rpm Revista Portuguesa de Musicologia
PORTUGUESE JOURNAL OF MUSICOLOGY

nova série | *new series*
7/2 (2020), pp. 329-350 ISSN 2183-8410
http://rpm-ns.pt

Singing in the Brain: When Melody and Words Come Together

Ana Isabel Pereira
LAMCI – CESEM
Faculdade de Ciências Sociais e Humanas
Universidade NOVA de Lisboa
anapereira@fcsh.unl.pt

Helena Rodrigues
LAMCI – CESEM
Departamento de Ciências Musicais
Faculdade de Ciências Sociais e Humanas
Universidade NOVA de Lisboa
helenarodrigues@fcsh.unl.pt

Resumo

Nas últimas décadas, o estudo do binómio música e linguagem tem suscitado interesse em vários ramos das ciências cognitivas, incluindo a psicologia, linguística, antropologia, musicologia, neurociência cognitiva e educação. Indubitavelmente, as canções são o material de eleição para estudar a relação entre música e linguagem. Este artigo explora contribuições das neurociências que podem ser consideradas interessantes para o campo da educação musical, especialmente no que respeita a relação da melodia e palavras em canções. A influência destas componentes na perceção e *performance* vocal de canções é ainda uma questão em debate tanto na área das neurociências como da educação musical. A fim de se estabelecer um pano de fundo para esta discussão, são apresentadas, resumidamente, as semelhanças evolutivas entre música e linguagem, bem como os mecanismos partilhados de aprendizagem para ambos os domínios, conforme ilustrado por vários estudos. Uma vez que o *pitch* e o ritmo são componentes importantes das canções, é ainda abordada a pesquisa comparativa de ambos os elementos no cruzamento da música e linguagem. Nesta interseção, é dado um enfoque especial à Teoria de Aprendizagem Musical, proposta por Edwin Gordon, em que se defende o uso de canções apresentadas tanto com texto como com sílaba neutra desde a infância. Considerando que as canções são um dos recursos mais utilizados na educação musical, questiona-se se os avanços científicos nas neurociências podem informar a pedagogia musical. Novos caminhos de investigação são sugeridos na interseção das duas disciplinas.

Palavras-chave

Canções; Educação musical; Neurociências; Canções com texto; Canções com sílaba neutra.

Abstract

For the past decades, the study of the binomial music and language has been of interest in several branches of cognitive sciences, including psychology, linguistics, anthropology, musicology, cognitive neuroscience, and education. Undoubtedly, songs are the perfect medium to study the relationship between both domains. This article will explore some contributions from neurosciences that could be deemed interesting to the field of music education, focusing on the relationship between melody and words in songs. The influence of these components on song perception and production is an ongoing matter of debate both in neurosciences and music education. The background for this discussion will be set by first mentioning the evolutionary commonalities between music and language, and a discussion on the shared leaning mechanisms for music and language. Since pitch and rhythm are important songs' components, the comparative research of these elements across music and language will also be approached. In the intertwining of both fields, a special focus will be given to Music Learning Theory, a framework proposed by Edwin Gordon, who advocates the use of songs presented both with text and neutral syllable since infancy. Considering that songs are one of the most used resources in music education, it is questioned if the scientific advances in the neurosciences can inform musical pedagogy, thus new paths of investigation are suggested at the intersection of the two disciplines.

Keywords

Songs; Music education; Neurosciences; Songs with text; Songs with neutral syllable.

> Dire & chanter etoient autrefois la même chose, dit Strabon [...]. Sur la maniere dont se lièrent les premieres sociétés, étoit-il étonnant qu'on mit en vers les premieres histoires, & qu'on chantât les premières loix? Etoit-il étonnant que les premiers Grammairiens soumissent leur art à la Musique & fussent la fois professeurs de l'un & de l'autre? (ROUSSEAU 2012, 407)

EVIDENCE HAS SUPPORTED THE CLAIM THAT SONGS are a universal human feature, recognizing patterns of variability and regularity across cultures (MEHR et al. 2019; SAVAGE - BROWN - SAKAI - CURRIE 2015). Singing is engrained in human society, thus it comes with no surprise that songs are an important part of music classes curricula in different countries all around the world. For example, in Portugal, the elementary school guidelines for music education clearly state that singing practice is the basis for music education (ME n.d.). Considering that the human voice is the first musical instrument that everyone possesses, to communicate and participate in music making through singing is within everyone's capabilities, and it should be nurtured. Nonetheless, the outcomes of the teaching and learning of songs depend on several factors. In this article, we will focus on the relationship between melody and words (also referred to in literature as text or lyrics), specifically how it might impact the way songs are taught and sung. As Johanne TAFURI (2008) points out, singing songs is not solely the result of the simple juxtaposition of both components, but rather a third with its own significance, and which arises from the interaction between them. One might say that songs represent the intertwining between music and language. Research in the neurosciences has shown that music and language present overlapping and shared mechanisms, more specifically for lexical/phonological processing in language and melodic processing in music. Yet, different levels of processing have been detected depending on the experimental conditions (ETTLINGER - MARGULIS - WONG 2011; SCHÖN et al. 2010). Music and language also share the same auditory channel. Furthermore, speech and musical sounds rely on the same acoustic parameters: frequency, duration, intensity, and timber (BESSON - CHOBERT - MARIE 2011; ETTLINGER - MARGULIS - WONG 2011; PATEL 2008). Therefore, the association of melody and text in songs offers interesting research questions concerning songs' structural properties (is the whole greater than the sum of its parts, as Aristotle and the Gestalt psychologists claim?) and their underlying processing and production mechanisms. Definitely, the study of music and language plays a role in music education, and the effort to link this area with the neurosciences is important. To approach this binomial from a wider view, let us start at the very beginning.

Evolutionary Commonalities Between Music and Language

It seems to be consensual that music and language are universal features of human society.[1] Both are used to convey and influence emotions, and some authors also mention that one of the central features shared by both is cultural transmission (KIRBY 2012; REBUSCHAT et al. 2012; SLOBODA 1985, ZATORRE - BAUM 2012). Different questions have been raised on this matter: Has music evolved from language? Or has language evolved from music? Have music and language evolved in parallel to each other, as if they were different systems? Or was there a single precursor for both music and language? The evolutionary commonalities between music and language have been strongly defended by different authors. The first original idea comes from Charles Darwin, who suggested that our ancestors developed a capacity for an early musical protolanguage, which provided the basis for the later evolution of language. The author also posits that sexual display was responsible for the emergence of this protomusic: 'It appears that the progenitor of man, either the males or females or both sexes, before acquiring the power of expressing their mutual love in articulate language, endeavoured to charm each other with musical notes and rhythm' (DARWIN 1871, 571).

Similarly, Jean-Jacques Rousseau advocates that 'music and language share a common ancestor and that language evolved out of music for the sake of a rational organization of human societies' (JENTSCHKE 2016, 3). On the other hand, Steven Mithen, one of the pioneers of cognitive archeology, suggests that the Neanderthals 'had an advanced form of communication that made extensive use of variations in pitch, rhythm, and timbre but did not use words or grammatical rules' (MITHEN 2005, 107). Therefore, the author defends a form of holistic communication of a type that has no analogy in the modern world. Steven MITHEN (2005) also posits that music has emerged from the remnants of 'Hmmmmm' (an acronym for holistic, manipulative, multi-modal, musical and memetic communication) after language evolved. He explains that:

> Compositional language [the use of words and a set of grammatical rules] took over the role on information exchange so completely that 'Hmmmmm' became a communication system almost

[1] The word 'language' will be used throughout this article as the use of speech.

