months of age (Werker & Tees, 1984, etc.). Phonemic perception relies on a sensitivity to acoustic spectra and rapid temporal changes on the order of 25-50ms. Fine-grained perception during this narrow time frame allows us to hear differences in consonant sounds: *t* versus *k* for instance (Tallal & Piercy, 1973; Rosen, 1992; Telkemeyer et al., 2009). This same time window is also critical for musical timbre perception: a harp versus a piano playing the same pitch (Hall, 1991; Hukin & Darwin, 1995; Robinson & Patterson, 1995; Shepard, 1980). Vowel perception relies on sensitivity to the acoustic spectra, a skill also needed when listening to a sustained sound from a musical instrument after the initial attack. Phonemic discrimination, therefore, is largely a sensitivity to timbre. Less research has been conducted on timbre sensitivity in infants as compared to other acoustic parameters, but 3- to 4-day old newborns are able to use timbre to organize auditory streams, much like adults (McAdams & Bertoncini, 1997). Even infants born 10 weeks premature can recognize both a change in phoneme (ba vs. ga) and a change in speaker, although the change in phoneme elicits a more robust response (Mahmoudzadeh et al., 2017). In fact, infants appear to be more adept than adults at detecting small changes in timbre under certain conditions (Lau et al., 2020). This sensitivity to timbre remains as they develop, with 6-month-old infants having long-term memory for the timbre of folksongs, and 7- to 8.5-month-old infants showing evidence of being able to differentiate tones that differ only in their spectral structure (Trainor, Wu & Tsang, 2004; Trehub, Endman & Thorp, 1990). In fact, timbre seems to be such a salient feature of sound to infants that it can interfere with their ability to recognize familiar sounds on the basis of other features: infants take longer to learn words when spoken by different speakers versus just one speaker (although note that multiple speakers seem to help infants learn the distinction between similar phonetic categories. Rost & McMurray, 2009), and – unlike adults - they seem not to recognize a familiar melody when played on an instrument different than the one they were familiarized with (Jusczyk, Pisoni & Mullennix, 1992; Trainor et al., 2004).

In addition to their fine-grained timbre awareness, newborns are also sensitive to the rhythm of language. They are able to distinguish between two different languages based on rhythmic class – stress-timed, syllable-timed, or mora-timed – whether or not the contrast includes their native language (Nazzi, Bertoncini & Mehler, 1998). Newborns also show a preference for their native language, but this appears to be a preference for the rhythmic class of their native language since they are unable to distinguish their native language from others of the same rhythmic class until 4 months of age (Moon, Cooper & Fifer, 1993; Bosch & Sebastian-Galles, 1997; Gervain & Mehler 2010; Nazzi et al., 1998). Rhythm also seems to help facilitate memory for language in very young infants: two-day-old infants show a larger response to errors in the words of nursery rhymes when they were familiarized with the nursery rhyme spoken in rhythm versus spoken as simple prose (Suppanen et al., 2019). Adding to the evidence that rhythm is a highly salient feature of language for infants, 4-month-old Czech infants were able to discriminate between Czech and a modified version of Czech where the rhythm was altered, but no other features were changed (Paillereau et al., 2021).

Finally, pitch also plays a role as infants learn how their native language is composed. Within the first few days of their lives, newborns appear to be able to discriminate affective prosody, as well as the characteristic prosody of their native language (Cheng et al., 2012; Friederici, 2006). The prosody of their native language even influences the "melody" of their crying (Mampe et al., 2009; Prochnow et al., 2019, the melodic complexity of which increases over the first few months of their life. Over the first four months, both the melodic contour and the interval content increase in complexity, and infants who do not show this developmental trajectory in their crying exhibit poorer language performance two years later (Wermke & Mende, 2009; Wermke at al., 2007; Armbrüster et al., 2020).

Caregivers seem to know intuitively that the melody, rhythm, and timbre of language are important and salient to infants. Motherese (or infant-directed speech) – the exaggerated timing, melodic contours, and timbral changes we use when speaking to babies – appears to be a human universal (Piazza et al., 2017). There has been significant debate over the purpose of motherese, with some arguing it is primarily a means of emotional communication (Trainor et al., 2000), while others point out that it is also a way of holding infants' attention (Fernald, 1989). Parents appear to vary the pitch contours in their speech depending on whether the infant is smiling and/or maintaining eye contact or not (Stern et al., 1982). Other researchers note that adults lengthen the vowels in content words, and exaggerate word and sentence boundaries when speaking to infants, emphasizing its role in helping infants learn and understand the different components of their native language (Kuhl et al., 1997; Saint-Georges et al., 2013). Saint-George and colleagues note that, "Mothers adjust their infant-directed speech to infants' age, cognitive abilities and linguistic level," (Saint-George et al., 2013, p. 9), and infants prefer motherese over normal speech (Fernald, 1985), highlighting the fact that this special, highly musical way of speaking likely has multiple functions.

Similar to the exaggerated way we talk to infants using motherese, we also sing differently to infants. Depending on the function of the song (a play song versus a lullaby, for instance), caregivers will exaggerate the timing, pitch, and volume contours of their singing when singing to infants (e.g. Trainor et al., 1997; Nakata & Trehub, 2011). Moreover, this special way of singing to infants also appears to be a human universal, much like infant-directed speech (l'Etoile, 2006). And just as infants prefer infant-directed speech, they also prefer infant-directed singing (Trainor, 1996; Masataka, 1999).

Despite preferring singing that exaggerates its acoustic features, infants show exquisitely sensitive perceptual abilities when it comes to musical stimuli, matching their precocious abilities with the rhythm, melody, and timbre of speech. As mentioned earlier, they appear to outperform adults when detecting small timbre changes when the pitch is also fluctuating unpredictably. The reverse of this also appears to be the case: infants can detect small pitch changes better than adults when the timbre is fluctuating unpredictably (Lau et al., 2020). Newborns can also detect a deviant pitch when the timbre is changing, and 5- to 8-month-old infants can discriminate pitch changes as small as a third of a half-step (Háden et al., 2009; Olsho et al., 1982). Newborns are also sensitive to pitch patterns and can use these patterns predictively (Háden et al., 2015).

As discussed earlier, different cultures map the relationship between pitches differently ("low" versus "high"; "thick" versus "thin"). Infants whose culture uses the high/low distinction appear to be equally sensitive to a thick/thin association (Dolscheid et al., 2014). Just as infants are sensitive to the phonemes of all languages, they may also be equally equipped to perceive pitches across different continua that emphasize different components of a tone (its pitch versus its timbre). More cross-cultural work will be needed to tease out these abilities and distinctions.

When it comes to rhythm, newborns are able to detect the beat in music, and are sensitive to an important rhythmic event being removed from a drumming pattern (Winkler et al., 2009). They can also distinguish small differences (60-100ms) in the length of two notes (Čėponienė et al., 2002; Cheour et al., 2002), and at two months, they are able to group separate sounds into rhythmic "units," much like adults (Demany, McKenzie & Vurpillot, 1977). Far more research has been done on the linguistic capabilities of newborns as compared to their music perception abilities, but the existing research indicates they are equally sensitive to the details of both music and language.

19.3.2 Music and Language Co-Development through 24 Months

As infants gain experience with their native language, music, and culture, their perceptual abilities become more refined and culture specific. This is also the case with their abilities to produce music and language, which are highly intertwined even up to two years of age (Dowling, 1984). For both music and language, however, this refinement is strikingly similar, especially given how much more emphasis language receives in Western culture (where the vast majority of research in these domains has been conducted) versus music.

Infants retain their famous ability to discriminate all phonemes until about 6 months, after which they begin to lose their ability to differentiate non-native contrasts (Cheour et al., 1998; Rivera-Gaxiola et al., 2005). Even at 6 months, however, they already show evidence of being attuned to the vowel sounds in their native language over other languages (Kuhl et al., 1992). By 8 months, they no longer distinguish non-native vowel contrasts, and by 10-12 months they stop discriminating between non-native consonant contrasts as well (Polka & Werker, 1994; Werker & Tees, 1984). When it comes to music, 6-month-olds can detect changes in a melody made up of both Western scales and Javanese scales, the latter of which is difficult for Western adults (Lynch et al., 1990). But like with linguistic contrasts, by 9 months, Western infants behave like adults: they can no longer detect changes in Javanese melodies (Lynch & Eilers, 1992). At the time of this writing, far more research has been conducted comparing phoneme perception across many different languages versus comparing infants' perceptions of different non-Western musical contrasts. Western infants have not been tested to see whether they can detect fine-grained contrasts important to timbre-based music, for instance, nor have infants from timbre-based musical cultures been tested to probe their pitch discrimination. Presumably, 6-month-old infants would be able to detect changes and contrasts in any culture's music (and then lose this ability due to enculturation three to six months later), but more cross-cultural work and sensitivity to cultural differences in this area are needed to further investigate these possibilities.

The same perceptual narrowing seen with phonemic discrimination is also seen in the rhythmic domain: English-speaking infants show a preference for stress-initial words by 7.5 months (Jusczyk, Hohne & Bauman, 1999); are especially sensitive to the stress patterns in their native language by 9 months (for a review, see Jusczyk, 2000); and are reliant on the rhythmic characteristics of language to segment words until 10.5 months, at which point they are able to use non-stress-based cues to segment words (Jusczyk et al., 1999). In the musical domain, 6-month-olds can detect changes in both simple and complex meters (which are rare to non-existent in traditional Western music), something that is difficult for Western adults (Hannon & Trehub, 2005a). By 12 months, however, Western infants behave like Western adults, losing the ability to detect changes in complex meters. In contrast, Balkan infants, whose traditional music includes complex meters, retain this ability (Hannon & Trehub, 2005b).

The vast majority of research on language perception has been done in Western cultures with non-tonal languages, but there is a growing body of research that investigates the development of linguistic competence in cultures that speak tonal languages (for reviews, see Lee & Cheng, 2020; Tsao & Liu, 2020). Six-month-old infants learning English show sensitivity to lexical tones similar to Chinese-learning infants, but at 9 months, the English-learning infants show a decreased sensitivity, mirroring the perceptual narrowing seen for non-native phonemic contrasts (Mattock & Burnham, 2006). Interestingly, Chinese infants remain sensitive to Thai tonal contrasts at 9 months, even though these are not present in their native language (Mattock & Burnham, 2006), suggesting that sensitivity to tonal contrasts in general is preserved for longer in infants from tone-language speaking cultures. However, sensitivity to language-specific tonal contrasts is evident from 4 months of age in infants from tonal language cultures (Yeung, Chen, & Werker, 2013).

