5

Speech processing in congenital amusia: A review of the literature and future directions

Manuela M. Marin
University of Innsbruck, Department of Psychology

ABSTRACT Human music, like language, is a cultural universal whose origins remain elusive. Some individuals have a congenital impairment in music processing and singing, which might be biologically rooted. Congenital amusia, a neurodevelopmental disorder, has been studied for around 15 years by only a few research groups worldwide. Although amusia was originally thought of as a music-specific disorder, it was demonstrated relatively quickly that it also affects the perception of speech intonation in laboratory conditions. This finding has spurred a fruitful research program, which has not only contributed to the ongoing debate about modularity by investigating pitch processing in music and speech, but which has also targeted various other aspects of linguistic processing. After a comprehensive introduction into the currently known characteristics of amusia, empirical studies on speech processing in non-tone and tone language speakers diagnosed with amusia will be reviewed. Research on English- and French-speaking amusics has primarily focused on the processing of phonology, linguistic and affective prosody, as well as on verbal memory. Amusics are impaired in several aspects of phonological processing, in differentiating between statements and questions, and further, in perceiving emotional tone. In the last few years, a growing body of research involving tone language speakers has reported impairments in processing of lexical tone and speech intonation among amusics. Hence, research on Mandarin and Cantonese speakers has provided evidence for the notion that amusia is not a disorder specific to non-tone language speakers, and further, it has also increased our knowledge about the core deficits of amusia. Future research will need to replicate current findings in speakers of languages other than those already studied and also clarify how this auditory disorder is linked to other learning disorders.

KEYWORDS congenital amusia, modularity, neurodevelopmental disorder, pitch processing, speech processing

1 INTRODUCTION

> Music is the divine way to tell beautiful, poetic things to the heart.
> — Pablo Casals

Humans are social creatures who enjoy communicating by several means, and information sharing using *speech* (i.e., the vocalized component of language) is the most obvious way of communication experienced in daily life. Yet human auditory communication includes more than speech, namely vocalizations and music, notably song and instrumental music, which all use structured sound sequences propagating in time. Over the last decades, a considerable amount of basic research has dealt with the comparison between language and music, following either a theoretical (Lerdahl & Jackendoff, 1985; Martin, 1972), empirical (Gorin, Kowialiewski, & Majerus, 2016; Tillmann, 2012) or computational (Rohrmeier, Zuidema, Wiggins, & Scharff, 2015) stance of studying cognitive processes. Due to the rise of the cognitive neurosciences, these efforts were complemented by investigating the neuronal basis underlying language and music processing (Besson, Barbaroux, & Dittinger, 2017; Glushko, Steinhauer, DePriest, & Koelsch, 2016; Patel, 2003; Steinbeis & Koelsch, 2008). Other current approaches include the comparison of developmental processes regarding speech and music in infants and young children (Brandt, Gebrian, & Slevc, 2012; Marin & Peltzer-Karpf, 2009; McMullen & Saffran, 2004; Trehub, Trainor, & Unyk, 1993) as well as the link between speech and music processing in developmental disorders (Don, Schellenberg, & Rourke, 1999; Flaugnacco et al., 2014; Lee, Sie, Chen, & Cheng, 2015; Vaiouli & Andreou, 2018). Within these areas of investigation it is now generally accepted that language and music share many attributes, cognitive resources and underlying neural networks aside from their distinctive natures (Asaridou & McQueen, 2013; Morrill, McAuley, Dilley, & Hambrick, 2015).

The cognitive ability to process and produce both complex hierarchically-structured speech and music is frequently found on a list of traits explaining humans' unique standing among animals (Hauser & Watumull, 2017; Martins, Gingras, Puig-Waldmueller, & Fitch, 2017). Music, like language, is a universal social phenomenon (Blacking, 1973) whose origins remain elusive (Merker, Morley, & Zuidema, 2015) because they are less clearly rooted in biology than is the case for language, for which the selective advantage is apparent (Cross & Morley, 2009). However, we have archaeological evidence that music played a role in human cultures already 40,000 years ago (Higham et al., 2012). A growing line of research follows a comparative approach across species to illuminate the biological roots of music: *zoomusicology* studies music-like aspects of sound communication in animals (Doolittle & Gingras, 2015). Cross-species research may not only help elucidate the nature and origins of music (Fitch, 2006; Honing, ten Cate, Peretz, & Trehub, 2015) and language (ten Cate, 2017), but as Patel (2017) argued, studying music across species may also help better understand the cognitive mechanisms of speech. More recently, building on progress in the biological sciences, researchers have shown growing interest in the genetic basis of music, especially in explaining individual differences in musicality (Gingras, Honing, Peretz, Trainor, & Fisher, 2015; Mariath et al., 2017). This relatively unexplored avenue of research is gaining impetus from more advanced research on the underlying genetics of the human capacity for language (Fisher, 2017).