> entirely concerned with the expression of emotion and the forging of group identities, tasks at which language is relatively ineffective. Indeed, having been relieved of the need to transmit and manipulate information, 'Hmmmmm' could specialize in these roles and was free to evolve into the communication system we now call music (MITHEN 2005, 266).

Steven Mithen distinguishes music from language describing the latter as being referential and manipulative, i.e., 'some utterances refer to things in the world, while others make the hearer think and behave in particular ways', whereas the former being essentially manipulative because 'it induces emotional states and physical movement by entrainment' (MITHEN 2005, 25). The same distinction has been expressed by Steven Brown but using other words: 'music emphasizes sound as emotive meaning and language emphasizes sound as referential meaning', therefore 'music and language are reciprocal specializations of a dual-natured referential emotive communicative precursor' (BROWN 2001, 271). This author believes that music and language have evolved from a common ancestor, which he coined as a 'musilanguage' stage. A critique pointed by Steven MITHEN (2005) to this model is that it makes no reference to human evolution and provides no specified date in the past for its existence.[2] The 'musilanguage' model is then based on the analysis of phrase structure and phonological properties of musical and linguistic utterances.[3] For the author, the final step of this model is the development of interactive properties by a coevolutionary process, where songs fit in:

> […] the rebinding of music and language in the form of novel functions that evolve parallel to their separation. The emergence of these interactive functions reflects coevolution of the underlying linguistics and musical systems. Thus, we can imagine verbal song[4] as evolving through a series of stages that parallel biological developments in both systems (BROWN 2001, 296).

A direct divergent perspective from both the 'musilanguage' model proposed by Steven Brown and the musical protolanguage defended by Steven Mithen is put forth by Christian Lehmann, Lorenz Welker, and Wulf Schiefenhövel: 'singing is primarily a mode of saying, and […] music without words is a "derived" trait'. Moreover, for the authors: 'the anatomical specialization of the vocal tract for producing an ambit of three octaves and a wide variety of timbres, and the precise auditory-laryngeal control of intonation are more closely related to our capacity to sing than to our capacity to speak.' (LEHMANN - WELKER - SCHIEFENHÖVEL 2009, 324)

---

[2] See a discussion on both the Brown's musilanguage' model and the Mithen's 'Hmmmmm' theory in BOTHA (2009).

[3] For a full explanation, see BROWN (2001).

[4] Steven BROWN (2001) defines verbal songs as songs with words or as words with music.

Thus, one factor that strongly supports this model would be that speaking voice only uses certain capacities of the vocal apparatus and makes less use of breathing control abilities than singing does. Therefore, the authors corroborate the idea that 'we can speak because we can sing' (LEHMANN - WELKER - SCHIEFENHÖVEL 2009, 328). On the other hand, Stefan KOELSCH (2011, 15) suggests that music and language are 'different aspects of the same domain, or two poles of a rather continuous dimension, rather than being two strictly separate domains', a phenomenon that he coins as the music-language continuum. He argues that:

> Once an individual puts emphasis in his/her utterances, the speech becomes more song-like, and many art- forms, such as Rap-music or Recitatives, are both song and speech. In addition, […] music is often structured according to a syntactic system, and during music listening, meaning emerges from the interpretation of musical information. Therefore, any clear-cut distinction between music and language (and thus also any pair of separate definitions for language and music is likely to be inadequate, or incomplete, and a rather artificial construct. (KOELSCH 2011, 15)

Although representing a totally distinct perspective, it should be mentioned the one put forward by Steven Pinker, a cognitive psychologist. He claims that music is a non-adaptive by-product of language, that is, language was the reason that the neural pathways developed as they did, and music subsequently used those pathways. The author claims that music is not a legitimate biological adaptation and it serves no role in human evolution:

> As far as biological cause and effect are concerned, music is useless. It shows no signs of design for attaining a goal such as long life, grandchildren, or accurate perception and prediction of the world. Compared with language, vision, social reasoning, and physical know-how, music could vanish from our species and the rest of our lifestyle would be virtually unchanged. Music appears to be a pure pleasure technology, a cocktail of recreational drugs that we ingest through the ear to stimulate a mass of pleasure circuits at once. (PINKER 1997, 528)

This view, which is a minoritarian perspective among scholars, has been criticized by many authors, especially because it narrows the function of music to entertainment (LEVITIN 2006; TREHUB 2003). As Tecumseh FITCH (2006, 85) stresses, 'by exploring and comparing the cognitive, neural, and genetic mechanisms underlying musical and linguistic abilities', it will be possible to test many interesting questions raised by Darwin, especially the musical protolanguage hypothesis. Indeed, in the last decades, the neurosciences have played an important role in exploring these shared mechanisms of music and language.

Shared Learning Mechanisms for Music and Language

The fact that both music and language share the same auditory and communicative channels, which involve complex sound systems and sensorimotor sequencing demands, is of particular importance in song teaching. For instance, when considering the role of words (text) in songs, several studies report its dominance in song acquisition (DAVIDSON 1985; DAVIDSON - MCKERNON - GARDNER 1981; MOOG 1976; RUTKOWSKI 1998; WELCH - SERGEANT - WHITE 1995-6). The importance of words in children's social lives might explain this dominance (WELCH 1998; WELCH - SERGEANT - WHITE 1995-6) or also the fact that words (text) combine elements such as pitch, rhythm, and timbre, thus enriching the combination of vowels and consonants as if it were an additional musical timbre (MOOG 1976). Moreover, children's stage of language development might also influence learning and vocal performance (WELCH 2006). Therefore, one of the questions debated has been: does text mask the properties of a melody, such as melodic contour or rhythm, thus affecting children's singing accuracy? Graham WELCH et al. (1998) suggest that the melody should be taught separately from text in order to correct the text bias. However, their study did not compare the influence of both strategies on singing accuracy, that is, teaching a song with melody and text combined (a song with text) and teaching a song by starting with the melody (the equivalent to a song with a neutral syllable), adding the text later. Nonetheless, some researchers have argued that 'the presence of melody enhances recall for words when there is a learned association between them' (GINSBORG - SLOBODA 2007, 424), that is, both components are better recalled when rehearsed together. Yet, empirical evidence is contradictory regarding the interaction between melody and words in song singing, specifically whether children sing better with text or a neutral syllable. For example: better with a neutral syllable (GOETZE 1985); significantly higher scores with text for the youngest group (JACOBI-KARNA, 1996); slightly better with a neutral syllable (LEVINOWITZ 1989); no significant differences but slightly higher scores with text (RUTKOWSKI 1993); no significant differences but better with a neutral syllable (SMALE 1987); better with text (WELCH et al. 1995-6). Results from a recent study conducted with 137 Portuguese children aged four to nine years-old corroborate the fact that vocal performance is better when melody and text are taught together for a large group, and significantly better for children aged four and five. However, results also show that a smaller group within this sample benefited from learning words and melody separately (PEREIRA 2019). The comparison between studies is limited since different procedures have been used to investigate the effect of this variable. Nevertheless, this empirical evidence coming from music education suggests, at least, the existence of shared learning mechanisms for music and language.