Research on perceptual narrowing in tonal languages shows a more complex developmental trajectory than that revealed by studies on non-tonal languages. For instance, Mandarin has four tone contours, labeled T1-4. Some tonal contrasts are easier to perceive than others: around 12 months of age, Mandarin-learning infants can accurately discriminate between the easier tonal contrast (T1 versus T3), but not the more difficult ones (T2 versus T3, or T2 versus T4). However, infants appear to have some sensitivity to the easier T1/T3 contrast even before it shows up in their behavioral responses: infants as young as 6 months of age display an adult-like mismatch negativity (MMN) response when confronted with an oddball among repeated stimuli for the easier tonal contrast, showing that their brains are sensitive to the difference between the two (Cheng & Lee 2018; Tsao, 2008). It is not until around 18 months that an immature version of the MMN response (P-MMR) is seen for the more difficult T2 versus T3 contrast (Lee & Cheng, 2020). Interestingly, some studies show *increasing* sensitivity to tonal contrasts in infants in non-tonal language cultures between 14-18 months of age (Götz et al., 2018; Liu & Kager, 2014). This U-shaped developmental curve is supported by evidence that non-tonal language speaking adults can show good lexical tone discrimination (Götz et al., 2018).

In all languages, these developments in perception set the stage for the understanding of syntax and semantics. Word order rules are an important feature of syntax in many languages, something infants are sensitive to by 8 months, but at this age it is largely prosody and word frequency that support this ability (Gervain et al., 2008; Hochmann, Endress & Mehler, 2010; Nespor et al., 2008). Infants start babbling around 7 months of age (Locke et al., 1995), and at 9 months, they appear to understand their first words. Shortly thereafter, they begin talking, typically sometime between 11 and 13 months (Friederici, 2006). Once infants start talking, meaning and syntax take over the focus of learning, and they are able to communicate both linguistically and musically in much more sophisticated ways. Infants experience an enormous increase in their vocabulary between 18 to 24 months, matched by a high point in their syntactic learning between 18 to 36 months (Friederici, 2006; Kuhl, 2010). Around 18 months, they are speaking in two-word utterances (Friederici, 2006; Gervain & Mehler, 2010) and at least one study shows evidence that at 15 months, infants are capable of imitating short song fragments (Benetti & Costa-Giomi, 2020).

Note that the further away the features of language are from music (semantics and syntax), the later they develop. We focus on musical features first, refining our perceptual abilities to align with the priorities of the culture we grow up in. For most cultures, this refinement includes a distinction between music and language. As we continue to fine-tune our understanding, music and language become more and more distinct in our perception, our production, and in how the brain treats these two types of acoustic stimuli.

19.3.3 Music and Language Co-Development in Childhood

Children continue to refine their perception of both music and language throughout childhood, reaching adult levels in both domains by about 12 years of age (Costa-Giomi, 2003; Hahne, Eckstein & Friederici, et al., 2004). They also develop in their ability to speak and create music, and these two areas become more and more distinct. Musical development has often been seen as following a slower and more difficult trajectory than language development, but this is because of mismatched expectations: musical development is often measured against that of professional musicians rather than the general adult population. When musical development is compared to the abilities of non-musician adults, it becomes clear that acquiring the perceptual and production abilities of one's musical culture are no more difficult or protracted than acquiring linguistic abilities.

Before the age of two, singing and speech are not reliably distinguishable from each other in young children's vocal output (Dowling, 1984). But once children reach the age of two, syntax and semantics take over in their development, and music and language become more distinct from each other. As early as 14 months, toddlers begin to show the rudiments of syntactic understanding, abilities that are continually refined throughout early childhood (Skeide & Friederici, 2016). By age 6, they have mastered the basic syntax of their native language (with more complex constructions are still being learned and refined until around age 10), although haemodynamic activity shows that semantic and syntactic processing largely overlaps until around 7 to 9 years of age (Scott, 2004; Nuñez et al., 2011; Hahne et al., 2004, Eckstein & Friederici, 20046; Skeide & Friederici, 2016).

Knowledge of Western musical syntax is shown by knowledge of scale and key membership (which notes sound like they belong in a given key), as well as conventions of harmony. By the age of 3, children show evidence of rudimentary knowledge of key membership and harmony (Corrigall & Trainor, 2009), but similar to language, this knowledge is very much dependent on context until at least age 5 (Koelsch et al., 2003; Trainor & Trehub, 1994; Trehub et al., 1986). Four to 6-year-olds are faster at detecting changes in diatonic melodies than non-diatonic ones, and are faster and more accurate at making timbre determinations for diatonic notes than non-diatonic ones (Politimou et al., 2021), evidence of implicit knowledge of key membership. Like adults, five-year-olds also show an electrophysiological response (an ERAN) to unexpected chords, although they can't detect a melodic change that implies a change of harmony (Koelsch et al., 2003; Trainor & Trehub, 1994). This ability arrives around age 7, when children's knowledge of tonal structures is comparable to that of adults. Their sensitivity to implied harmonies, however, continues to develop until around age 12, mirroring their development in sensitivity to more complex linguistic syntax (McMullen & Saffran, 2004; Speer & Meeks, 1985; Trainor & Trehub, 1994; Costa-Giomi, 2003). Children's pitch discrimination abilities also reach adult levels around this same time, between the ages of 8-10 (Werner & Marean, 1996).

Children growing up in tone-language cultures reach some of these developmental milestones in music sooner than children in non-tonal language cultures. For instance, 3-5-year-old children learning Mandarin are better at discriminating pitch differences than English-learning children (Creel et al., 2018). However, both groups showed similar performance on timbre contrasts. Timbre continues to be quite salient to children, facilitating memory for a paired visual stimulus more easily than a paired melodic contour (Creel, 2016). Even adults continue to show evidence of memory facilitation based on timbre (Halpern & Müllensiefen, 2008; Radvansky & Potter, 2000). In fact, this is not just true of music: timbre may be more important than pitch, even in tonal languages. Whispered speech is perfectly intelligible, even in tonal languages, where listeners are able to achieve 85% accuracy in understanding what was said (Abramson, 2015). Remove timbral distinctions, however (making all of the phonemes the same) and speech becomes completely unintelligible, even in non-tonal languages (Patel, 2010).

It is in early childhood that it becomes easier to study music and language production in addition to perception. Development in these domains is also remarkably similar between music and language. Two-year-olds' speech is simple, and they tend to eliminate function words, but not content words (Gerken, Landau & Remez, 1990). By the age of 6, however, their speaking abilities are comparable to adults (Scott, 2004). In the realm of singing, 2-year-olds can repeat brief melodic phrases that have an identifiable rhythm and contour, and 3-year-olds sing what are called "outline" songs – getting the gist, but leaving out many of the finer details, much like their speaking at this age (Dowling, 1999; Davidson, 1994; Hargreaves, 1996; Moog, 1976). The traditional view of singing abilities in toddlers suggests that they are more accurate in reproducing melodic contours than precise pitches or intervals and that they sing in a limited range (i.e. Flowers & Dunne-Sousa, 1990), but more recent research with carefully constructed testing games suggests that 3-year-olds are able to sing much more accurately and with a larger range than previously believed (Gudmundsdottir, 2020). Researchers have noted that young children tend to switch to a new key in the middle of a melody to accommodate their vocal range (e.g. Flowers & Dunne-Sousa, 1990). Gudmundsdottir points out, however, that children may not understand that precise pitches and a stable tonal center are important values in Western music (Gudmundsdottir, 2020). This recalls the singing of Balinese musicians discussed earlier, who adjust their tuning depending on the character of the scene, irrespective of the instrumental accompaniment. Western adults may consider adhering to a stable tonal center and precise pitches to be a self-evident feature of good singing, but it is clear that this is a cultural value, not an inherent feature of music. Furthermore, children may be shy or reluctant to sing in front of adults, adding an additional difficulty in accurately measuring children's singing abilities (Gudmundsdottir, 2020).

Singing ability continues to develop through the elementary school years, with ability seeming to peak around 6th grade (Demorest & Pfordresher, 2015). Interestingly, in a study that compared the singing ability of kindergartners, 6th graders, and adults, the adults performed more like the kindergartners (Demorest & Pfordresher, 2015). Singing can clearly be improved with practice and training, even in very young children (e.g. Welch et al., 1991), and because non-musician adults in Western cultures tend not to sing once they are finished with music classes in school, it is not surprising that ability would deteriorate.

Musical skills also appear to support and enhance language abilities in a variety of areas. Significant correlations have been found in particular between children's rhythmic abilities and perception, and various measures of linguistic competence. Five- to 7-year-olds show a strong correlation between performance on rhythmic discrimination tasks and their grammatical production abilities (Gordon et al., 2015). Rhythmic perception and production have also been found to be associated with phonological awareness in 3- and 4-year-olds (Politimou et al., 2019; Bonacina et al., 2020). At this same age, melodic perception predicts grammar acquisition, with children who have more informal musical experience at home showing a stronger association (Politimou et al., 2019).

19.3.4 The Bias of Pitch-Based Cultures

Nearly all of the studies cited above on the development of music and language abilities were done by Western researchers on mostly Western subjects. The uniqueness and degree to which this population is not representative of human universals, that it is indeed WEIRD, has been discussed at length by other researchers (Henrich et al., 2010). Specifically in terms of music, Western researchers bring a pitch-centered bias to their view of music, but not all cultures center pitch in their music-making. In addition to Siberian nomadic song discussed earlier, the traditional music of Tuva and other Turkic ethnic groups (in Russia and inner Asia), shakuhachi music in Japan, and didgeridoo music of Australian aboriginal cultures are all timbre-based, just to name a few. Pitch is of marginal importance in these musical traditions, but precise timbral distinctions are paramount. When Nenet musicians listen to Westerners reproduce their music, it sounds hopelessly "out of tune" because Westerners cannot accurately perform the timbral variations inherent in their music (Ojamaa 2005, cited in Nokolsky 2020).