Genetics and genetic abnormality also play a role in our understanding of *neurodevelopmental disorders*, a term referring to impairments of the growth and development of the central nervous system. *Congenital amusia* (hereafter *amusia*) is a hereditary disorder (Kalmus & Fry, 1980; Peretz, Cummings, & Dubé, 2007; Pfeifer & Hamann, 2018) mostly affecting pitch processing in music (Ayotte, Peretz, & Hyde, 2002; Foxton, Dean, Gee, Peretz, & Griffiths, 2004). This auditory disorder was first described by a team led by Isabelle Peretz (2002) in a patient called Monica, who was impaired in musical abilities, such as singing and melodic discrimination and recognition. Monica had no history of brain lesion, no hearing loss, normal cognitive abilities and was exposed to music from an early age onwards. This influential discovery has led to an intensive research program over the past 15 years, which has primarily aimed at identifying the core deficit of amusia, its neuronal basis, and how this disorder affects speech processing. Amusia is currently being studied by a half-dozen research groups worldwide.

This chapter provides a review of the literature on amusia and speech processing (perception and production) within the framework of music and language comparisons. After briefly introducing the reader to current research on pitch processing in amusia, I will present a section on speech processing in *non-tone language speakers*, followed by a section on amusics who are *tone language speakers* (i.e., they use pitch to communicate lexical or grammatical meaning). This chapter will highlight influential studies leading to the notion that amusia is not a music-specific disorder, introduce the reader to common concepts and research paradigms, and also include ideas for future directions.

2 WHAT IS CONGENITAL AMUSIA?

Musical tones are characterized by perceptual attributes such as pitch, loudness, duration and timbre. *Pitch* is the perceptual property of a sound that is mostly related to the *frequency* of a sound, describing whether a sound is low or high. Frequency can be understood as the speed of a vibration and refers to the number of wave cycles that occur in one second. Musical tones varying in pitch can be assigned positions on a *musical scale* (i.e., a set of notes ordered by pitch). A sequence of tones taken from a musical scale can form a *melody*, and consecutive tones form *melodic intervals* that are defined by the ratio of their frequencies. This ratio determines the perceived pitch difference. In Western music, a *semitone* forms the smallest interval between adjacent tones on a diatonic scale. Several musical tones in a row can form a *melodic contour*, a pitch pattern that describes the global nature of a tone sequence over time (e.g., rising or falling). Since pitch is among the fundamental building blocks of music (McDermott & Hauser, 2005), it becomes obvious that being impaired in pitch processing has far-reaching effects for the appreciation of music.

Congenital amusia is commonly diagnosed with the *Montreal Battery for the Evaluation of Amusia* (MBEA; Peretz, Champod, & Hyde, 2003), which assesses different aspects of musical pitch and time processing. The MBEA comprises six subtests (contour, interval, scale, rhythm, meter and memory), involving short melodies following the rules of Western tonal music. After having listened to pairs of melodies in the pitch-related subtests, participants perform a same/different discrimination task based on detecting possible manipulations of the second melody such as violations of contour, interval among other things. Thus, this test probes important music perception skills. In general, amusics perform significantly worse than healthy controls on the three pitch-related subtests of the MBEA but not necessarily on the meter and rhythm subtests (Peretz & Coltheart, 2003), which has led some researchers to diagnose amusia mostly by focusing on these pitch-related subtests (e.g., Liu, Patel, Fourcin, & Stewart, 2010). People diagnosed with amusia should not be mistaken for people in the population whose music perception abilities are normal but who cannot sing in tune (Pfordresher & Brown, 2007). Importantly, the first studies on amusia already hinted at a *fine-grained pitch processing deficit*, rather than at a specific music processing deficit. For example, amusics have difficulties detecting pitch changes that are smaller than two semitones (Hyde & Peretz, 2004), and they are less sensitive to the perception of pitch direction (i.e., pitches moving up or down; Foxton et al., 2004).