The extent to which the systems of music and language overlap is an active area of research and debate.[5] A few decades ago, music and language were seen as different psychological faculties. The lateralization of music and speech was then the current theory, and it was thought that the music functions were localized in the right-hemisphere of the brain whereas the speech functions were located on the left-hemisphere (JÄNCKE 2012). The advances of modern brain imaging techniques have challenged this old theory. A considerable amount of research has already explored the relationship between music and language, seeking their common mechanisms. Sebastian JENTSCHKE (2016) groups these studies in three categories, which might report:

> (1) an anatomical overlap (e.g., similar pattern of brain activation when processing musical and linguistic stimuli); (2) functional interactions between cognitive processes (e.g., ERP[6] responses in one domain are modulated when similar processes in the other domain are carried out concurrently); or (3) transfer effects (i.e., enhanced processing in one domain, following training in the other domain, for example as a consequence of musical training or experience with a tonal language) (JENTSCHKE 2016, 5).

Research has revealed that each activated brain area is a vast region that can accommodate more than one distinct processing network (PERETZ 2012). Studies pointing to the overlapping of neurophysiological, perceptual and cognitive processes in music and language, suggest that it might not imply sharing neural circuits (BESSON et al. 2011; PATEL 2011). It might rely on entirely different (non-overlapping) neural populations that serve multiple functions (JENTSCHKE 2016), even if overlapping in neural networks is observed. Although revealing hemispheric dominances, research also suggest the existence of a shared learning mechanism for the system of spoken language and for music but with partially distinct neural networks. Examples can be found in studies with patients with acquired amusia,[7] where brain networks can be specialized for music functions without overlapping with networks involved in language (ETTLINGER - MARGULIS - WONG 2011; KOELSCH et al. 2002; SAITO et al. 2012; SCHÖN et al. 2010). In the case of song singing, Isabelle PERETZ (2012) argues:

5 See the contributions in, e.g., *Frontiers in Psychology*, special issue, 2 (2012); JENTSCHKE (2016); SAITO et al. (2012); and extensive overviews on PATEL (2008); REBUSCHAT et al. (2012).

6 ERP stands for Event-Related Potential, i.e., a type of brain wave that is associated with a response to a specific stimulus, such as a particular wave pattern observed when a patient hears a clicking sound. See <http://medical-dictionary.thefreedictionary.com/Event-Related+Potential> (accessed December 12, 2018).

7 Acquired amusia is the occurrence of selective deficits in the auditory processing of music after in an individual previously able to process music normally. For a comprehensive view on a cognitive framework to understand and assess amusia see, e.g., CLARK - GOLDEN - WARREN (2015).

> The number of networks involved is particularly large in the case of production tasks since the output system also involves the perceptual systems for auditory monitoring. Many of these processing components might be shared between music and speech, especially when singing contain lyrics. (PERETZ 2012, 260)

For example, in Yoko SAÏTO et al. (2006), adult participants were asked to recite and sing the text of a well-learned song both *a cappella* and along a pre-recorded voice. Results reveal that singing solo and with the recording activated specific brain areas that were not involved when speaking the text, but no specific areas were activated in the speaking condition. Although the 'melody' condition was not tested (singing without text, i.e., using a neutral syllable), results suggest the existence of specific brain areas related to melody production. Accordingly, Erin MCMULLEN and Jenny SAFFRAN (2004) have discussed that several musical and linguistic processes appear to be modularized in adults,[8] but this might not be true for children, for whom this process would be emergent and not present since birth. Stefan KOELSCH (2011, 16) also suggests that 'the human brain, particularly at an early age, does not treat language and music as strictly separate domains, but rather treats language as a special case of music'. Curiously, studies with infants—who are not yet able to speak like adults—have shown that a precocious distinction might exist between the singing and speaking voices. Results reveal that a greater extension of vowels in vocalizations was produced in response to the singing condition than to the speaking condition (REIGADO - RODRIGUES 2018; REIGADO - ROCHA - RODRIGUES 2011). It would be interesting to investigate which brain areas are activated in children for both conditions, and to compare it with adults. As Aniruddh PATEL (2011, 214) posits, the 'comparative music-language research is still in its infancy'.

Nevertheless, studies involving people suffering from brain damage have contributed to the ongoing discussion on the relationship between music and language in songs, since it is possible to observe how the brain works in the absence of certain functions.[9] Isabelle PERETZ and Robert ZATORRE (2005, 107) believe that 'the convergence between lesion and neuroimaging data is the optimal strategy to build a sound model of how the brain processes music'. These researchers also suggest that individual differences might exist on the mapping of the human brain processes, as well as at what level they occur. Eventually, this might depend on the relation between environmental inputs (e.g., musical training, experience), innate factors (e.g., musical aptitude) and development,

[8] Modularity (or domain-specificity) refers to the existence of 'a distinct mechanisms that deals with a particular aspect of the input and does this either exclusively or more effectively than any other mechanism. […] A domain may be as broad and general as auditory scene analysis and as narrow and specific as tonal encoding of pitch' (PERETZ 2012, 256).

[9] A huge amount of research has been done in music education concerning songs' production, less developed in songs' perception, but this article's purpose is to explore neurosciences' findings and not the other way around.

which varies greatly among people. Therefore, we believe that evidence should be carefully interpreted. In order to illustrate possible behaviors concerning song's production, i.e., song's vocal performance (with text and/or neutral syllable), a few studies from the neurosciences are pointed: (1) an aphasic patient without amusia that is able to sing familiar song melodies and learn new ones but cannot produce these melodies' words intelligibly, either spoken or sung (HÉBERT - RACETTE - GAGNON - PERETZ 2003).[10] The authors suggest that neurally separable modules might exist for both singing and speaking; (2) an adult diagnosed with primary progressive aphasia (a severe speech disorder) that was able to sing in tune, but the text was often unintelligible. It was observed that after singing the first word correctly, this patient switched to a neutral syllable (such as 'la') to continue singing. The authors suggest that 'the verbal production, be it sung or spoken, is mediated by the same (impaired) language output system and that this speech route is distinct from the (spared) melodic route' (PERETZ - GAGNON - HÉBERT - MACOIR 2004, abstract); (3) a girl with autistic disorder and severe mental handicap who can sing songs but cannot speak (MOGHARBEL - SOMMER - DEUTSCH - WENGLORZ - LAUFS 2005-6). The authors suggest that she has a mental representation of the phonetic form of songs' text, which belongs to reproduction in the same way as the melody and rhythm; and (4) out of a group of eleven amusic participants, two of them are able to sing proficiently songs with text, while the other nine are not. The authors suggest that this was probably due to a strong association between melody and text in memory or by relying on an integrated representation of melody and text. In addition, among those two special cases, one of them was unable to sing a song with neutral syllable, whereas the other produced a complete performance of the song (DALLA BELLA et al. 2009).[11] The authors then suggest that when the task requires the association of a new speech segment, such as the neutral syllable 'la' or 'ta', to a familiar melody, the information retrieval from memory may become impossible. This poor memory representation was not due to melodic complexity, since the authors took that into account. Probably, to sing the song with a neutral syllable was a task never experienced or trained before. Thus, results might also be a consequence of the task's nature. Would the results differ if the words and the melody of the song had been taught separately, and only after combined? Following these reported differences on individuals with the same impairments, it is possible to assume different ways of perception and production among individuals. Even among people with no speech or hearing impairment, what happens in the brain concerning the perception

---

[10] Aphasics patients are people suffering from partial or total loss of the ability to articulate ideas or comprehend spoken or written language, resulting from damage to the brain from injury or disease. Aphasic (n.d.), *WordNet 3.0, Farlex clipart collection* (2003-2008), available at < https://www.thefreedictionary.com/aphasic> (accessed December 14, 2018)

[11] Amusia is a form of aphasia characterized by the loss of ability to produce (motor agnosia) or to recognize (sensory amusia) musical sounds. Amusic (n.d.) *Farlex Partner Medical Dictionary* (2012), available at <https://medical-dictio nary.thefreedictionary.com/Amusic (accessed December 14, 2018). In this study, the amusic participants suffer from a perceptual pitch deficit.

and production of songs may differ among them. It should also be reminded that different experimental conditions and tasks might condition results and subsequent comparisons. In brief, among people suffering from the same impairment or with no impairments, different perception and production behaviors might occur. Findings reveal that the interaction between text and melody in a song may assume different forms.