Western music relegates timbral changes to second-class status, but it can be valuable to view Western subjects' musical abilities through the lens of timbral primacy. Children who cannot match pitch, maintain pitch, sing accurate intervals, or maintain a consistent tonal center have traditionally been deemed poor singers (i.e. Welch, 1983; Welch, 1985; Flowers & Dunne-Sousa, 1990). Rather than labeling them as deficient, it may be more productive to recognize that the alignment of their abilities with the musical priorities of their community are still developing. Recall that Tuvan children learn to imitate environmental sounds with their voices, producing extremely precise timbral distinctions with amazing accuracy (Levin & Suzukei, 2006). Likewise, in cultures ranging from the Sioux in North America to populations in Siberia, the "personal song," in which each member of a community performs an ever-changing and developing timbre-based song, is widespread. This personal song, accompanying life's activities, is extremely elastic in terms of pitch, but has a recognizable contour and timbre particular to each individual (Nikolsky et al., 2020).

This kind of personal singing to accompany one's day-to-day life is highly reminiscent of the spontaneous, improvised singing documented in 3- and 4-year-old children. Dean (2020) equipped 15 toddlers with recording devices that could record up to 16 continuous hours, and captured an endearing kaleidoscope of singing throughout the day, particularly when the children were alone. All of the children in the study sang and most of the singing was improvisatory, incorporating nonsense words, syllables, humming, chanting, and vocal play. This differed from their singing when interacting with others; in the latter case, they were more likely to sing songs they had *learned*, or at least use words that conveyed meaning (rather than the exploratory improvisations of solitary singing). The specificities of pitch appear to be the least important ingredient in these songs, echoing Gudmundsdottir's statement that young children may not have learned yet that exact pitches and stable tonal centers are an important musical value in their culture's music (Gudmundsdottir, 2020).

In one of the few cross-cultural studies of singing accuracy, researchers found that the Tsimane' people of the Bolivian Amazon could sing the correct contour in echoing a model, but not the correct pitches or intervals (Jacoby et al., 2019). To encourage them to sing more accurately, the Tsimane' were given "explicit" feedback intended to improve their pitch accuracy. This "explicit feedback," however, consisted only of prerecorded words such as "Excellent," "Good," "Okay," or "Try again." The feedback failed to improve the pitch accuracy of the Tsimane', but it is not at all clear that the Tsimane' understood the feedback as referring to their pitch accuracy (versus volume, timbre, contour, rhythm, or another parameter). Researchers coming from a pitch-centric music culture take for granted the importance of pitch accuracy, but this value should not be assumed. The primacy of pitch is often seen as a universal standard, but it may be the other way around: timbre, contour, and rhythm may be the most important ingredients of music universally, with specific and precise pitches being an anomaly. Echoing calls for more cross-cultural work in music cognition (Jacoby et al., 2020), a more nuanced understanding of the relative value of various music parameters in different cultures' music will result in more meaningful research questions and results.

19.3.5 Emergent modularity

It can be difficult for adults to understand how music and language could start off as deeply intertwined in our experience as infants, given how easily we can distinguish between them now. Some have argued that the brain treats them separately, even in infancy, with the left hemisphere specialized for language and the right hemisphere specialized for music (e.g. Dehaene-Lambertz, Dehaene & Hertz-Pannier, 2002). However, a number of studies have found bilaterial activation for both music and language in the infant brain (Fava et al., 2014; Perani 2012; Dehaene-Lambertz, 2000; Kotilahti et al., 2010; Minagawa-Kawai et al. 2011). Some have found no differences in activation for speech versus music (Fava et al., 2014) and some have found right hemisphere lateralization for speech (Perani et al., 2011).

It appears that as we develop and undergo enculturation, functional specialization for language and music emerges over time (cf. Karmiloff-Smith, 1992; McMullen & Saffran, 2004). This would predict that auditory stimuli that *can* be understood as language would be left-lateralized in their processing only if and when they *are* understood as language. In fact, this is exactly what has been found. The degree of left-lateralization in adult second-language learners is a function of proficiency: the better someone is able to understand a second language, the greater the degree of left-lateralization (Dehaene et al., 1997; Perani et al., 1996, 1998). Sound contrasts that are not meaningful in a speaker's native language result in left-lateralization only after subjects undergo extensive training to hear the non-native contrast (Best & Avery, 1999; Zhang et al., 2009). Sine-wave speech, which initially sounds like meaningless whistles, can be perceived as speech after training. Only once it is heard as speech does it activate speech areas (Möttönen et al., 2006). Similarly, the whistled language *Silbo Gomero*, a version of Spanish that extracts its prosodic content, is understood clearly as speech in those who speak it. Those who are proficient with the language show brain activation in speech perception areas of the brain, activation that is absent in those who do not speak the language (Carreiras et al., 2005).

Even in everyday language, the musical components of speech are always present, helping to clarify meaning. Garden-path sentences, whose meaning can be quite ambiguous and confusing in print, are rarely misunderstood when spoken (Kreiner & Eviatar, 2014). Kreiner and Eviatar (2014) argue that prosody is the root of syntactic comprehension, noting that syntactic and prosodic boundaries largely correspond in speech, helping to aid comprehension. Just as infants use the rhythm and melody of speech to scaffold future syntactic and semantic comprehension, adults use these same features of spoken language both in speaking and listening to ensure communicative clarity.

Although music and language seem unambiguously separate to adults, it is clear that we begin life with them deeply intertwined. There are striking parallels between music and language perception, production, and development from birth through childhood. And indeed, the two domains continue to interact and inform each other even in adulthood. As stated above, more work needs to be done cross-culturally and on musical development in general before we can fully appreciate the entire range of abilities we are born with and acquire as we grow and mature.

19.4 LINKED DEVELOPMENTAL DISORDERS

As argued above, the striking parallels in musical and linguistic development suggest that these domains are only gradually differentiated, initially relying on the same underlying processes. But are musical and linguistic development truly linked, or do these similar trajectories instead reflect development of two cognitively/neurally distinct domains that have similar principles? Relevant evidence comes not just from typically developing children, but also from individuals with abnormal musical or linguistic development.

The most well studied deficit of musical development is congenital amusia, which is associated with deficits in pitch perception/memory and sometimes also with abnormal processing of rhythmic aspects of music. A condition mostly studied in adults, congenital amusia is not an auditory processing deficit *per se* (for example, auditory cortical responses to pitch are normal; Moreau, Jolicœur, & Peretz, 2013; Norman-Haignere et al., 2016), but instead seems to reflect a lack of conscious access to pitch information (see Peretz, 2016, for a review). Perhaps surprisingly, given the discussion above, the clear music processing deficits in congenital amusia are not accompanied with significant language deficits, which could suggest an early neural specialization for music processing (Peretz & Hyde, 2003).

However, it turns out that individuals with congenital amusia do show subtle deficits in processing pitch-based aspects of language. These include discriminating lexical tones (in amusic tone language speakers; Liu et al., 2021; Nan, Sun, & Peretz, 2010; Wang & Peng, 2014), discriminating and imitating intonation patterns (such as the rising pitch at the end of "Intonation matters?" vs. the falling pitch at the end of "Intonation matters."; Hutchins, Gosselin, & Peretz, 2010; Liu, Patel, Fourcin, & Stewart, 2010), and categorizing emotional prosody (Thompson, Marin, & Stewart, 2012; Pralus et al., 2019). Subtle abnormalities in speech perception extend to non-pitch-based aspects of language as well (Jones et al., 2009; Zhang, Shao, & Huang, 2017) and can impact on speech intelligibility in noisy environments (Liu, Jiang, Wang, Xu, & Patel, 2015).

Still, despite their (sometimes dramatic) musical deficits, amusics rarely report any speech processing difficulties in real-world contexts (e.g., Ayotte, Peretz, & Hyde, 2002). That is, the language deficits associated with congenital amusia noted above are relatively subtle and usually can only be observed in controlled laboratory tasks. This is probably because language and music involve many redundant cues, and these studies examined adults who have had ample developmental time to prioritize the cues that work best for them. That is, the prolonged maturation of our musical and linguistic abilities allows individuals with a pitch processing deficit to learn to rely more on non-pitch cues, thus achieving normal language performance despite recruiting different underlying neural processes (e.g., Jasmin et al., 2020; Lolli et al., 2015).

Abnormal language processing in congenital amusia is part of the story; equally importantly, there appear to be musical processing deficits associated with developmental *language* disorders as well (see Chapter X, this volume, for discussion). For example, the phonological processing difficulties associated with developmental dyslexia have often been attributed to more basic underlying auditory processing impairments (Hämäläinen, Salminen, & Leppänen, 2013), which likely impact music processing as well. Indeed, developmental dyslexia is associated with musical deficits, especially involving musical rhythm, in both children (Colling et al., 2017; Huss et al., 2011; Overy et al., 2003; Thomson & Goswami, 2008) and adults (e.g., Couvignou, Peretz, & Ramus, 2019). There is also evidence that musical training might mitigate the auditory deficits associated with developmental dyslexia (e.g., Flaugnacco et al., 2015; Frey et al., 2019). Of course, music and language difficulties in dyslexia do not seem inexorably linked; for example, consider the existence of dyslexic musicians with persistent phonological and reading deficits (Weiss, Granot, & Ahissar, 2014). Note, however, that the pattern of musical and auditory performance in dyslexic musicians is not entirely normal (Zuk et al., 2017), which may suggest that dyslexic musicians have learned to prioritize musical cues that rely less on the auditory features that are more problematic for them (just as amusics may achieve near-normal language processing by learning to prioritize more reliable linguistic cues over development).

Abnormal musical processing is associated with other developmental language deficits as well. Timing and rhythm processing, in particular, seems to be problematic across a range of (related) disorders (see Ladányi et al., 2020, for a review) including developmental language disorder (aka specific language impairment; Bedoin, Brisseau, Molinier, Roch, & Tillmann, 2016; Cumming, Wilson, Leong, Colling, & Goswami, 2015; Przybylski et al., 2013; Richards & Goswami, 2019) and developmental stuttering (Chang, Chow, Wieland, & McAuley, 2016; Wieland, McAuley, Dilley, & Chang, 2015). Abnormal language development does not only impact rhythm; developmental language disorder is also associated with deficits in pitch-matching and melody reproduction (Clément, Planchou, Béland, Motte, & Samson, 2015) and in the processing of harmonic structure (Jentschke, Koelsch, Sallat, & Friederici, 2008).