Besides the use of the MBEA, researchers have employed psychoacoustic pitch perception tests to shed further light on the nature of the underlying low-level pitch processing deficit. A recent meta-analysis (Vuvan, Nunes-Silva, & Peretz, 2015) included a wide range of tests comprising acoustic tasks, such as *pure tone pitch change detection* (Does one tone in a set differ from the others?) and *identification* (Which of the tones is different from the others?), as well as *pure tone pitch change discrimination* (Are two sets of tones the same or different?) and *identification* (Is the pitch change up or down?). Vuvan et al. also evaluated the performance of amusics in non-acoustic tasks (including tone and non-tone language speakers), mostly those involving the perception of pitch as expressed by *speech intonation* (i.e., the perception of the pattern or melody created by pitch changes in connected speech). Such tasks can be labeled as either *speech intonation identification* (Is the sentence a statement or a question?) or *discrimination* (Is the intonation of the second sentence the same as the first?). The authors analyzed 42 studies with data stemming from five research groups who used the MBEA to diagnose amusia. Controls performed much better than amusics in the respective pitch tasks. Moreover, the effect sizes could be predicted by the size of the pitch change used in the task (i.e., the larger the pitch change, the smaller the performance gap). Furthermore, the performance gap was smaller in tone language speakers, and interestingly, no effect was found when comparing studies using either tone or speech stimuli. The latter finding is of particular importance because it suggests that amusia is a pitch processing deficit that equally impacts musical tone *and* speech intonation processing. Finally, the authors reported a moderate to strong correlation between the performance on the MBEA and the respective pitch tasks. Although this meta-analysis provided several interesting insights into the underlying pitch processing deficit of amusia, it is necessary to exercise caution when interpreting the results: most research groups have a limited pool of participants due to the very low prevalence of amusia, which lies between 1.5 and 4.2%, depending on the way amusia is diagnosed (Henry & McAuley, 2010; Peretz, 2016). Thus, there is likely a large overlap of tested people in the individual studies coming from the same research group. This exemplifies the need to establish new pools of amusics and the collaboration between research groups to increase sample size and statistical power.

In the last few years, several other elaborate studies have further examined the pitch processing deficit underlying amusia. Most sounds in nature are not pure tones (sine tones) but *harmonic complex tones*, consisting of several *harmonics* (partials) which are sine waves and multiples of a *fundamental frequency* (F0). The lower *resolved* harmonics of such a complex tone largely determine the sensation of pitch and can even be heard out individually, whereas the higher harmonics do not strongly contribute to pitch sensation (Bernstein & Oxenham, 2003). Cousineau, Oxenham and Peretz (2015) demonstrated that amusics are impaired in pitch perception only when the resolved harmonics are not present in a complex tone, whereas pitch discrimination did not differ from controls when the tones only contained higher harmonics. In addition, amusics had normal periphery auditory perception, with intact basic temporal, spectral and intensity coding. In another series of psychoacoustic tests, Whiteford and Oxenham (2017) found that amusics were impaired in pitch processing of very high frequencies, which are not commonly used in musical communication. They also showed deficits in the processing of frequency modulations (see also Albouy, Cousineau, Caclin, Tillmann, & Peretz, 2016), and most strikingly in the processing of amplitude modulations, which refers to a deficit in the perception of loudness changes. Their data also provided more support for the view that amusics have a pitch impairment that is independent from a concurrent short-term memory problem related to pitch processing (Tillmann, Leveque, Fornoni, Albouy, & Caclin, 2016; for counterevidence see Jiang et al., 2013).

Another line of research has focused on the hypothesis that amusia is particularly associated with the inability to identify the direction of pitch changes (i.e., contour processing), which may be independent from poor pitch memory (see also *Melodic Contour Deafness Hypothesis*, Patel, Foxton, & Griffiths, 2005; for counterevidence see Vuvan et al., 2015). Lu, Sun, Ho and Thompson (2017) developed a novel task (*Self-paced Audio-visual Contour Task*) that probes contour processing while keeping memory load minimal. Visual and melodic contours were simultaneously presented and participants were asked to judge the congruency between the two contours. Specifically, after the presentation of each individual tone forming the contour, the participants had to judge whether the tone was lower or higher than the preceding tone. The authors found strong support for the notion that amusics have a contour processing deficit in addition to a pitch discrimination and short-term memory deficit.

Zhou, Liu, Jing and Jiang (2017) studied neuronal responses during melodic contour processing in a cross-modal priming paradigm. They showed that for Mandarin-speaking amusics musical meaning conveyed by changes in pitch direction (in this case a scale moving upwards or downwards) did not evoke an N400 effect in semantically incongruent melody-image pairs, whereas an N400 effect was observed in healthy controls. The N400 is a component of the event-related potential measured by means of electroencephalography (EEG) and sensitive to the processing of meaning (Kutas & Hillyard, 1980). Zhou et al.'s findings suggest that the amusic brain does not process changes in pitch direction as a semantically meaningful cue, which was shown for both implicit and explicit tasks. However, using more complex music as primes (stimuli that use other acoustic cues than pitch to convey meaning) revealed that amusics showed an N400 effect as controls did, indicating that they are able to process musical meaning based on other cues than pitch direction. Thus, recent empirical evidence stemming from different research groups clearly indicates that amusia, among other things, may be characterized by a *pitch direction processing deficit*.

A number of researchers have attempted to identify the neuronal underpinnings of amusia, especially in the context of pitch processing. Peretz (2016), reviewing the neurobiology of amusia, is a proponent of the view that "[t]he core deficit of congenital amusia is characterized by a lack of awareness of acquired musical pitch knowledge. The amusic brain shows abnormalities in neural transmission between the auditory cortex and the inferior frontal gyrus in the right cerebral hemisphere" (p. 857). Her opinion is mostly based on research involving Western language speakers and suggests that the amusic brain responds normally to small pitch changes in music, but that conscious detection is impaired. In other words, as Peretz puts it, the amusic brain is "[i]n tune, but unaware" (2016, p. 859).