Transfer effects are out of scope of this article.[12] Yet, related to the findings discussed above, it is interesting to mention that 'because speech and music share sub-cortical pitch-processing circuits, speech processing benefits' (PATEL 2012, 125). Pitch, as well as rhythm, is among the most important common auditory features between music and language processed in the brain (see contribution in, e.g., KRUMHANSL 2000).[13] This topic is also of interest to song singing and teaching.

Pitch and Rhythm: Common Auditory Features Between Music and Language

Every human being is born with two distinct sound systems, where pitch and rhythm play an important role: (1) linguistic, which includes the vowels, consonants, pitch contrasts of the native language, and rhythm, i.e., patterns of timing and accentuation that characterize the flow of syllables in sentences; and (2) musical, which includes timbres, rhythm, and pitches of the music's culture (PATEL 2008). Pitch and rhythm have been studied separately both in music and language (speech), but less research has been devoted to the comparative study of each element across domains. The purpose of this section is to highlight a few thoughts from the comparative research that might be relevant to music education and song singing, rather than providing an extensive literature review.[14]

What does research say about the relationship between the pitch in melodies and the pitch pattern of the spoken lyrics/text? In music, individual pitches can be combined in a stable way to shape a song's melody, for example. In language, and although humans are capable of speaking on a single tone (monotone), it rarely happens. Instead, a set of voice pitch modulations are observed, also known as prosody. Prosody refers to a set of speech parameters, including intonation (fundamental frequency, i.e., pitch variation across a sentence), stress and rhythm (PATEL 2008; ZATORRE - BAUM 2012).[15] Julia MERRIL and Pauline LARROUY-MAESTRI (2017) have found that

---

[12] See meta-analyses on transfer effects from music training to literacy development in GORDON - FEHD - MCCANDLISS (2015).

[13] There is still no universally accepted definition for either concepts. Possible definitions for pitch and rhythm are put forth by PATEL (2008): pitch is considered as the primary basis for sound categories in music, such as intervals or chords; while rhythm can be considered as the systematic patterning of sound in terms of timing, accent and grouping.

[14] For a comprehensive overview, mainly on instrumental music and not songs, see PATEL (2008).

[15] Naturally, within this area of research, more studies have focused on tonal languages than others since pitch defines the lexical distinction between words' meanings (WONG - DIEHL 2002; YUNG 1991).

pitch is one of the most important acoustical features when determining if a vocalization stands out as song or speech. Their findings were based in an experiment with adult listener's while seeking to clarify which vocal features influenced adult listeners' perception of speechsong compositions (which are neither typical for speech nor song) by Arnold Schoenberg. On the other hand, research has reported the illusion effect in which a spoken phrase is perceptually transformed to sound like song rather than speech, by repeating it several times over (DEUTSCH - HENTHORN - LAPIDIS 2011). It was hypothesized that due to repetition, the pitches belonging to the spoken phrase would increase in perceptual salience and would also be distorted to conform to a well-formed tonal melody. As Diana DEUTSCH (2013, 313) posits, 'it appears that the neural circuitries, underlying the perception of speech and song can accept the same input, but process it differently, so as to produce different outputs'. Corroborating this idea, Robert Zatorre and Shari Baum argue that pitch information processing differs greatly between speech and music. Evidence indicates that despite some shared cognitive processes and neural substrates for processing pitch, two systems exist: 'one focuses on contour, which may overlap across domains, and another, perhaps specific to music, involving more accurate pitch encoding and production' (ZATORRE - BAUM 2012, 5). To illustrate, the authors mention that considering certain styles within the western tonal musical system, relatively small deviations from certain intervals in a melody can be considered as errors by listeners, whereas deviations of a similar magnitude in speech contour are not.[16] The distinction between the two mechanisms is also supported by studies with people suffering from congenital amusia, i.e., who exhibit a deficit in detecting pitch changes in melodies (amusia), a condition present from birth (congenital). Based on past findings, the authors noted that these patients:

> […] have little difficulty perceiving large changes in pitch contours typical of speech […]. When measured with stimuli that have small pitch deviations, however, these individuals show impairments, whether the stimuli are speech or not […], indicating a selective deficit at the level of fine-grained pitch distinctions […], which are not as critical for speech as they are for music […] (ZATORRE - BAUM 2012, 4).

These findings refer to neuropsychological data. Nonetheless, evidence from neuroimaging has shown the existence of brain responses to fine-grained pitch differences in amusic individuals (e.g., PERETZ - BRATTICO - JÄRVENPÄÄ - TERVANIEMI 2009), which is not integrated in a conscious percept probably because is not supported by long-term memory representations of tonality schemas.

[16] It is expected that both musical training and previous exposure to melodies/songs impact on judging accuracy. See LARROUY-MAESTRI (2018).

Aniruddh Patel suggests the comparison between the degree of correspondence of linguistic intonation—which is not built around a stable set of pitch intervals—and the melody in songs by implementing two-steps: (1) to identify the melodic contour of the text when it is spoken; and (2) to compare it with the melodic contour to which the text is set in a song. He also suggests that 'it would be interesting to know if songs with melodies congruent with the spoken pitch contour have an advantage in terms of how easily they are learned or remembered' (PATEL 2008, 218). In music education, when teaching the words of a song, the melodic rhythm should be kept intact since it will facilitate fitting them onto the melody (or vice versa), that is, a direct link between the linguistic prosody and musical meter:

> While in spoken verse the underlying metrical scheme is grounded in the prosody of the language in which it is composed, in sung verse the structure is created by the mapping of specific prosodic units of the text (syllables, moras, tones, etc.) onto the rhythmic-melodic structure provided by the tune (PROTO - CANETTIERI - VALENTI 2015, abstract).

The idea put forth by Aniruddh PATEL (2008) could be interesting to explore in song teaching but using a different approach. For example, the songs' words could be taught using the melodic contour of the melody (e.g., if the melody rises, the speech contour would rise as well), which, of course, might not match the melodic contour of those words when spoken. However, to work the text expressively, intentionally using the voice-pitch modulations, could enhance singing, although the fine-grained pitch information is not the same as using the singing voice, understandably. Nevertheless, if the idea put forth above by Robert ZATORRE and Shari BAUM (2012) is accurate, that is, the system focused on the contour overlaps across the music and language domains, then this focus on the expressiveness of the text might have positive effects on singing accuracy. On the other hand, it would be interesting to test if songs that use words whose spoken contour is similar to the contour of the melody are better sung.