Related impacts to music and language emerge in other disorders not specific to music or language. For example, attention deficit hyperactivity disorder (ADHD) often involves (or is comorbid with) language impairment (Cohen et al., 2000), and is associated with abnormal rhythm processing (Puyjarinet et al., 2017). In contrast, musical processing seems mostly unimpaired in autism, even including understanding of emotional aspects of music (which is somewhat surprising given characteristic abnormalities in social processing; Heaton, 2009). However, the ability to discriminate affective musical states *is* linked to language abilities in autism (Heaton et al., 2008) and musical interventions in autism can lead to improvements in communicative behavior (Geretsegger, Elefant, Mössler, & Gold, 2014), suggesting developmental links between the social/emotional and communicative components of music and language (see also Chapter X, this volume).

Finally, congenital hearing loss obviously impacts music and speech processing, but congenitally deaf individuals can successfully learn full-fledged sign languages. At first blush, this could seem to be a problematic counterexample for the tight music/language links defended here. However, note that music is not only an auditory stimulus, but also a kinesthetic and visual one. Individuals with profound hearing loss often enjoy music, attend concerts and dance clubs, etc. (see Holmes, 2017, for discussion). In fact, some types of music need not involve sound at all; for example, Christine Sun Kim's *Face Opera II* (2013) involves a group of prelingually deaf performers "singing" via coordinated ASL facial expressions (Holmes, 2016). In terms of development, sign-language-exposed infants show a rhythmic "babbling" (e.g., Petitto, Holowka, Sergio, & Ostry, 2001) that may reflect developing temporal processing abilities involved in their later linguistic acquisition. (See Chapter X, this volume, for additional discussion of music and language in hearing impaired children.)

In sum, the existence of developmental deficits specific to music or language processing is sometimes taken as an argument against deep connections between these domains. However, the specific patterns of difficulties associated with these deficits, combined with the complexity of both music and language, paint a more nuanced picture.

19.5 OPEN QUESTIONS AND FUTURE DIRECTIONS

Music and language acquisition involve a feedback loop between our neurological development and experience-dependent exposure and training, but many details remain unresolved. How are milestones in music and language acquisition tied to our brain's maturation? What is happening in the brain that enables targets to be hit? Fuster (2013) views language as intrinsically bound up with prediction and planning and therefore heavily reliant on our prefrontal cortex, which is slow to mature. To what degree are gains in linguistic and musical comprehension and production bootstrapped to this gradual process? Yet young children are also generally more adept at language acquisition than adults, whose cognition otherwise outperforms them (Kuhl, 2010). Why is that so? What creates the developmental window which, once closed, puts language out of reach? With regard to experience, to what degree can factors such as cultural and socio-economic background influence the developmental timeline? Tooley et al. (2021) have found that higher economic status lengthens the time period for brain maturation. Studies have likewise documented the developmental costs of economic hardship: faced with acute poverty of stimulus, impoverished children suffer from language delays (Nelson et al., 2011). In an imaging study, Merz et al. (2020) found that economically disadvantaged children had key cortical regions involved in language processing that were underdeveloped relative to their peers. Likewise, in a longitudinal analysis of children from birth to four years of age, those from low-income households had reduced gray matter volume, especially in regions involved in executive function (Blair et al., 2016). The coordination of our inner lives with the world around us is intricate, subtle, and multi-faceted; understanding its general outlines, nuances, and range of possible progressions and outcomes remains an important frontier.

Cross-cultural studies would shed more light on this feedback loop, but these remain challenging. Western influence has intruded upon all but the most remote corners of the globe. As noted, researchers have spent time in isolated communities such as the Piraha and Tsimane' of the Amazon (Everett, 20054, McDermott, 2016) and the Mafa of Cameroon (Fritz, 2009), but in our interconnected world, it will only get harder to find subjects sequestered from Western practice. Ethnic communities and recent immigrants in urban areas offer a potential alternative (Jacoby et al., 2020), but the urgency of documenting those untouched by the West is accelerating.

Not only is going beyond Western subjects essential; so is avoiding Western biases in interpreting the results. Dowling has observed that infant's first songs display "a high variability of interval sizes and a drift in the 'tonal center'" (Dowling, 1984, 145). Likewise, as noted earlier, Gudmundsdottir (2020) has found that 3-year-olds are largely indifferent to Western music's priorities of stable pitch and fixed tonal center. Communist Russia's early music education program did not expect much in the way of pitch discrimination for children that age (Nikolsky et al., 2020). Nevertheless, Western investigators are often too quick to impose their pitch-centricity on the young. Nikolsky remarks:

> Children's singing is no more pitch-'defective' than their first attempts to speak are phoneme-'defective.' One cannot make mistakes in the production of pitch classes, if he/she is unaware of pitch classes. What appear as infant's 'poor singing' to a researcher, in fact, might not be 'poor' at all to an infant…" (Nikolsky et al., 2020, p. 180).

An extensive amount of evidence points to the partial entanglement of music and language even in adults (Patel, 2011). However, the jury is still out on whether these are the result of neural reuse, in which the same neural network can serve a variety of tasks, or because specialized networks lie so close together within the same region of the brain that they are hard to distinguish. Evidence has supported both views: for instance, interference tasks (Fennell et al., 2021) suggest that music and language compete for working memory. Meanwhile, researchers have teased out distinctive activation patterns within the same cortical regions (Abrams et al., 2011, Norman-Haignere et al., 2015; Ogg et al., 2019; Rogalsky et al., 2011; Sammler et al., 2013). Because music and language are so deeply entangled early in life, childhood development may offer further insights.

Meanwhile, animal studies promise to deepen our understanding of aural cognition. Human hearing begins to develop in utero, making it difficult to study, but the ear canals of mice are blocked during the first ten days of life. A recent study (Kline et al., 2021) suggests that the famous principle "neurons that fire together wire together" operates in the mouse brain in response to sound: neurons in the higher-order cortex responding to harmonically regular sounds become wired to each other when they share the same onset. The mouse thus learns through exposure to fuse the composite sounds that make up a mouse cry and other environmental stimuli. A follow-up experiment will broadcast white noise during the critical learning period to see if it interferes with the developmental progression: if so, it will indicate that a crucial feature of mouse audition is experience-dependent. Although it is not yet possible to know how closely mouse and human hearing are related, the results have implications for our own pitch, timbre, and phonemic perception, which likewise show a post-natal progression (Butler, Follard & Trainor, 2013).

No brain tissue survives from prehistoric times, and artifacts are scarce; our closest relatives in the animal kingdom communicate very differently from the way we do. As a result, research into the origins of music and language largely relies on inferences from the present. To that end, early child development may offer important clues. If ontogeny were to recapitulate phylogeny, the early entanglement and co-development of music and language would suggest that these two modes of communication may have largely been indistinguishable at first, before eventually separating (Tomlinson, 2015; Brown, 2017). We did not originate as specialists: whatever biological primitives exist must be open-ended enough to allow for Siberian timbral music and Gregorian chant (Patel and von Rueden, 2021), as well as the speech of the !Xoo, Rokata, Piraha, and Papau peoples (Evans and Levinson, 2009; Dabrowska, 2015; Tomasello and Ibbotson, 2016). Cognitively, the common denominators may have been our sociability, creativity, and ability to learn, and neurologically, our brains' early susceptibility to experience and on-going capacity for rewiring. Furthermore, the question of which came first—music or language—may be misdirected: they were partners, serving often overlapping functions—from social bonding to self-expression and nurturing the young, with different communities solving their communication needs in different ways. Certainly, as we learn more about the genetic and cultural factors influencing childhood development, we will gain a clearer view into our distant past (for further discussion, see Part IV).

19.6 SUMMARY

For both music and language, children begin as generalists and become specialists. This enables them to adapt to any environment in which they are raised. As they enter the world, they have rudimentary processing abilities and physical skills, but there is a lot they can't do: they can't match pitch or clap to a steady beat, and don't understand what words mean. It is through exposure and training that children's nascent cognitive abilities become attuned to their native culture. Whatever biological predispositions exist are rarely mandates: they can be overridden, and interact with each other in complex ways, yielding, for instance, different sensations of dissonance and different mappings for pitch perception.

The acquisition of music and language is a quintessential case of emergent modularity. The two are highly entangled in newborns – indeed, infants first attend to the musical features of their native language. Over time, thanks to the brain's quest for processing efficiency, musical and linguistic cognition becomes refined and separated. It is unlikely that we are born with well-formed music and language modules: indeed, linked developmental disorders point to communal rather than siloed functions. Rather, brains likely pull together the necessary neural resources based on the stimuli to which they are exposed: as a result, Western educated children learn to sing their pitch-centric music in tune, while the children of Siberia adopt timbral melodies in which there are no "wrong notes."

This process occurs throughout childhood at approximately equal rates for music and speech: it takes about the same amount of time for children to become conversant in their native language and musical culture, and about the same amount of time to become a Shakespearean actor or a professional violist.

From an evolutionary perspective, music and language do not have to be genetically prescribed; all that is required is brains that can invent them. Along with our social natures, humans have the "drastic increase in cortical connective complexity" (Fuster, 2013, 176) that this requires. Music and language operate the way they do because they fit our brains. They are means we use to explore what our minds are capable of: their structures, nuances, consistencies, and surprises reflections of our richly networked and plastic neural architecture, brought into the world largely unformed.

Parents around the world are eager to get children talking. But, ironically, it may be that we only have language and music because we are born too soon: both need to be learned and require ample time to do so; both are subject to extraordinary variability. We tend to marvel at how precociously our children learn to talk and make music; instead, we perhaps should be thankful it happens so *late*. Our unfledged entry into the world may give us a unique opportunity to be less programmed and more flexible, enabling us to invent, cultivate, and teach the words and songs that we speak and sing throughout our lifetimes.