Several brain imaging studies (for details see Peretz, 2016) have indicated that the amusic brain processes auditory information normally up to the *primary auditory cortex* (A1) and the *superior temporal gyrus* (STG), but it appears as if the top-down feedback from the *inferior frontal gyrus* (IFG) to the STG is not happening to the extent it does in healthy controls. Therefore, it is likely that the right fronto-temporal network is not intact in amusics due to a reduced *arcuate fasciculus* (AF) connectivity between these regions (Figure 1). Peretz (2016) further suggests that amusia, together with other cognitive learning disorders showing a disconnection between the perceptual and frontal cortices (e.g., *dyslexia*, a learning difficulty affecting reading ability, and *prosopagnosia*, an inability to recognize faces), may have a common origin and that a genetic variant may be shared between these disorders.

Some studies have provided evidence against the frontal-temporal pathway hypothesis and reported impairments regarding the auditory brainstem response (Lehmann, Skoe, Moreau, Peretz, & Kraus, 2015) or identified issues regarding the tractography algorithm used to de-

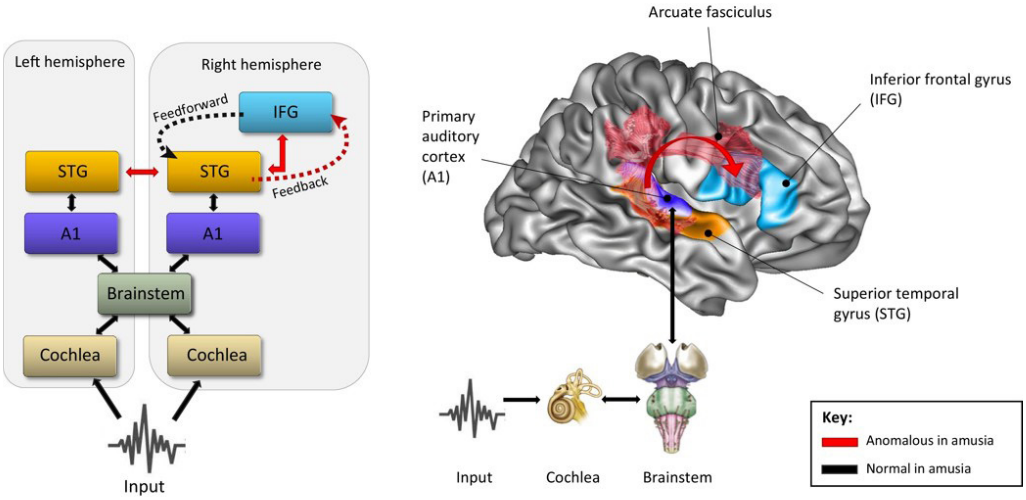

Figure 1. A schematic (left panel) and anatomical representation (right panel) of the anomalous recurrent processing in the right frontotemporal network in amusics (reprinted from *Trends in Cognitive Sciences*, Vol. 20, Peretz, Neurobiology of congenital amusia, p. 863, Copyright (2016), with permission from Elsevier).

tect the AF (Chen et al., 2015). Nevertheless, some further support for intact processing in the auditory cortex (Norman-Haignere et al., 2016) has accumulated using fMRI (for counterevidence using magnetoencephalography see Albouy et al., 2013), alongside recent reports that the amusic brain is characterized by a global connectivity deficiency of the white matter network, going clearly beyond a sole deficiency of the frontal-temporal pathway (Wang, Zhang, Wan, & Peng, 2017; Zhao et al., 2016). Moreover, it may be possible that the amusic brain differs between tone and non-tone language speakers, as recently suggested by Zhang, Peng, Shao and Wang (2017). With the rise of more advanced neuroscientific methods, future research will need to determine the precise neural processing deficiencies of amusia, especially how perceptual and memory processes are linked and how they are affected by missing feedback from other brain areas. This research will also need to integrate current findings suggesting that amusia is more than a fine-grained pitch processing deficit.

A vital question with respect to neurodevelopmental disorders is whether effective training programs can be developed to improve the impairment. Liu, Jiang, Francart, Chan and Wong (2017) have recently provided the first evidence for training effects on pitch direction perception. Twenty Chinese amusics were split into training and control groups. Over two weeks, the training group took part in a 10-session program which was individually tailored to improve the perception of pitch direction identification in speech and music. Amusics in the training group performed better in the pitch direction identification task after the training, but this improvement did not transfer to musical abilities. Nevertheless, this report will motivate further training studies on the *plasticity* of the amusic brain, especially since previous research has failed to show clear training effects (Anderson, Himonides, Wise, Welch, & Stewart, 2012; Goulet, Moreau, Robitaille, & Peretz, 2012; Wilbiks, Vuvan, Girard, Peretz, & Russo, 2016).