And what does research say about the relationship between the rhythm in melodies and in spoken lyrics/text? There exists a considerable amount of literature on how language prosody is reflected in songs (PALMER - KELLY 1992; GORDON - MAGNE - LARGE 2011). The prevailing notion is that, at least for languages with distinct patterns of stronger or weaker syllables (such as English and Portuguese), the syllabic accent patterns in speech are related to the musical metrical accent patterns, i.e., stronger syllables in speech are expected to occur on hierarchical prominent musical beats (LERHDAL - JACKENDOFF 1983; PATEL 2008). Each language has its characteristic patterns of timing and accentuation. As Aniruddh Patel posits, the failure to acquire native rhythm originates a foreign accent in speech. This author also suggests an interesting line of research to answer the above-posed question, which 'pertains to cultural differences in the prevalence of certain types of musical rhythms' (PATEL 2008, 157). For example, Yamomoto noted that it is rare to find Japanese children's songs in triple meter (based in triple rhythms within the 6/8 signature) while in Britain they are commonly found (PATEL, 2008). The author then suggests that it might be due to differences in English versus Japanese speech rhythms. Related to the enculturation processes, a question for future research could be: Does songs' vocal performance differ between British and Japanese children depending on the meter of the songs taught (duple/triple)? A vast majority of studies with children used song material in duple meter (GUERRINI 2006; HORNBACH - TAGGART 2005; JACOBI-KARNA 1996; KLINGER - CAMPBELL - CROSBY 1998; LEIGHTON - LAMONT 2006; NICHOLS 2016). Fewer studies have used triple meter songs (MANG 2006; WISE - SLOBODA 2008). Two studies used songs in both meters, but the purpose was not to compare song's vocal performance considering this variable (DEMOREST - PFORDRESHER 2015; WELCH et al. 2011). To our knowledge, no research has been done to investigate the influence of meter on song singing. Further studies could investigate this line of research, within the same culture and cross-culturally (such as between Japanese and British children).

Another example: Erin HANNON and Sandra TREHUB (2005) sought to investigate the perception of temporal changes in folk melodies with simple (regular meter, e.g., 2/4) and complex metrical structure (irregular meter, e.g., 7/8) of north American, Bulgarian or Macedonian adults, and six months old north American infants. The researchers assumed that north American adults and infants would have had more exposure to simple duration ratios, whereas the others would have been exposed to both simple and complex duration ratios. Results reveal that north American infants were more similar to Bulgarian and Macedonian adults than to north American adults since they differentiated alterations that violated the metrical structure of musical patterns from those that preserved the metrical structure both when the original pattern had simple or complex meter.[17] The authors suggest that:

> These findings imply that the temporal perception and production biases of North American adults arise from extended exposure to the simple metrical structures that predominate in Western music. Human listeners may begin with flexible processing of metrical structure, which facilitates perception of temporal nuances in various kinds of music. Months, perhaps years, of exposure to the dominant metrical categories of a specific musical culture may prompt perceptual reorganization or narrowing of the metrical frameworks that can be handled with ease (HANNON - TREHUB 2005, 53).

[17] Future replications of Erin HANNON and Sandra TREHUB (2005) should use vocal stimuli instead of instrumental, since research has shown that timbres could influence perception in different manners (KRUMHANSL - IVERSON 1992; WESSEL 1989).

Facing these results, could it be considered that a continued exposure since early years to meters not commonly present in children's music culture would benefit learning? In fact, findings from the study reported above align with what Edwin Gordon defends:

> Because listening serves as a preparation for learning to sing and to audiate, the more music a child younger than eighteen months hears and the greater variety and balance among listening experiences in terms of tonalities, meters, and musical styles, the better the child will be to learn to sing, move, and audiate[18] (GORDON 2012, 246-7).

As seen, there are still new lines of research to pursue in comparative research across music and language, specifically concerning pitch and rhythm. In the crossover, it has been reported that rhythm disorders can occur independently from pitch disorders (DI PIETRO - LANGANARO - LEEMAN - SCHNIDER 2004, cit. in PERETZ 2012), and the pitch dimension might be impaired but not the time dimension, meaning that individuals can sing in-time although out of tune (ALCOCK - WADE - ANSLOW - PASSINGHAM 2000; DALLA BELLA - GUIGUÈRE - PERETZ 2009). In the neurosciences and music education, both variables should continue to be considered separately as well as intertwined when investigating its influence on songs' perception and production. These areas of research should consider human behavior from the maternal womb.

To Sing Songs With Text and With a Neutral Syllable: Is it Important?

Evidence on brain plasticity has shown that several different processes and types of plasticity exist, and that some of them are more dominant in certain life periods and less dominant in others (KOLB - GIBB 2011).[19] Children's brain development is influenced by different environmental factors and rapidly changing. In the field of music education, Edwin GORDON (2003; 2012) acknowledges this fact and stresses that some periods in children's lives are crucial for learning. He refers that the critical period characterized by surges of neurological connections and synapses begins prenatally and continues until eighteen months; and, in addition, this period will interact with the sensitive period which ends at around five years old. Edwin Gordon explains that:

[18] Audiation is a word coined by Edwin Gordon. According to this author (GORDON 2007; 2012), audiation is to music what thought is to language, and although music is not a language, the process of audiating and giving meaning to music is similar to thinking and giving meaning to language. Audiation occurs when a person hears and comprehends music for which sound is not physically present (as in recall), is no longer present (as in listening), or may have never been present (as in creativity and improvisation). It should not be confounded with imitation, which is often referred to as inner hearing, and provides the necessary readiness for audiation. One can imitate without comprehending. Also, aural perception, memory, imagery, and recognition are different from audiation, and some become part of the audiation process.

[19] This process is designated as neuroplasticity, i.e., structural and functional changes in the brain that are brought about by training and experience, or in other words, it describes how training and experience reorganize neural pathways in the brain. See CHUDLER (n.d.).

> Nature provides humans with an abundance of cells that encapsulate chromosomes which carry molecular DNA and genes. Unless cells are stimulated to make syntactic connections, they are lost, not to be regained. […] if a very young child has no opportunity to develop a music listening vocabulary, cells that would have been used to establish the music listening sense will at best be directed to another sense […]. No amount of effort later will be able to offset this handicap completely (GORDON 2012, 246)

Edwin Gordon also advocates that children should be exposed both to songs with text and a neutral syllable (such as bah, la, and keeping the speech-rhythm information). The latter would be particularly important until the age of eighteenth months, since beyond that age words (with semantic meaning) will become the main focus. According to Gordon teaching songs with a neutral syllable seems to favor children in two aspects: (1) it provides them with a greater focus on the intrinsically musical aspects of the melody, such as pitches, intervals, intensities; and (2) it allows them to concentrate on audiating tonal and rhythm patterns, tonality and meter. In spite of the importance of songs presented with a neutral syllable, this pedagogue and researcher considers that songs with text, such as folk songs or other songs from the literature, are also important, but they should come at the right place of the sequence (GORDON 2007).[20] The reason pointed is that adding text contributes to a shift of attention to what is most familiar to children. Thus, they might not listen to the melody, and as a consequence do not learn to audiate (GORDON - JORDANOFF 1993).

So, when it comes to songs in music education, there are two different aspects to be considered: (1) songs with a neutral syllable may exist per se, that is, it is not necessary to have words in songs. Songs with text are used to tell stories, to help routines, to approach values and attitudes, among others, and they are most valuable. Yet, songs presented with a neutral syllable also provide an enormous potential to engage children in different classroom activities without focusing on a specific message; (2) in order to teach a song with text, two different strategies might be used, namely, melody and text combined; and melody and text separately, and then combined. In this case, teaching the melody is the equivalent to a song presented with a neutral syllable. As mentioned before, some authors also acknowledge that teaching the text and melody separately would benefit singing accuracy (WELCH et al. 1998). A recent study does not corroborate this recommendation at least for children aged four and five years-old, but it recognizes that for some children the strategy might help (PEREIRA 2019). When it comes to this variable (text/neutral syllable) more research is still needed, especially longitudinal studies.

[20] The sequence referred to is a distinctive feature of Edwin Gordon's theory comparing to other music pedagogues: it addresses how students learn music, which in turn affects how they are taught music and what they learn (GORDON 2003; 2012).