REFERENCES

Abrams, D.A., Bhatara, A., Ryali, S., Balaban, E., Levitin, D.J., Menon, V. (2011). Decoding temporal structure in music and speech relies on shared brain resources but elicits different fine-scale spatial patterns. *Cerebral Cortex 21,* 1507–1518.

Abramson, A. S. (2015). Tonal experiments with whispered Thai. In *Papers in linguistics and phonetics to the memory of Pierre Delattre* (pp. 31-44). De Gruyter Mouton.

Agawu, K. (2016). *The African Imagination in Music*. New York: Oxford University Press.

Alekseyev, EY. (1976) Problems in the genesis of musical mode (on the example of Yakut folksong):analysis (Проблемы формирования лада (на материале якутской народной песни): исследование). Muzyka (Музыка), Moscow.

Armbrüster, L., Mende, W., Gelbrich, G., Wermke, P., Götz, R., & Wermke, K. (2020). Musical intervals in infants' spontaneous crying over the first 4 months of life. *Folia Phoniatrica et Logopaedica*, 1-12.

Arom, S. (1991). *African Polyphony and Polyrhythm*. Cambridge: Cambridge University Press.

Asano, R., Boeckx, C., & Seifert, U. (2021). Hierarchical control as a shared neurocognitive mechanism for language and music. *Cognition, 216,* 104847.

Ayotte, J., Peretz, I., & Hyde, K. (2002). Congenital amusia: A group study of adults afflicted with a music-specific disorder. *Brain, 125*(2), 238-251.

Bedoin, N., Brisseau, L., Molinier, P., Roch, D., & Tillmann, B. (2016). Temporally regular musical primes facilitate subsequent syntax processing in children with specific language impairment. *Frontiers in Neuroscience 10*:245.

Beliayeva-Ekzempliarskaya, S. (1925). "Musical experience in preschool age [Музыкальное переживание в дошкольном возрасте]," in Collection of Works of the Physiolo-psychological Department [Сборник работ физиолого-психологической секции], ed S. Beliayeva-Ekzempliarskaya (Moscow: The State Institute of Musical Science [Гос ин-та музыкальной науки.]), 3–29.

Benetti, L., & Costa-Giomi, E. (2020). Infant vocal imitation of music. *Journal of Research in Music Education*, *67*(4), 381-398.

Bergen, B.K., 2001. Nativization processes in L1 Esperanto. *Journal of Child Language 28,* 575–595.

Best, C. T. (2019). The diversity of tone languages and the roles of pitch variation in non-tone languages: Considerations for tone perception research. *Frontiers in Psychology, 10:364*.

Best, C. T., & Avery, R. A. (1999). Left-hemisphere advantage for click consonants is determined by linguistic significance and experience. *Psychological Science*, *10*(1), 65-70.

Bidelman, G. M., Hutka, S., & Moreno, S. (2013). Tone language speakers and musicians share enhanced perceptual and cognitive abilities for musical pitch: evidence for bidirectionality between the domains of language and music. *PloS one*, *8*(4), e60676.

Blair, C., & Raver, C. C. (2016). Poverty, Stress, and Brain Development: New Directions for Prevention and Intervention. *Academic Pediatrics, 16*(3), S30-S36.

Bonacina, S., Krizman, J., White-Schwoch, T., Nicol, T., & Kraus, N. (2020). Distinct rhythmic abilities align with phonological awareness and rapid naming in school-age children. *Cognitive Processing*, *21*(4), 575-581.

Bosch, L., & Sebastián-Gallés, N. (1997). Native-language recognition abilities in 4-month-old infants from monolingual and bilingual environments. *Cognition*, *65*(1), 33-69.

Brandt, A., Gebrian, M., & Slevc, L. R. (2012). Music and early language acquisition. *Frontiers in Psychology*.

Brown, S. (2017). A joint prosodic origin of language and music. *Frontiers in Psychology, 8*.

Butler, B. E., Folland, N. A., & Trainor, L. J. (2013). Development of pitch processing: Infants' discrimination of iterated rippled noise stimuli with unresolved spectral content. *Hearing Research*, *304*, 1–6.

Carreiras, M., Lopez, J., Rivero, F., & Corina, D. (2005). Neural processing of a whistled language. *Nature*, *433*(7021), 31-32.

Čėponienė, R., Kushnerenko, E., Fellman, V., Renlund, M., Suominen, K., & Näätänen, R. (2002). Event-related potential features indexing central auditory discrimination by newborns. *Cognitive Brain Research*, *13*(1), 101-113.

Chang, S.-E., Chow, H. M., Wieland, E. A., & McAuley, J. D. (2016). Relation between functional connectivity and rhythm discrimination in children who do and do not stutter. *NeuroImage: Clinical, 12,* 442–450.

Cheng, Y. Y., & Lee, C. Y. (2018). The development of mismatch responses to Mandarin lexical tone in 12-to 24-month-old infants. *Frontiers in Psychology*, *9*, 448.

Cheng, Y., Lee, S. Y., Chen, H. Y., Wang, P. Y., & Decety, J. (2012). Voice and emotion processing in the human neonatal brain. *Journal of Cognitive Neuroscience*, *24*(6), 1411-1419.

Cheour, M., Čėponienė, R., Lehtokoski, A., Luuk, A., Allik, J., Alho, K., and Näätänen, R. (1998). Development of language-specific phoneme representations in the infant brain. *Nature Neuroscience 1,* 351–353.

Cheour, M., Čėponienė, R., Leppänen, P., Alho, K., Kujala, T., Renlund, M., & Näätänen, R. (2002). The auditory sensory memory trace decays rapidly in newborns. *Scandinavian Journal of Psychology*, *43*(1), 33-39.

Chomsky, N. (1965). *Aspects of the Theory of Syntax*. Cambridge, Mass: MIT Press.

Clément, S., Planchou, C., Béland, R., Motte, J., & Samson, S. (2015). Singing abilities in children with Specific Language Impairment (SLI). *Frontiers in Psychology 6*.

Cohen, N. J., Vallance, D. D., Barwick, M., Im, N., Menna, R., Horodezky, N. B., & Isaacson, L. (2000). The interface between ADHD and language impairment: An examination of language, achievement, and cognitive processing. *The Journal of Child Psychology and Psychiatry and Allied Disciplines, 41*(3), 353-362.

Colling, L. J., Noble, H. L., & Goswami, U. (2017). Neural entrainment and sensorimotor synchronization to the beat in children with developmental dyslexia: An EEG study. *Frontiers in Neuroscience, 11*, 360.

Corrigall, K. A., and Trainor, L. J. (2009). Effects of musical training on key and harmony perception. *Annals of the New York Academy of Sciences, 1169,* 164–168.

Costa-Giomi, E. (2003). Young children's harmonic perception. *Annals of the New York Academy of Sciences,* 999, 477–484.

Couvignou, M., Peretz, I., & Ramus, F. (2019). Comorbidity and cognitive overlap between developmental dyslexia and congenital amusia. *Cognitive Neuropsychology, 36*(1-2), 1-17.

Creel, S. C. (2016). Ups and downs in auditory development: Preschoolers' sensitivity to pitch contour and timbre. *Cognitive Science*, *40*(2), 373-403.

Creel, S. C., Weng, M., Fu, G., Heyman, G. D., & Lee, K. (2018). Speaking a tone language enhances musical pitch perception in 3–5-year-olds. *Developmental Science*, *21*(1), e12503.

Cross, I. (2003). Music as a Biocultural Phenomenon. *Annals of the New York Academy of Sciences, 999,* 106–111.

Cross, I. (2008). Musicality and the human capacity for culture. In *Musicae Scientiae, 12*, 147–167.

Cross, I. (2009). The evolutionary nature of musical meaning. *Musicae Scientiae*, *13*(2), 179–200.

Cumming, R., Wilson, A., Leong, V., Colling, L. J., & Goswami, U. (2015). Awareness of rhythm patterns in speech and music in children with specific language impairments. *Frontiers in Human Neuroscience 9*.

Dąbrowska, E. (2015). What exactly is Universal Grammar, and has anyone seen it? *Frontiers in Psychology* 6:852.

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.

de l'Etoile, S. K. (2006). Infant-directed singing: A theory for clinical intervention. *Music Therapy Perspectives*, *24*(1), 22-29.

Dean, B. (2020). Spontaneous singing in early childhood: An examination of young children's singing at home. *Research Studies in Music Education*, 1321103X20924139.

Dehaene-Lambertz, G. (2000). Cerebral specialization for speech and non- speech stimuli in infants. *J. Cogn. Neurosci*. 12, 449–460.

Dehaene-Lambertz, G., Dehaene, S., and Hertz-Pannier, L. (2002). Functional neuroimaging of speech perception in infants. *Science* 298, 2013–2015.

Dehaene, S., Dupoux, E., Mehler, J., Cohen, L., Paulesu, E., Perani, D., ... & Le Bihan, D. (1997). Anatomical variability in the cortical representation of first and second language. *Neuroreport*, *8*(17), 3809-3815.

Demany, L., McKenzie, B., & Vurpillot, E. (1977). Rhythm perception in early infancy. *Nature*, *266*(5604), 718-719.

Demorest, S. M., & Pfordresher, P. Q. (2015). Singing accuracy development from K-adult: A comparative study. *Music Perception: An Interdisciplinary Journal*, *32*(3), 293-302.

Deutsch, D., Lapidis, R., & Henthorn, T. (2008). The speech-to-song illusion. *Journal of the Acoustical Society of America*, *124*(4), 2471.

Dittinger, E., Barbaroux,M., D'Imperio, M, Jäncke, L. Elmer, S., Besson, M. (2016). Professional Music Training and Novel Word Learning: From Faster Semantic Encoding to Longer-lasting Word Representations. *J Cogn Neurosci*. 28 (10): 1584–1602.

Dolscheid, S., Hunnius, S., Casasanto, D., & Majid, A. (2014). Prelinguistic infants are sensitive to space-pitch associations found across cultures. *Psychological Science*, *25*(6), 1256-1261.

Dolscheid, S., Shayan, S., Majid, A., & Casasanto, D. (2013). The Thickness of Musical Pitch: Psychophysical Evidence for Linguistic Relativity. *Psychological Science, 24(5), 613-621.*

Dowling, W. J. (1984). Development of Musical Schemata in Children's Spontaneous Singing. *Advances in Psychology*, *19*(C), 145–163.