3 SPEECH PROCESSING IN NON-TONE LANGUAGE SPEAKERS WITH CONGENITAL AMUSIA

In the following section, I will review studies that investigated amusia in the context of phonology, linguistic prosody, affective prosody, syntax and verbal short-term memory. These studies go beyond the mere study of pitch processing in music and speech and target possible effects of amusia on various aspects of speech processing from a broader perspective.

3.1 Phonology

Phonological processing involves the use of sounds to process written and spoken language, comprising *phonological awareness*, *phonological working memory* and *phonological retrieval*. Phonological awareness refers to the awareness of sounds and oral language patterns within the words of one's language. More specifically, it alludes to the ability to work with syllables, to blend and segment *phonemes* in words (a phoneme, in a particular language, is one of the smallest units of speech making one word different from another). Phonological working memory refers to the short-term maintenance of speech sounds. Phonological retrieval, frequently assessed by *rapid automatized naming tests*, refers to the ability to recall phonemes associated with specific *graphemes* (the smallest meaningful contrastive units in a writing system; Wagner & Torgesen, 1987).

In healthy individuals, research has shown that phonological processing is positively associated with musical ability (Anvari, Trainor, Woodside, & Levy, 2002; Dawson, Aalto, Simko, Vainio, & Tervaniemi, 2017; Degé & Schwarzer, 2016; Lamb & Gregory, 1993; Linnavalli, Putkinen, Lipsanen, Huotilainen, & Tervaniemi, 2018; Marin, 2009; Moreno, Friesen, & Bialystok, 2011; Slevc & Miyake, 2006). The common explanation is that phonological processing relies on the effective processing of basic, low-level auditory features, namely changes in pitch and temporal variations. These changes in the auditory signal are shared attributes of speech and music, which are thus likely to share networks of auditory processing in the brain.

Remarkably few studies have been designed to test phonological processing in amusia. Jones, Lucker, Zalewski, Brewer and Drayna (2009) were the first to study whether a congenital pitch deficit observed for musical processing is associated with decreased phonological processing. The authors assessed *tune-deafness* (one of the precursor terms to amusia) by using the *Distorted Tunes Test* (Kalmus & Fry, 1980), which requires participants to indicate whether a popular melody sounds out-of-tune or not. Phonological processing was assessed with the *Comprehensive Test of Phonological Processing* (Lennon & Slesinski, 2003) and two other tests on auditory word comprehension and syllable segmentation, which in total led to five measures of phonological awareness (auditory word discrimination, elision, phoneme reversal, blending words and blending non-words) and three measures of phoneme awareness (segmenting words, segmenting non-words and syllable segmentation). Sixty-nine English-speaking participants took part in the main experiment, 35 of which were tune-deaf. In general, Jones et al.'s results indicated that tune-deaf participants performed significantly worse than controls in all eight tests probing phonological processing.

Sun, Lu, Ho and Thompson (2017), building upon the work of Jones et al. (2009), tested a group of English-speakers diagnosed with amusia using the MBEA. The authors assessed how pitch and rhythm processing in music are related to phonological processing in speech. They employed a pure-tone pitch discrimination task, a rhythm discrimination task as well as four subtests of the Comprehensive Test of Phonological Processing (elision, memory for digits, non-word repetition and rapid-digit naming). Twenty amusics and twenty controls with intact reading abilities performed similarly well on all four phonological subtests. However, the authors identified a subgroup of amusics with particularly high pitch discrimination thresholds. Further analysis considering the two subgroups of amusics (i.e., with normal or abnormal pitch discrimination thresholds) and controls revealed group differences only for the elision test, which assesses the ability to remove phonological segments from spoken words to form new words. Amusics with abnormal pitch discrimination thresholds performed worse than amusics with normal pitch discrimination thresholds and controls (the latter two groups performed similarly well). IQ and reading ability could not explain why amusics with abnormal pitch discrimination performed worse. Regarding rhythm discrimination ability and its relation to phonological processing the authors reported a moderate correlation; however, further statistical analyses revealed that pitch discrimination ability was a stronger predictor of the performance in the elision test than rhythm discrimination ability and phonological short-term memory. Taken together, their results suggest that pitch discrimination ability, and not performance on the MBEA, is a reliable predictor of phonological awareness. Early stage deficits in auditory processing of pitch may thus have a negative impact on higher level processes of discriminating, segmenting and blending phonemes.

The studies by Jones et al. (2009) and Sun et al. (2017) present a first step towards our understanding of how impairments in basic pitch processing impact phonological processing. A comparison of results across the two studies is difficult because they differ in their way of assessing the ability to process musical pitch, and Jones et al. did not determine any pitch thresholds. Unfortunately, none of the studies employed a pitch direction discrimination task; therefore, future studies may include a wider spectrum of tests to assess general and musical pitch processing more comprehensively (Lu et al., 2017). At the moment, the choice of a specific pitch test is often unexplained and seems to depend on the research groups' views on the nature of amusia. However, an open and broad approach towards measuring pitch processing would shed light on the question of whether the ability to process musical and/or general pitch is linked to phonological awareness, which would have important implications for issues surrounding the relationship between music and speech processing.