These two perspectives on songs, provide several ideas for future research. Acknowledging that the brain is dynamically and rapidly changing since birth and especially during the first years of life, it might be interesting to investigate if the exposure to songs with text and also with a neutral syllable would make a difference in the way text and melody interact in the brain throughout life. Assuming Edwin Gordon's claim, could the exposure to both type of songs shape the brain in different ways? Would it influence the way that neurophysiological, perceptual and cognitive processes overlap? Would more exposure to songs presented with a neutral syllable until the age of eighteen months influence the process of pitch encoding, for instance? What differences could be observed in children's brain when singing with words, singing with a neutral syllable (such as bah or la), and reciting the words? Is it possible that continued exposure to both ways of presenting a song influences the way children sing and will sing throughout their lives? Considering song teaching specifically, would those two strategies contribute to sharpen the auditory system (melody and text combined; or melody and text separately, and then combined)? Perhaps the answer to these questions might influence pedagogical practices, and also provide more insights on the relationship between music and language in the brain.

Despite the questions posed above, to present children with a variety of musical material is essential in developing their musical listening vocabulary, never excluding musical material from children's culture, of course. Variety should exist since birth and it will set the stage for the ability to discriminate, which is an important part of the learning process. Therefore, songs are not the only material to be used in music education. Apart from songs with text and with a neutral syllable, in different tonalities/modes and meters, the inclusion of chants presented with text (such as rhymes) and with a neutral syllable, on different meters, and instrumental music should be a regular practice. As Edwin GORDON (2003; 2012) states, the quality of informal music guidance until five years old is determinant in the way children use musical vocabulary when they are older, which will reflect on their singing, as well on other activities such as chanting, moving, creating, improvising.

Final remark

Although the existing models on music processing in the brain (KOELSCH 2011; PATEL 2011) are not linked to music education, connections could be made.[21] Current research using neuroimaging techniques or neuropsychological data has already provided valuable insights on the relationship between music and language in the brain as described along this article. Nonetheless, the future relationship between neuroscience and music pedagogy needs to be continually mapped and refined (COLLINS 2013). This relationship should work in both directions. On one hand, findings from the

[21] See contribution in HODGES - GRUHN (2012); RICHARD (2008); and an overview in RAUSCHER - GRUHN (2007).

neurosciences should be translated into music education practices as a way of improving songs' instruction and providing clues to understand the outcomes (songs' vocal performances); and, in addition, to make music teachers aware of how complex the 'singing in the brain' process is (perception and production). On the other hand, research in music education can provide interesting suggestions for future research in the cognitive neuropsychology field.

References

ALCOCK, Katherine J., Derick WADE, Philip ANSLOW, and Richard E. PASSINGHAM (2000), 'Pitch and Timing Abilities in Adult Left Hemisphere-dysphasic and Right-hemisphere Subjects', *Brain Language,* 75, pp. 47-65

BESSON, Mireille, Julie CHOBERT, and Céline MARIE (2011), 'Transfer between Music and Speech: Common Processing, Attention, and Memory', *Frontiers in Psychology,* 2/94, DOI: 10.3389/fpsyg.2011.00094

BOTHA, Rudolf (2009), 'On Musilanguage/"Hmmmmm" as an Evolutionary Precursor to Language', *Language and Communication,* 29/1, pp. 61-76

BROWN, Steven (2001), 'The "Musilanguage" Model of Music Evolution', in *The Origins of Music*, edited by Steven Brown, Björn Merker, and Nills L. Wallin (Cambridge, MA, MIT Press), pp. 271-300

CHUDLER, Eric H. (n.d.), *Brain Plasticity: What is it?* <https://faculty.washington.edu/chudler/plast.html> (accessed December 14, 2018)

CLARK, Camilla N., Hannah L. GOLDEN, Jason D. WARREN (2015), 'Acquired Amusia', *Handbook of Clinical Neurology,* 129, pp. 607-31

COLLINS, Anita (2013), 'Neuroscience Meets Music Education: Exploring the Implications of Neural Processing Models on Music Education Practice', *International Journal of Music Education*, 31/2, pp. 217-31

DALLA BELLA, Simone, Jean-François GIGUÈRE, Isabelle PERETZ (2009), 'Singing in Congenital Amusia', *Journal of the Acoustical Society of America,* 126/1, pp. 414-24

DARWIN, Charles (1871), *The Descendent of Man, and Selection in Relation to Sex* (London, John Murray)

DAVIDSON, Lyle (1985), 'Tonal Structures of Children's Early Songs', *Music Perception*, 2/3, pp. 361-74, DOI: 10.2307/40285304

DAVIDSON, Lyle, Patricia MCKERNON, and Howard GARDNER (1981), 'The Acquisition of Song: A developmental Approach', in *Documentary Report of the Ann Arbor Symposium: Applications of Psychology to the Teaching and Learning of Music*, edited by James Mason, Paul Lehman, and Wilbert McKeachie (Reston, VA, Music Educators National Conference), pp. 301-14

DEMOREST, Steven. M., and Peter Q. PFORDRESHER (2015), 'Singing Accuracy Development from K-Adult: A Comparative Study', *Music Perception*, 32/3, pp. 293-302, DOI: 10.1525/MP.2015.32.3.293

DEUTSCH, Diana (2013), 'The Processing of Pitch Combinations', in *The Psychology of Music*, edited by Diana Deutsch (London, Academic Press), pp. 249-325

DEUTSCH, Diana, T. HENTHORN, and R. LAPIDIS (2011), 'Illusory Transformation from Speech to Song', *The Journal of the Acoustical Society of America,* 129, pp. 2245-52

ETTLINGER, Marc, Elizabeth H. MARGULIS, and Patrick C. M. WONG (2011), 'Implicit Memory in Music and Language', *Frontiers in Psychology,* 2/211, DOI: 10.3389/fpsyg.2011.00211

FITCH, W. Tecumseh (2006), 'On the Biology and Evolution of Music', *Music Perception,* 24/1, pp. 85-8

GINSBORG, Jane, and John A. SLOBODA (2007), 'Singers' Recall for the Words and Melody of a New, Unaccompanied Song', *Psychology of Music*, 35/3, pp. 421-40, DOI: https://doi.org/10.1177/03057 35607072654

GOETZE, Mary (1985), 'Factors Affecting Accuracy in Children's Singing' (Ph.D. dissertation, The University of Colorado at Boulder)

GORDON, Edwin Elias (2003), *A Music Learning Theory for Newborn and Young Children* (Chicago, GIA Publications)

GORDON, Edwin Elias (2007), *Study Guide for Learning Sequences in Music: A Contemporary Music Learning Theory* (Chicago, GIA Publications)

GORDON, Edwin Elias (2012), *Learning Sequences in Music: A Contemporary Music Learning Theory* (Chicago, GIA Publications)

GORDON, Edwin Elias, and Christine JORDANOFF (1993), 'Kodály and Gordon: Same and Different', *Kodály Envoy*, XX/1, pp. 22-8

GORDON, Reyna L., Hilda M. FEHD, and Brice D. MCCANDLISS (2015), 'Does Music Training Enhance Literacy Skills? A Meta-Analysis', *Frontiers in Psychology,* 6/1777, pp. 1-16, DOI: 10.3389/fpsyg.2015.01777

GORDON, Reyna L., Cyrille L. MAGNE, and Edward W. LARGE (2011), 'EEG Correlates of Song Prosody: A New Look at the Relationship Between Linguistic and Musical Material', *Frontiers in Psychology,* 2/352, pp. 1-13, DOI: 10.3389/fpsyg.2011.00352