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.

Eagleman, D. (2020). *Livewired: The Inside Story of the Ever-Changing Brain*. New York: Pantheon Books.

Eitan, Z., & 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*(3), 405–422.

Evans, N., & Levinson, S. C. (2009, October). The myth of language universals: Language diversity and its importance for cognitive science. *Behavioral and Brain Sciences*.

Everett, D. L. (2005). Cultural constraints on grammar and cognition in Pirahã: Another look at the design features of human language. *Current Anthropology*.

Fava, E., Hull, R., Baumbauer, K., & Bortfeld, H. (2014). Hemodynamic responses to speech and music in preverbal infants. *Child Neuropsychology*, *20*(4), 430-448.

Fennell, A. M., Bugos, J. A., Payne, B. R., & Schotter, E. R. (2021). Music is similar to language in terms of working memory interference. *Psychonomic Bulletin and Review*, *28*(2), 512–525.

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.

Fitzroy, A. B., & Sanders, L. D. (2013). Musical expertise modulates early processing of syntactic violations in language. *Frontiers in Psychology*, *3*, 603.

Flaugnacco, E., Lopez, L., Terribili, C., Montico, M., Zoia, S., & Schön, D. (2015). Music training increases phonological awareness and reading skills in developmental dyslexia: A randomized control trial. *PLoS ONE 10*(9), e0138715.

Flowers, P. J., & Dunne-Sousa, D. (1990). Pitch-pattern accuracy, tonality, and vocal range in preschool children's singing. *Journal of Research in Music Education*, *38*(2), 102-114.

Frey, A., François, C., Chobert, J., Velay, J. L., Habib, M., & Besson, M. (2019). Music training positively influences the preattentive perception of voice onset time in children with dyslexia: A longitudinal study. *Brain Sciences, 9*(4), 91.

Friederici, A. D. (2006). The neural basis of language development and its impairment. *Neuron* 52, 941–952.

Fritz, T. (2009). *Emotion investigated with music of variable valence-neurophysiology and cultural influence* (Doctoral dissertation, Max Planck Institute for Human Cognitive and Brain Sciences Leipzig).

Fuster, J.M. (2013). *The Neuroscience of Freedom and Creativity: Our Predictive Brain*. Cambridge, England: Cambridge University Press.

Geretsegger, M., Elefant, C., Mössler, K. A., & Gold, C. (2014). Music therapy for people with autism spectrum disorder. *Cochrane Database of Systematic Reviews 17*(6), CD004381.

Gerken, L., Landau, B., and Remez, R. (1990). Function morphemes in young children's speech perception and production. *Dev.Psychol*. 26, 204–216.

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.

Gordon, R. L., Shivers, C. M., Wieland, E. A., Kotz, S. A., Yoder, P. J., & Devin McAuley, J. (2015). Musical rhythm discrimination explains individual differences in grammar skills in children. *Developmental science*, *18*(4), 635-644.

Götz, A., Yeung, H. H., Krasotkina, A., Schwarzer, G., & Höhle, B. (2018). Perceptual reorganization of lexical tones: effects of age and experimental procedure. *Frontiers in psychology*, *9*, 477.

Gudmundsdottir, H. R. (2020). Revisiting singing proficiency in three-year-olds. *Psychology of Music*, *48*(2), 283-296.

Háden, G. P., Németh, R., Török, M., & Winkler, I. (2015). Predictive processing of pitch trends in newborn infants. *Brain research*, *1626*, 14-20.

Háden, G. P., Stefanics, G., Vestergaard, M. D., Denham, S. L., Sziller, I., & Winkler, I. (2009). Timbre-independent extraction of pitch in newborn infants. *Psychophysiology*, *46*(1), 69-74.

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.

Halpern, A. R., & Müllensiefen, D. (2008). Effects of timbre and tempo change on memory for music. *Quarterly Journal of Experimental Psychology*, *61*(9), 1371-1384.

Hämäläinen, J. A., Salminen, H. K., & Leppänen, P. H. (2013). Basic auditory processing deficits in dyslexia: Systematic review of the behavioral and event-related potential/field evidence. *Journal of Learning Disabilities 46*(5), 413–427.

Hannon, E. E., & Trainor, L. J. (2007). Music acquisition: effects of enculturation and formal training on development. *Trends in cognitive sciences*, *11*(11), 466-472.

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. Deliege and J. Sloboda (Oxford: Oxford University Press), 145–170.

Heaton, P. (2009). Assessing musical skills in autistic children who are not savants. *Philosophical Transactions of the Royal Society B: Biological Sciences, 364*(1522), 1443-1447.

Heaton, P., Allen, R., Williams, K., Cummins, O., & Happé, F. (2008). Do social and cognitive deficits curtail musical understanding? Evidence from autism and Down syndrome. British *Journal of Developmental Psychology, 26*(2), 171-182.

Henrich, J. (2020).The Weirdest People in the World: How the West Became Psychologically Peculiar and Particularly Prosperous. New York, NY: Farrar, Straus and Giroux.

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). Most people are not WEIRD. *Nature*, *466*(7302), 29-29.

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.

Holmes, J. A. (2016). Singing beyond hearing. *Journal of the American Musicological Society, 69*(2), 542-547.

Holmes, J. A. (2017). Expert listening beyond the limits of hearing: Music and deafness. *Journal of the American Musicological Society*, *70*(1), 171-220.

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., & Goswami, U. (2011). Music, rhythm, rise time perception and developmental dyslexia: perception of musical meter predicts reading and phonology. *Cortex, 47*(6), 674-689.

Hutchins, S., Gosselin, N. & Peretz, I. Identification of changes along a continuum of speech intonation is impaired in congenital amusia. *Front Psychol 1,* 236 (2010).

Jacoby, N., Margulis, E. H., Clayton, M., Hannon, E., Honing, H., Iversen, J., … Wald-Fuhrmann, M. (2020). Cross-cultural work in music cognition: Challenges, insights, and recommendations. *Music Perception*, *37*(3), 185–195.

Jacoby, N., Undurraga, E. A., McPherson, M. J., Valdés, J., Ossandón, T., & McDermott, J. H. (2019). Universal and non-universal features of musical pitch perception revealed by singing. *Current Biology*, *29*(19), 3229-3243.

Jasmin, K., Dick, F., Stewart, L., & Tierney, A. T. (2020). Altered functional connectivity during speech perception in congenital amusia. *ELife, 9,* e53539.

Jentschke, S., & Koelsch, S. (2009). Musical training modulates the development of syntax processing in children. *NeuroImage*, *47*(2), 735–744.

Jentschke, S., Koelsch, S., Sallat, S., & Friederici, A. D. (2008). Children with specific language impairment also show impairment of music-syntactic processing. *Journal of Cognitive Neuroscience 20*(11), 1940–1951.

Jones, J. L., Lucker, J., Zalewski, C., Brewer, C. & Drayna, D. Phonological processing in adults with deficits in musical pitch recognition. *J Commun Dis 42,* 226–234 (2009).

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.

Jusczyk, P. W., Pisoni, D. B., & Mullennix, J. (1992). Some consequences of stimulus variability on speech processing by 2-month-old infants. *Cognition*, *43*(3), 253-291.

Karmiloff-Smith, A. (1992). *Beyond Modularity: A Developmental Perspective on Cognitive Science*. MIT Press.

Kline, A.M., Aponte, D.A., Tsukano, H. *et al*. (2021). Inhibitory gating of coincidence-dependent sensory binding in secondary auditory cortex. *Nat Commun* 12, 4610.

Koelsch, S. (2012). *Brain and music*. Chichester, UK: Wiley Blackwell.

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., ... & Fellman, V. (2010). Hemodynamic responses to speech and music in newborn infants. *Human Brain Mapping*, *31*(4), 595-603.

Kreiner, H., & Eviatar, Z. (2014). The missing link in the embodiment of syntax: Prosody. *Brain and Language*, *137*, 91-102.

Kreutzer, N. (2001). Acquisition among from rural Shona-speaking Zimbabwean children from birth to 7 years. *J. Res. Music Educ.* 49, 198–211.

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.

Ladányi, E., Persici, V., Fiveash, A., Tillmann, B., & Gordon, R. L. (2020). Is atypical rhythm a risk factor for developmental speech and language disorders?. *Wiley Interdisciplinary Reviews: Cognitive Science, 11*(5), e1528.

Lau, B., Oxenham, A., & Werner, L. (2020). Infants can outperform adults in pitch and timbre perception. https://psyarxiv.com/vbws6/

Laycock, D., (1979). Multilingualism: linguistic boundaries and unsolved problems in Papua New Guinea. In:Wurm, S.A. (Ed.), New Guinea and Neighboring Areas: A Sociolinguistic Laboratory. Mouton, The Hague, pp. 81–99.

Lee, C. Y., & Cheng, Y. Y. (2020). Neurophysiological Studies of Mandarin Lexical Tone Acquisition in Early Childhood. *Speech Perception, Production and Acquisition*, 101-116.

Levin, T.C., & Suzukei, V. (2006). *Where rivers and mountains sing: Sound, music, and nomadism in Tuva and beyond*. Indiana University Press.

Liu, F., Jiang, C., Wang, B., Xu, Y., & Patel, A. D. (2015). A music perception disorder (congenital amusia) influences speech comprehension. *Neuropsychologia 66,* 111–118.

Liu, F., Patel, A. D., Fourcin, A., & Stewart, L. (2010). Intonation processing in congenital amusia: Discrimination, identification and imitation. *Brain 133*(6), 1682–1693.

Liu, F., Yin, Y., Chan, A. H., Yip, V., & Wong, P. C. (2021). Individuals with congenital amusia do not show context-dependent perception of tonal categories. *Brain and Language, 215,* 104908.

Liu, L., & Kager, R. (2014). Perception of tones by infants learning a non-tone language. *Cognition*, *133*(2), 385-394.

Locke, J. L., Bekken, K. E., Mcminnlarson, L., & Wein, D. (1995). Emergent control of manual and vocal-motor activity in relation to the development of speech. *Brain and language*, *51*(3), 498-508.

Lolli, S., Lewenstein, A. D., Basurto, J., Winnik, S., & Loui, P. (2015). Sound frequency affects speech emotion perception: Results from congenital amusia. *Frontiers in Psychology, 6,* 1340.