If we identified the key variable causing the observed deficit in phonological awareness in amusia, we might also be able to better understand the mechanisms underlying another developmental disorder, namely those of dyslexia (see also Turker in this volume). It is generally agreed that deficits in phonological awareness and poor reading abilities are related (Gillon, 2018; Ozernov-Palchik & Gaab, 2016; Ramus et al., 2003), but it is still unclear how a deficit in phonological awareness may eventually develop into a reading disorder (Raschle, Stering, Meissner, & Gaab, 2014). For example, deficits in *rapid auditory processing* have been described in pre-reading children with a familial risk for dyslexia (Raschle et al., 2014). Interestingly, impaired encoding of rapid pitch information has also been identified as a possible cause of amusia (Albouy et al., 2016). Albouy et al. argue that developmental disorders such as dyslexia, specific language impairment and congenital amusia may be characterized by a low-level sensory dysfunction. Clearly, future research will need to extend this growing area of investigation by taking on a broader view on auditory development and its disorders.

3.2 **Linguistic prosody**

In human communication, it is not only important *what* people say but also *how* they say it. The term *prosody* refers to the melodic and rhythmical aspects of speech, sometimes being described as "the music of everyday speech" (Wennerstrom, 2001, p. 1). Prosody is derived from ancient Greek *προσῳδία*, meaning "song sung to music, variation in pitch of the speaking voice, or pronunciation of a syllable on a certain pitch" (Oxford English Dictionary, 2007). In linguistic terms, prosody is the study of those aspects of speech that are above that of the individual phoneme (i.e., *suprasegmental*) and those that may also refer to sequences of words. Prosody is characterized by changes in vocal pitch, loudness, tempo and voice quality (Crystal & Quirk, 1964; Fonagy, 1978). *Linguistic prosody* is specific to a language and can signal phrase boundaries, put emphasis on individual parts of a word (*word stress*) or words within a sentence (*sentence focus*), or signal whether a sentence is a statement, question or command. *Affective prosody* can be partly understood cross-culturally

(Sauter, Eisner, Ekman, & Scott, 2010) and signals emotions and attitudes of the speaker in normal speech as well as *affect bursts* without semantic meaning (see section 3.3). In a meta-analysis, Belyk and Brown (2014) reported that although linguistic and affective prosody processing largely overlap in the right hemisphere of the healthy brain, affective prosody more likely activates Brodmann area 47 and linguistic prosody the ventral part of area 44.

Studying linguistic prosody perception in amusia is an interesting topic because amusics generally do not report any speech impairments in their daily lives. However, the first studies on music-speech comparisons in amusia already concerned themselves with speech intonation perception, probably motivated by the idea that melody processing in music may be related to the processing of speech melody. A considerable body of research has assessed how amusics perceive and produce linguistic prosody (Vuvan et al., 2015), which has largely shaped our understanding of the general pitch processing deficit underlying amusia (see section 2). Therefore, it is worth reviewing some of the most influential studies on this topic to demonstrate how the field has moved away from considering amusia as a domain-specific disorder (Peretz & Coltheart, 2003) by questioning initial results on speech perception and developing complex experimental tasks.

Ayotte et al. (2002) observed that amusics were *not* impaired in the processing of intonation patterns in sentences but in the processing of discrete tone sequences based on the intonation patterns of the same sentences, suggesting that amusia is a domain-specific disorder. Arguing that the pitch contours were not exactly matched across sentences and tone sequences, which may explain this dissociation, Patel et al. (2005) created discrete and gliding pitch-analogs of the intonation contour present in the sentences. Employing a discrimination task, they used two types of sentence-pairs (statement-question pairs and focus-shift pairs) that had the same semantic content but different intonation contours. Amusics performed worse on the discrimination of pitch-analogs mimicking the intonation patterns of the sentences than on the original sentences, corroborating findings by Ayotte et al. (2002).

Although Patel et al. (2005) observed no deficit in speech intonation processing among amusics, Lochy et al. (2004) had already noted that some French-speaking amusics may show an impairment. Motivated by this observation, Patel, Wong, Foxton, Lochy and Peretz (2008) studied British and French-Canadian amusics and investigated the perception of statement-question pairs in more detail. Specifically, they examined whether deficits in the processing of the final pitch change (measured as *glide size* and *glide rate*), which differentiate statements from questions, can explain individual differences in amusics' performance. They found that around 30% of the tested amusics could not differentiate between a statement and a question characterized by a falling or rising pitch. The same amusics, however, were able to discriminate discrete pitch analogs based on the intonation contours of sentences. In addition, these amusics were not impaired when asked to detect changes in emphasis placed on individual words within a sentence. Taken together, these results indicate that the underlying problem of intonation processing in speech is related to an impairment of processing pitch direction.