GUERRINI, Susan C. (2006), 'The Developing Singer: Comparing the Singing Accuracy of Elementary Students on Three Selected Vocal Tasks', *Bulletin of the Council for Research in Music Education,* 167, pp. 21-31

HANNON, Erin E., and Sandra E. TREHUB (2005), 'Metrical Categories in Infancy and Adulthood', *Psychological Science,* 16/1, pp. 48-55

HÉBERT, Sylvie, A. RACETTE, L. GAGNON, and Isabelle PERETZ (2003), 'Revisiting the Dissociation Between Singing and Speaking in Expressive Aphasia', *Brain,* 126, pp. 1838-50

HODGES, Donald A., and Wilfried GRUHN (2012), 'Implications of Neurosciences and Brain Research for Music Teaching and Learning, in *The Oxford Handbook of Music Education*, edited by G. E. McPherson and G. F. Welch (Oxford, Oxford University Press), vol. 1, pp. 205-23

HORNBACH, Christina M., and Cynthia C. TAGGART (2005), 'The Relationship between Developmental Tonal Aptitude and Singing Achievement among Kindergarten, First-, Second-, and Third-grade Students', *Journal of Research in Music Education*, 53/4, pp. 322- 31

JACOBI-KARNA, Kathleen Louise (1996), 'The Effects of the Inclusion of Text on the Singing Accuracy of Preschool Children' (Ph.D. dissertation, The University of Arizona)

JÄNCKE, Lutz (2012), 'The Relationship Between Music and Language', *Frontiers in Psychology*, 3/123, DOI: 10.3389/fpsyg.2012.00123

JENTSCHKE, Sebastian (2016), 'The Relationship between Music and Language', in *The Oxford Handbook of Music Psychology*, edited by Susan Hallam, Ian Cross, and Michael Thaut (Oxford, Oxford University Press), pp. 1-18

KIRBY, Simon (2012), 'Darwin's Musical Protolanguage: An Increasingly Compelling Picture', in *Language and Music as Cognitive Systems*, edited by Patrick Rebuschat, Martin Rohrmeier, John A. Hawkins, and Ian Cross (Oxford, Oxford University Press), pp. 98-102

KLINGER, Rita, Patricia Shehan CAMPBELL, and Thomas GOOLSBY (1998), 'Approaches to Children's Song Acquisition: Immersion and Phrase-by-Phrase', *Journal of Research in Music Education*, 46/1, pp. 24-34

KOELSCH, Stefan (2011), 'Toward a Neural Basis of Music Perception: A Review and Updated Model', *Frontiers in Psychology*, 2/110, DOI:10.3389/fpsyg.2011.00110

KOELSCH, Stefan, Thomas C. GUNTER, D. Yves CRAMON, Stefan ZYSSET, Gabriele LOHMANN, Angela D. FRIEDERICI (2002), 'Bach Speaks: A Cortical "Language-Network" Serves the Processing of Music', *Neuroimage,* 17/2, pp. 956-66

KOLB, Brian, and Robbin GIBB (2011), 'Brain Plasticity and Behavior in the Developing Brain', *Journal Academy Child Adolescence Psychiatry,* 20/4, pp. 265-76

KRUMHANSL, Carol (2000), 'Rhythm and Pitch in Music Cognition', *Psychological Bulletin,* 126/1, pp. 159-79

KRUMHANSL, Carol Lynne, and Paul IVERSON (1992), 'Perceptual Interactions Between Musical Pitch and Timbre', *Journal of Experimental Psychology: Human Perception and Performance,* 18/3, pp. 739-51

LARROUY-MAESTRI, Pauline (2018), '"I Know When I Hear It": On Listeners' Perception of Mistuning', *Music & Sience,* 1, DOI:10.1177/2059204318784582

LEHMANN, Christian, Lorenz WELKER, and Wulf SCHIEFENHÖVEL (2009), 'Towards an Ethology of Song: A Categorization of Musical Behavior', *Musica Scientiae,*13/2 *special issue*, pp. 321-38

LEIGHTON, Geraldine, and Alexandra LAMONT (2006), 'Exploring Children's Singing Development: Do Experiences in Early Schooling Help or Hinder?', *Music Education Research*, 8/3, pp. 311-30, DOI: 10.1080/14613800600957461

LERDAHL, Fred, Ray JACKENDOFF (1983), *A Generative Theory of Tonal Music* (Cambridge, MA, MIT Press)

LEVINOWITZ, Lily Muhler (1989), 'An Investigation of Preschool Children's Comparative Capability to Sing Songs with and without Words', *Bulletin of the Council for Research in Music Education*, 100, pp. 14-9

LEVITIN, Daniel (2006), *This Is Your Brain on Music* (London, Dutton Penguin)

MANG, Esther (2006), 'The Effects of Age, Gender and Language on Children's Singing Competency', *British Journal of Music Education*, 23/2, pp. 161-74, DOI: 10.1017/S0265051706006905

MCMULLEN, Erin, and Jenny R. SAFFRAN (2004), 'Music and Language: A Developmental Comparison', *Music Perception,* 21/3, pp. 289-311

MEHR, Samuel, A., Manvir SINGH, Dean KNOX, Daniel M. KETTER, Daniel PICKENS-JONES, et al. (2019), 'Universality and Diversity in Human Song', *Science*, 366/6468, DOI: https://doi.org/10.1126/science.aax0868>

MERRIL, Julia, and Pauline LARROUY-MAESTRI (2017), 'Vocal Features of Song and Speech: Insights from Schoenberg's Pierrot Lunaire', *Frontiers in Psychology*, 8/108, DOI: 10.3389/fpsyg.2017.01108

MITHEN, Steven (2005), *The Origins of Music, Language, Mind and Body* (London, Phoenix)

MOGHARBEL, Christliebe El, Grit SOMMER, Werner DEUTSCH, Marcus WENGLORZ, and Ingo LAUFS (2005-6), 'The Vocal Development of a Girl Who Sings But Does Not Speak', *Musica Scientiae,* 10/1 special issue, pp. 235-56

MOOG, Helmut (1976), 'The Development of Musical Experience in Children of Preschool Age', *Psychology of Music*, 4/2, pp. 38-45, DOI: 10.1177/030573567642005

NICHOLS, Brian E. (2016), 'Task-Based Variability in Children's Singing Accuracy', *Journal of Research in Music Education*, 64/3, pp. 309-21, DOI: 10.1177/0022429416666054

PALMER, Caroline, and Michael H. KELLY (1992), 'Linguistic Prosody and Musical Meter in Song', *Journal of Memory and Language,* 31/4, pp. 525-42

PATEL, Aniruddh D. (2008), *Music, Language, and the Brain* (Oxford, Oxford University Press)

PATEL, Aniruddh D. (2011), 'Music and the Brain: Three Links to Language', in *Oxford Handbook of Music Psychology*, edited by Susan Hallam, Ian Cross, and Michael Thaut (Oxford, Oxford University Press), pp. 208-16

PATEL, Aniruddh D. (2012), 'The OPERA Hypothesis: Assumptions and Clarifications', *Annals of the New York Academy of Sciences*, 1252, pp. 124-8, DOI: doi:10.1111/j.1749-6632.2011.06426.x

PEREIRA, Ana Isabel (2019), 'The Influence of Singing with Text and a Neutral Syllable on Portuguese Children's Vocal Performance, Song Recognition, and Use of Singing Voice' (PhD thesis, Universidade NOVA de Lisboa), available at <http://hdl.handle.net/10362/91276>