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.

Magne, C., Schön, D., & Besson, M. (2006). Musician Children Detect Pitch Violations in Both Music and Language Better than Nonmusician Children: Behavioral and Electrophysiological Approaches. *Journal of Cognitive Neuroscience*, *18*(2), 199–211.

Mahmoudzadeh, M., Wallois, F., Kongolo, G., Goudjil, S., & Dehaene-Lambertz, G. (2017). Functional maps at the onset of auditory inputs in very early preterm human neonates. *Cerebral Cortex*, *27*(4), 2500-2512.

Mampe, B., Friederici, A. D., Christophe, A., & Wermke, K. (2009). Newborns' cry melody is shaped by their native language. *Current biology*, *19*(23), 1994-1997.

Masataka, N. (1999). Preference for infant-directed singing in 2-day-old hearing infants of deaf parents. *Developmental psychology*, *35*(4), 1001.

Mattock, K., & Burnham, D. (2006). Chinese and English infants' tone perception: Evidence for perceptual reorganization. *Infancy*, *10*(3), 241-265.

McAdams, S., and Bertoncini, J. (1997). Organization and discrimination of repeating sound sequences by new- born infants. *J. Acoust. Soc. Am*. 102, 2945–2953.

McDermott, J. H., Schultz, A. F., Undurraga, E. A., & Godoy, R. A. (2016). Indifference to dissonance in native Amazonians reveals cultural variation in music perception. *Nature*, *535*(7613), 547–550.

McMullen, E., and Saffran, J. R. (2004). Music and language: a developmental comparison. *Music Percept*. 21, 289–311.

McPhee, Colin (1966). *Music in Bali: A Study in Form and Instrumental Organization in Balinese Orchestral Music*. New Haven: Yale University Press.

Mehr, S. A., Singh, M., Knox, D., Ketter, D. M., Pickens-Jones, D., Atwood, S., … Glowacki, L. (2019). Universality and diversity in human song. *Science*, *366*(6468).

Merz, E. C., Maskus, E. A., Melvin, S. A., He, X., & Noble, K. G. (2020). Socioeconomic Disparities in Language Input Are Associated With Children's Language-Related Brain Structure and Reading Skills. *Child Development*, *91*(3), 846–860.

Milovanov, R., Huotilainen, M., Välimäki, V., Esquef, P. A. A., & Tervaniemi, M. (2008). Musical aptitude and second language pronunciation skills in school-aged children: Neural and behavioral evidence. *Brain Research*, *1194*, 81–89.

Minagawa-Kawai, Y., Cristià, A., Vendelin, I., Cabrol, D., & Dupoux, E. (2011). Assessing signal-driven mechanisms in neonates: brain responses to temporally and spectrally different sounds. *Frontiers in Psychology*, *2*, 135.

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.

Moreau, P., Jolicœur, P., & Peretz, I. (2013). Pitch discrimination without awareness in congenital amusia: evidence from event-related potentials. *Brain and Cognition, 81*(3), 337-344.

Möttönen, R., Calvert, G. A., Jääskeläinen, I. P., Matthews, P. M., Thesen, T., Tuomainen, J., & Sams, M. (2006). Perceiving identical sounds as speech or non-speech modulates activity in the left posterior superior temporal sulcus. *Neuroimage*, *30*(2), 563-569.

Mrazek, Jan (1999). "Javanese Wayang Kulit in the Times of Comedy: Clown Scenes, Innovation, and the Performance's Being in the Present World." Pt. 1. *Indonesia* 68, 33 - 128.

Musacchia, G., Sams, M., Skoe, E., & Kraus, N. (2007). Musicians have enhanced subcortical auditory and audiovisual processing of speech and music. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(40), 15894–15898.

Nakata, T., & Trehub, S. E. (2011). Expressive timing and dynamics in infant-directed and non-infant-directed singing. *Psychomusicology: Music, Mind and Brain*, *21*(1-2), 45.

Nan, Y., Sun, Y. & Peretz, I. Congenital amusia in speakers of a tone language: association with lexical tone agnosia. *Brain 133,* 2635–2642 (2010).

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.

Nelson, K. E., Welsh, J. A., Trup, E. M. V., & Greenberg, M. T. (2011). Language delays of impoverished preschool children in relation to early academic and emotion recognition skills. *First Language*, *31*(2), 164–194.

Nespor, M., Shukla, M., van deVijver, R., Avesani, C., Schraudolf, H., and Donati, C. (2008). Different phrasal prominence realization in VO and OV languages. *Lingue e Linguaggio*. 7, 1–28.

Nikolsky, A. (2020). "Emergence of the Distinction between Verbal and Musical in Early Childhood." In N. Masataka (Ed.) The Origins of Language Revisited: Differentiation from Music and the Emergence of Neurodiversity and Autism (139-219). Singapore: Springer.

Nikolsky, A., Alekseyev, E., Alekseev, I., & Dyakonova, V. (2020). The Overlooked Tradition of "Personal Music" and Its Place in the Evolution of Music. *Frontiers in Psychology*, *10*.

Norman-Haignere, S. V., Albouy, P., Caclin, A., McDermott, J. H., Kanwisher, N. G., & Tillmann, B. (2016). Pitch-responsive cortical regions in congenital amusia. *Journal of Neuroscience, 36*(10), 2986-2994.

Norman-Haignere, S. V., Feather, J., Boebinger, D., Brunner, P., Ritaccio, A., McDermott, J. H., … Kanwisher, N. (2022). A neural population selective for song in human auditory cortex. *Current Biology*, *32*(6), 1454–1455.

Norman-Haignere, S., Kanwisher, N. G., & McDermott, J. H. (2015). Distinct cortical pathways for music and speech revealed by hypothesis-free voxel decomposition. *Neuron, 88*(6), 1281-1296.

Nuñez, 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.

Ogg, M., Moraczewski, D., Kuchinsky, S. E., & Slevc, L. R. (2019). Separable neural representations of sound sources: Speaker identity and musical timbre. *Neuroimage, 191*, 116-126.

Ojamaa, T. (2005). Throat rasping: problems of visualization. *The world of music*, 55-69.

Olsho, L. W., Schoon, C., Sakai, R., Turpin, R., & Sperduto, V. (1982). Auditory frequency discrimination in infancy. *Developmental Psychology 18*(5), 721-726.

Overy, K., Nicolson, R. I., Fawcett, A. J., & Clarke, E. F. (2003). Dyslexia and music: Measuring musical timing skills. *Dyslexia, 9*(1), 18–36.

Paillereau, N., Podlipský, V. J., Smolík, F., Šimáčková, Š., & Chládková, K. (2021). The development of infants' sensitivity to native versus non-native rhythm. *Infancy*, *26*(3), 423-441.

Parkinson, C., Kohler, P. J., Sievers, B., & Wheatley, T. (2012). Associations between auditory pitch and visual elevation do not depend on language: Evidence from a remote population. *Perception*, *41*(7), 854–861.

Patel, A. D. (2003). Language, music, syntax and the brain. *Nature Neuroscience, 6*(7), 674-681.

Patel, A. D. (2010). *Music, language, and the brain*. Oxford University Press.

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

Patel, A. D., & Von Rueden, C. (2021). Where they sing solo: Accounting for cross-cultural variation in collective music-making in theories of music evolution. *Behavioral and Brain Sciences, 44,* 77–79.

Perani, D. (2012). Functional and structural connectivity for language and music processing at birth. *Rendiconti Lincei*, *23*(3), 305-314.

Perani, D., Dehaene, S., Grassi, F., Cohen, L., Cappa, S. F., Dupoux, E., ... & Mehler, J. (1996). Brain processing of native and foreign languages. *NeuroReport-International Journal for Rapid Communications of Research in Neuroscience*, *7*(15), 2439-2444.

Perani, D., Paulesu, E., Galles, N. S., Dupoux, E., Dehaene, S., Bettinardi, V., ... & Mehler, J. (1998). The bilingual brain. Proficiency and age of acquisition of the second language. *Brain: a journal of neurology*, *121*(10), 1841-1852.

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.

Peretz, I. (2016). Neurobiology of congenital amusia. *Trends in Cognitive Sciences, 20*(11), 857-867.

Peretz, I., & Hyde, K. L. (2003). What is specific to music processing? Insights from congenital amusia. *Trends in Cognitive Sciences 7*(8), 362–367.

Petitto, L. A., Holowka, S., Sergio, L. E., & Ostry, D. (2001). Language rhythms in baby hand movements. *Nature 413*(6851), 35–36.

Piazza, E. A., Iordan, M. C., & Lew-Williams, C. (2017). Mothers consistently alter their unique vocal fingerprints when communicating with infants. *Current Biology*, *27*(20), 3162-3167.

Politimou, N., Dalla Bella, S., Farrugia, N., & Franco, F. (2019). Born to speak and sing: Musical predictors of language development in pre-schoolers. *Frontiers in Psychology*, *10*, 948.

Politimou, N., Douglass-Kirk, P., Pearce, M., Stewart, L., & Franco, F. (2021). Melodic expectations in 5-and 6-year-old children. *Journal of Experimental Child Psychology*, *203*, 105020.

Polka, L., and Werker, J. F. (1994). Developmental changes in perception of nonnative vowel contrasts. *J. Exp. Psychol. Hum. Percept. Perform.* 20, 421–435.

Pralus, A., Fornoni, L., Bouet, R., Gomot, M., Bhatara, A., Tillmann, B., & Caclin, A. (2019). Emotional prosody in congenital amusia: impaired and spared processes. *Neuropsychologia, 134*, 107234.

Pride, L. (1963). Chatino Tonal Structure. *Anthropological Linguistics* 5(2), 19-28.

Prochnow, A., Erlandsson, S., Hesse, V., & Wermke, K. (2019). Does a 'musical' mother tongue influence cry melodies? A comparative study of Swedish and German newborns. *Musicae Scientiae*, *23*(2), 143-156.

Przybylski, L., Bedoin, N., Krifi-Papoz, S., Herbillon, V., Roch, D., Léculier, L., & Tillmann, B. (2013). Rhythmic auditory stimulation influences syntactic processing in children with developmental language disorders. *Neuropsychology 27*(1), 121–131.