Building upon this hypothesis, Liu et al. (2010) employed a series of five tasks probing discrimination, identification and imitation of statements and questions that differed in the pitch direction of the final word (i.e., statement-question discrimination in natural speech and gliding tones, nonsense speech, statement-question identification and imitation). They also conducted two psychoacoustic tasks which determined thresholds for pitch change detection and pitch direction discrimination. The authors observed that amusics performed worse than controls in all speech intonation tasks, for which smaller pitch contrasts were used than in Patel et al. (2008). Moreover, amusics' pitch direction discrimination thresholds were negatively correlated with their performance in most tasks. Altogether, this comprehensive study provided strong evidence that amusia is *not* a domain-specific disorder and that brain networks for music and speech processing may overlap in amusia (see section 2).

Trying to understand why some studies reported no deficit in speech intonation processing (Ayotte et al., 2002; Patel et al., 2005; Peretz et al., 2002) whereas others did (Liu et al., 2010; Patel et al., 2008), Hutchins, Gosselin and Peretz (2010) hypothesized that the reason for the discrepancy may lie in the fact that speech uses larger pitch differences to communicate relevant information than music, which would explain why amusics' impairment is more salient with musical stimuli. In other words, fine-grained differences in pitch play a larger role in music than in speech. Hutchins et al. used sentences and tone sequences which followed the pitch contour of the sentences, and they manipulated the pitch of the final word/tone in 11 steps so that they represented a pitch continuum, indicating either a statement or a question at its endpoints. Participants judged whether a sentence was a statement or question or whether a tone sequence ended on a rising or falling contour. For both speech and music, amusics showed lower accuracy and consistency in their categorizations, with a similar performance for both stimulus types (see also Vuvan et al., 2015). Furthermore, the data showed that amusics could correctly identify unambiguous stimuli representing statements or questions, which may explain why amusics usually do not report any problems in their daily lives.

Another question of interest is whether discrimination ability is coupled with *imitation ability*, or whether these aspects of speech intonation and pitch processing are decoupled (Liu et al., 2010; Loui, Guenther, Mathys, & Schlaug, 2008). This question is of relevance because pitch processing may involve a dual-route in the brain: one pathway involving vocal production and another pathway involving representations that are necessary to perform conscious perceptual tasks. Hutchins and Peretz (2012) employed local pitch variations in sentences (i.e., pitch shifts within a sentence) and tested whether amusics and controls can perceive the similarity or difference of these variations when comparing pairs of sentences (discrimination task), and further, whether they can imitate pitch in speech after listening to the same sentences (imitation task). Their results supported the view that amusics are impaired in perceiving pitch changes in speech but have intact imitation abilities. These findings suggest that amusics are able to make use of a kind of motor-system based pitch representation when imitating speech intonation. As the authors point out, a decoupling of perception and production has also been reported for general visual (Goodale & Milner, 1992) as well as for speech processing (Hickok & Poeppel, 2004), which makes the existence of similar separate neural pathways for pitch processing also very likely.

In a follow-up study, Hutchins and Peretz (2013) examined perception and imitation of vocal pitch shifts in singing and speech within the context of the *dual-route model of pitch*. The *pitch shift reflex* "is a sophisticated system that produces a compensatory response in voice F0 [fundamental frequency] that is opposite in direction to a change in voice pitch feedback (pitch-shift stimulus)" (Bauer & Larson, 2003, p.

1048). The underlying psychological process is explained by the parity of motor and auditory representations of pitch, assuming that auditory input is transformed into a vocal-motor code representation. If a person sings or speaks and the voice pitch feedback is altered in pitch during the vocalization, the person makes adjustments in pitch even without conscious awareness of the pitch shift in the feedback. Hutchins and Peretz applied a self-matching method in which amusics and controls were first recorded while producing sung (three different notes) and spoken (three one-syllable words) materials. Then participants were asked to imitate their own recordings, including pitch shift trials in which imitations were either shifted up or down (25 or 200 cents) or left unaltered as a control. The results showed that amusics were generally less accurate than controls when imitating sung or spoken pitch prior to the onset of pitch shifts. Amusics performed particularly badly in the singing condition. In controls, the performance for sung and spoken materials was similar. Interestingly, only a few amusics did not show a normal pitch shift response, and nearly half showed a normal pitch shift response to small and large shifts. Their performance was predicted by vocal pitch matching ability and not by pitch perception ability. The authors conclude that their data provide support for the dual-route model of pitch, and further, that pitch perception and production are dissociated in amusia (see section 2).