PERETZ, Isabelle (2012), 'Music, Language, and Modularity in Action', in *Language and Music as Cognitive Systems*, edited by Patrick Rebuschat, Martin, Rohrmeier, John A. Hawkins, and Ian Cross (Oxford, Oxford University Press), pp. 254-68

PERETZ, Isabelle, Elvira BRATTICO, Miika JÄRVENPÄÄ, Mari TERVANIEMI (2009), 'The Amusic Brain: In Tune, Out of Tune, and Unaware', *Brain,* 132, pp. 1277-86

PERETZ, Isabelle, Lise GAGNON, Sylvie HÉBERT, Jöel MACOIR (2004), 'Singing in the Brain: Insights from Cognitive Neuropsychology', *Music Perception,* 21/3, pp. 373-90

PERETZ, Isabelle, and Robert ZATORRE (eds.) (2003), *The Cognitive Neuroscience of Music* (Oxford, Oxford University Press)

PERETZ, Isabelle, Robert J. ZATORRE (2005), 'Brain Organization for Music Processing', *Annual Review of Psychology*, 56, pp. 89-114

PINKER, Steven (1997), *How the Mind Works* (New York, W. W. Norton & Company)

PROTO, Teresa, Paolo CANETTIERI, Gianluca VALENTI (eds.) (2015), *Text and Tune: On the Association of Music and Lyrics in Sung Verse* (Bern, Peter Lang)

RAUSCHER, Frances H., and Wilfried GRUHN (2007), *Neurosciences in Music Pedagogy* (New York, NY, Nova Science Publishers)

REBUSCHAT, Patrick, Martin ROHRMEIER, John A. HAWKINS, and Ian CROSS (eds.) (2012), *Language and Music as Cognitive Systems* (Oxford, Oxford University Press)

REIGADO, João, ROCHA, António, RODRIGUES, Helena (2011), 'Vocalizations of Infants (9–11 Months Old) in Response to Musical and Linguistic Stimuli', *International Journal of Music Education,* 29/3, pp. 241-55

REIGADO, João, and Helena RODRIGUES (2018), 'Vocalizations Produced in the Second Year of Life in Response to Speaking and Singing', *Psychology of Music*, 46/5, pp. 626-37

ROUSSEAU, Jean-Jacques (2012), 'Essai sur l'origine des langues', in *Collection complete des oeuvres*, Genève, 1780-9, 8/4, available at <http://www.rousseauonline.ch/Text/essai-sur-l-origine-des-langues.php>

RUTKOWSKI, Joanne (1993), 'The Use of Words Versus a Neutral Syllable for Evaluating Children's Use of Singing Voice', Paper presented at the Eastern Regional meeting of the Music Educators' National Conference (Springfield, MA)

RUTKOWSKI, Joanne (1998), 'The Nature of Children's Singing Voices: Characteristics and Assessment', in *Sharing the Voices: The Phenomenon of Singing*, edited by Brian A. Roberts (St. John's, NF, Canada, Memorial University Press), pp. 201-9

SAÏTO, Yoko, Kenji ISHII, Naoko SAKUMA, Keiichi KAWASAKI, Keiichi ODA, Hidehiro MIZUSAWA (2012), 'Neural Substrates for Semantic Memory of Familiar Songs: Is There an Interface Between Lyrics and Melodies?' *PLOS ONE*, 7/9, e46354 DOI: https://doi.org/10.1371/journal.pone.0046354

SAVAGE, Patrick E., Steve BROWN, Emi SAKAI, Thomas E. CURRIE (2015), 'Statistical Universals Reveal the Structures and Functions of Human Music', PNAS, 112/29, pp. 8987-92

SCHÖN, Daniele, Reyna GORDON, Aurélie CAMPAGNE, Cyrille MAGNE, Corine ASTÉSANO, Jean-Luc ANTON, and Mireille BESSON (2010), 'Similar Cerebral Networks in Language, Music and Song Perception', *NeuroImage*, 51/1, pp. 450-61

SLOBODA, John (1985), *The Musical Mind: The Cognitive Psychology of Music* (Oxford, Oxford University Press)

SMALE, M. J. (1987), 'An Investigation of Pitch Accuracy of Four- and Five-year-old Singers' (Ph.D. dissertation, University of Minnesota)

TAFURI, Johanne (2008), *Infant Musicality: New Research for Educators and Parents* (Farnham, Ashgate SEMPRE studies in the psychology of music)

TREHUB, Sandra (2003), 'Musical Predispositions in Infancy: An Update', in *The Cognitive Neuroscience of Music*, edited by Isabelle Peretz and Robert Zatorre (Oxford, Oxford University Press), pp. 3-20

WELCH, Graham Frederick (1998), 'Early Childhood Musical Development', *Research Studies in Music Education*, 11/1, pp. 27-41

WELCH, Graham Frederick, Desmond. C. SERGEANT, and Peta J. WHITE (1995-6), 'The Singing Competences of Five-year-old Developing Singers', *Bulletin of the Council for Research in Music Education*, 127, pp. 155-62

WELCH, Graham Frederik (2006), 'Singing and Vocal Development', in *The Child as Musician: A Handbook of Musical Development,* edited by Gary McPherson (Oxford, Oxford University Press), pp. 311-29

WELCH, Graham Fredrick, Evangelos HIMONIDES, Jo SAUNDERS, Ioulia PAPAGEORGI, Tiija RINTA, and Costanza PRETI (2011), 'Researching the First Year of the National Singing Programme Sing Up in England: An Initial Impact Evaluation', *Psychomusicology: Music, Mind and Brain*, 21/1-2, pp. 83-97, DOI: 10.1037/h0094006

WESSEL, David (1989), 'Timbre and Perceptual Organization of Musical Patterns', *The Journal of the Acoustical Society of America,* 86, S58, DOI: 10.1121/1.2027569

WISE, Karen J., and John A. SLOBODA (2008), 'Establishing an Empirical Profile of Self-defined "Tone Deafness": Perception, Singing Performance and Self-assessment', *Musicae Scientiae*, 12/1, pp. 3-26, DOI: 10.1177/102986490801200102

WONG, Patrick C. M., and Randy L. DIEHL (2002), 'How Can the Lyrics of a Song in a Tone Language be Understood?', *Psychology of Music*, 30/2, pp. 202-9

YUNG, Bell (1991), 'The Relationship of Text and Tune in Chinese Opera', in *Music, Language, Speech and Brain*, edited by Johan Sundberg, Lennart Nord, and Rolf Carlson (London, MacMillan), pp. 408-18

ZATORRE, Robert J., and Shari R. BAUM (2012), 'Musical Melody and Speech Intonation: Singing a Different Tune?', *PLOS Biology*, 10/7, pp. 1-6

Ana Isabel Pereira holds a PhD in Musical Sciences (teaching and psychology of music) from NOVA FCSH. She is a guest lecturer at ESELx/IPL and at NOVA FCSH. She teaches music to children between three and ten years old. Her interests include the study of musical learning and development, and the evaluation of children's vocal performance. She graduated in environmental engineering from the IST/UL.

Helena Rodrigues is an Assistant Professor at NOVA FCSH. She is one of the founders of the Laboratory for Music and Communication in Infancy (LAMCI) at CESEM and Companhia de Música Teatral. She was a Researcher Fellow at the Royal Flemish Academy of Belgium for Science and the Arts. She coordinated the project Opus Tutti and GermInArte, both dedicated to music in infancy and early childhood.

Recebido em | *Received* 15/01/2019
Aceite em | *Accepted* 18/02/2021