Puyjarinet, F., Bégel, V., Lopez, R., Dellacherie, D., & Dalla Bella, S. (2017). Children and adults with Attention-Deficit/Hyperactivity Disorder cannot move to the beat. *Scientific Reports, 7*(1), 1-11.

Radvansky, G. A., & Potter, J. K. (2000). Source cuing: Memory for melodies. *Memory & Cognition*, *28*(5), 693-699.

Richards, S., & Goswami, U. (2019). Impaired recognition of metrical and syntactic boundaries in children with Developmental Language Disorders. *Brain Sciences, 9*(2), 33.

Rivera-Gaxiola, M., Klarman, L., Garcia-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.

Robinson, S. (2006). The phoneme inventory of the Aita dialect of Rotokas. *Oceanic Linguistics*, *45*(1), 206–209.

Rogalsky, C., Rong F., Saberi K., and Hickok G. (2011). Functional anatomy of language and music perception: temporal and structural factors investigated using functional magnetic resonance imaging. J. Neurosci. 31, 3843–3852.

Rosen, S. (1992). Temporal information in speech: acoustic, auditory and linguistic aspects. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 336, 367–373.

Rost, G. C., & McMurray, B. (2009). Speaker variability augments phonological processing in early word learning. *Developmental Science*, *12*(2), 339-349.

Saint-Georges, C., Chetouani, M., Cassel, R., Apicella, F., Mahdhaoui, A., Muratori, F., ... & Cohen, D. (2013). Motherese in interaction: at the cross-road of emotion and cognition? (A systematic review). *PloS one*, *8*(10), e78103.

Sammler, D., Koelsch, S., Ball, T., Brandt, A., Grigutsch, M., Huppertz, H. J., … Schulze-Bonhage, A. (2013). Co-localizing linguistic and musical syntax with intracranial EEG. *NeuroImage*, *64*(1), 134–146.

Savage, P. E., Loui, P., Tarr, B., Schachner, A., Glowacki, L., Mithen, S. and Fitch, W. T. (2021) Music as a coevolved system for social bonding. *Behavioral and Brain Sciences*.

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.

Seskauskaité, D. (2004). "Sutartinès and Balkan Diaphonic Songs." *International Review of the Aesthetics and Sociology of Music,* 35(1), 71-92.

Shepard, R. (1980). Multidimensional scaling, tree-fitting, and clustering. *Science* 210, 390–398.

Skeide, M. A., & Friederici, A. D. (2016). The ontogeny of the cortical language network. *Nature Reviews Neuroscience*, *17*(5), 323-332.

Slevc, L. R., & Okada, B. M. (2015). Processing structure in language and music: A case for shared reliance on cognitive control. *Psychonomic Bulletin & Review, 22*(3), 637-652.

Slevc, L. R., Rosenberg, J. C., & Patel, A. D. (2009). Making psycholinguistics musical: Self-paced reading time evidence for shared processing of linguistic and musical syntax. *Psychonomic Bulletin & Review, 16*(2), 374-381.

Speer, J. R., and Meeks, P. U. (1985). School children's perception of pitch in music. *Psychomusicology* 5, 49–56.

Stern, D. N., Spieker, S., & MacKain, K. (1982). Intonation contours as signals in maternal speech to prelinguistic infants. *Developmental Psychology*, *18*(5), 727.

Stone R. (1982). Let the inside be sweet: the interpretation of music event among the Kpelle of Liberia. Bloomington, IN: Indiana University Press.

Suppanen, E., Huotilainen, M., & Ylinen, S. (2019). Rhythmic structure facilitates learning from auditory input in newborn infants. *Infant Behavior and Development*, *57*, 101346.

Tallal, P., and Piercy, M. (1973). Defects of non-verbal auditory perception in children with developmental aphasia. *Nature* 241, 468–469.

Telkemeyer, S., Rossi, S., Koch, S. P., Nierhaus, T., Steinbrink, J., Poeppel, D., Obrig, H., and Wartenburger, I. (2009). Sensitivity of newborn auditory cortex to the temporal structure of sounds. *J. Neurosci*. 29, 14726–14733.

Thompson, W. F., Marin, M. M., & Stewart, L. (2012). Reduced sensitivity to emotional prosody in congenital amusia rekindles the musical protolanguage hypothesis. *Proceedings of the National Academy of Sciences 109*(46), 19027–19032.

Thomson, J. M., & Goswami, U. (2008). Rhythmic processing in children with developmental dyslexia: auditory and motor rhythms link to reading and spelling. *Journal of Physiology-Paris, 102*(1-3), 120-129.

Thurston, W.R. (1987). Processes of change in the languages of north-western New Britain. In: *Pacific Linguistics* B99, The Australian National University, Canberra.

Tierney, A. L., & Nelson, C. A. (2009). Brain Development and the Role of Experience in the Early Years. *Zero to Three*, *30*(2), 9–13.

Tomasello, P., & Ibbotson, M. (2016). Evidence Rebuts Chomsky's Theory of Language Learning. *Scientific American*, 1–17.

Tomlinson, G.(2015). *A Million Years of Music: The Emergence of Human Modernity*. New York: Zone Books.

Tooley, U. A., Bassett, D. S., & Mackey, A. P. (2021). Environmental influences on the pace of brain development. *Nature Reviews Neuroscience, 22,* 372–384.

Tracey, H. (1958). Towards an Assessment of African Scales. *African Music, 2*(1), 15-20.

Trainor, L. J. (1996). Infant preferences for infant-directed versus noninfant-directed playsongs and lullabies. *Infant Behavior and Development*, *19*(1), 83-92.

Trainor, L. J., and Trehub, S. E. (1994). Key membership and implied harmony in Western tonal music: developmental perspectives. *Percept. Psychophys*. 56, 125–132.

Trainor, L. J., Austin, C. M., and Desjardins, N. (2000). Is infant-directed speech a result of the vocal expression of emotion? *Psychological Science, 11,* 188–195.

Trainor, L. J., Clark, E. D., Huntley, A., & Adams, B. A. (1997). The acoustic basis of preferences for infant-directed singing. *Infant Behavior and Development*, *20*(3), 383-396.

Trainor, L. J., Wu, L., & Tsang, C. D. (2004). Long-term memory for music: Infants remember tempo and timbre. *Developmental Science*, *7*(3), 289-296.

Trehub, S. E., Cohen, A., Thorpe, L., and Morrongiello, B. (1986). Development of the perception of musical relations: semitone and diatonic structure. *Journal of Experimental Psychology: Human Perception and Performance, 12,* 295–301.

Trehub, S. E., Endman, M. W., & Thorpe, L. A. (1990). Infants' perception of timbre: Classification of complex tones by spectral structure. *Journal of Experimental Child Psychology*, *49*(2), 300-313.

Trehub, S.E., Becker, J, Morley, I. (2015). Cross-cultural perspectives on music and musicality. *Philosophical Transactions of the Royal Society B, 370:*20140096.

Tsao, F. M. (2008). The effect of acoustical similarity on lexical-tone perception of one-year-old Mandarin-learning infants. *中華心理學刊*, *50*(2), 111-124.

Tsao, F. M., & Liu, H. M. (2020). Lexical-Tonal Perception Development in Infancy. *Speech Perception, Production and Acquisition*, 177-197.

Walker, P., Gavin Bremner, J., Mason, U., Spring, J., Mattock, K., Slater, A., & Johnson, S. P. (2010). Preverbal infants' sensitivity to synaesthetic cross-modality correspondences. *Psychological Science*, *21*(1), 21–25.

Wang, X., & Peng, G. (2014). Phonological processing in Mandarin speakers with congenital amusia. *Journal of the Acoustical Society of America 136*(6), 3360–3370.

Weiss, A. H., Granot, R. Y., & Ahissar, M. (2014). The enigma of dyslexic musicians. *Neuropsychologia 54,* 28–40.

Welch, G. F. (1983). *Improvability of poor pitch singing experiments in feedback* (Doctoral dissertation, Institute of Education, University of London).

Welch, G. F. (1985). A schema theory of how children learn to sing in tune. *Psychology of Music*, *13*(1), 3-18.

Welch, G. F., Rush, C., & Howard, D. M. (1991). A developmental continuum of singing ability: Evidence from a study of five-year-old developing singers. *Early child development and care*, *69*(1), 107-119.

Werker, J. F., & Tees, R. C. (1984). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant behavior and development*, *7*(1), 49-63.

Wermke, K., and Mende, W. (2009). Musical elements in human infants' cries: in the beginning is the melody. *Musicae Scientiae, 13,* 151–175.

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.

Werner, L. A., and Marean, G. C. (1996). *Human Auditory Development.* Madison, WI: Brown Benchmark.

Wieland, E. A., McAuley, J. D., Dilley, L. C., & Chang, S. E. (2015). Evidence for a rhythm perception deficit in children who stutter. *Brain & Language 144,* 26–34.

Winkler, I., Háden, G. P., Ladinig, O., Sziller, I., & Honing, H. (2009). Newborn infants detect the beat in music. *Proceedings of the National Academy of Sciences*, *106*(7), 2468-2471.

Wray, A., & Grace, G. W. (2007). The consequences of talking to strangers: Evolutionary corollaries of socio-cultural influences on linguistic form. *Lingua*, *117*(3), 543–578.

Yeung, H. H., Chen, K. H., & Werker, J. F. (2013). When does native language input affect phonetic perception? The precocious case of lexical tone. *Journal of Memory and Language*, *68*(2), 123-139.

Yu, M., Xu, M., Li, X., Chen, Z., Song, Y., & Liu, J. (2017). The shared neural basis of music and language. *Neuroscience, 357*, 208-219.

Zhang, C., Shao, J., & Huang, X. (2017). Deficits of congenital amusia beyond pitch: Evidence from impaired categorical perception of vowels in Cantonese-speaking congenital amusics. *PLoS ONE,* 12(8), e0183151.

Zhang, Y., Kuhl, P. K., Imada, T., Iverson, P., Pruitt, J., Stevens, E. B., ... & Nemoto, I. (2009). Neural signatures of phonetic learning in adulthood: a magnetoencephalography study. *Neuroimage*, *46*(1), 226-240.

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