To summarize, research on amusia and linguistic prosody has not only broadened our understanding of the underlying pitch processing deficit in amusia, but has also led to a research program that focuses on speech perception alongside production, following an embodied view of cognition. Research on amusia has thereby contributed to our understanding of pitch processing in healthy individuals. Future studies may address the issue of whether amusics show deficits in linguistics prosody processing that extend beyond speech intonation processing, such as speech rhythm processing, which may be worth testing in those amusics who are also impaired in musical rhythm processing.

3.3 Affective prosody

Affective or *emotional prosody* refers to the non-verbal expression of emotion in speech. By means of varying features, such as pitch contour, intensity, duration and speech rate (Bachorowski, 1999; Bänziger, Patel, & Scherer, 2014), humans manage to communicate a wide range of emotions in social interactions. For instance, affective prosody plays a crucial role in infant-directed speech and has a positive impact on linguistic outcomes during language development (Spinelli, Fasolo, & Mesman, 2017). Affective vocal communication has also received attention in relation to the origins of language (see also Filippi & Gingras in this volume). None other than Darwin suggested in *The Descent of Man* (1871) that proto-musical emotional expressions, similar to singing, were a precursor of language (for a detailed discussion see Fitch, 2010), whereas Spencer (1911) argued the contrary, namely that affective prosody was the precursor of music. Brown (2017) has recently presented his *Musilanguage 2.0 model*, taking on an alternative perspective by putting affective and linguistic prosody at the core of his model and arguing that there were two sequential precursor stages, shared by language and music, prior to a bifurcation into two homologous communication systems at a later evolutionary stage.

Both music and language share acoustic cues in the communication of emotion (Juslin & Laukka, 2003), and overlapping brain networks in the processing of vocal and musical emotions have frequently been described in healthy individuals (Frühholz, Trost, & Kotz, 2016). Moreover, musical training and affective prosody perception are positively correlated (Lima & Castro, 2011). Motivated by these findings, several studies have addressed the issue of affective prosody perception in amusia, hypothesizing that if parts of the brain networks are shared between music and language, a pitch processing impairment may lead to deficits in affective prosody processing.

Thompson, Marin and Stewart (2012) were the first to show that Western amusics are impaired in decoding emotions from speech. Amusics and controls listened to semantically-neutral spoken phrases (e.g., “The broom is in the closet and the book is on the desk”) using headphones and were asked to report the conveyed emotion by choosing one out of six emotion categories: happy, sad, tender, irritated, afraid and neutral. Amusics performed worse than controls in classifying happiness, tenderness, sadness and irritation. They also reported some awareness of their emotion perception deficit in their daily lives. Lolli, Lewenstein, Basurto, Winnik and Loui (2015) further investigated the underlying cause of this perceptual impairment by working with the same set of stimuli and applying filtering techniques to manipulate the frequency content of the phrases. Amusics and controls performed similarly in the unfiltered condition, which did not corroborate findings by Thompson et al. (2012). More specifically, their findings indicated that among amusics pitch discrimination ability correlated with the performance in the emotion categorization task in the low-pass filtered speech condition (preserving the speech contour but not formants) but not in the high-pass filtered (preserving speech rate, stress patterns and rhythm) and unfiltered conditions. The data suggest that amusics with poor pitch discrimination abilities did not rely on high-frequency speech cues to perform the task.

Lima et al. (2016) investigated whether amusics' emotional prosody deficit extends to deficits in affective processing of other auditory or visual socio-emotional cues. They studied amusics' ability to decode emotions communicated by affective prosody (using different stimuli than Thompson et al., 2012, and Lolli et al., 2015), non-verbal vocalizations and dynamic facial expressions. Across modalities amusics showed impairment in affective processing, as demonstrated by lower recognition accuracy and more ambivalent responses. An analysis of the underlying acoustic cues of the vocal stimuli, focusing on pitch and non-pitch attributes, revealed that amusics' profile of acoustic cues predicting affective responses differed from the one of controls. Furthermore, Lima et al. reported that amusics were less able to judge the authenticity of laughter (pose vs. spontaneous) but showed normal contagion responses. In summary, this is the first study to suggest that amusia is a developmental disorder affecting the processing of socio-emotional stimuli other than auditory.

Future research on amusia and vocal affect processing may study, besides emotional communication, aesthetic aspects of speech and music (see Loutrari & Lorch, 2017, for a case study on the perception of musical and prosodic expressivity), which may reveal why some amusics are able to enjoy music (Omigie, Müllensiefen, & Stewart, 2012). Moreover, the neural underpinnings of the impairment of affective processing in amusia are still unclear. It needs to be determined whether the deficit observed in amusia is due to the fact that the right hemisphere, which is impaired in amusics, is mainly active in affective prosody processing in healthy individuals (Frühholz et al., 2016). Another open question is whether and how the ability to process affect in music is related to the processing of affect in speech. Several studies have assessed whether amusics show impairments in musical affect processing (Cousineau, McDermott, & Peretz, 2012; Gosselin, Paquette, & Peretz, 2015; Marin, Thompson, Gingras, & Stewart, 2015; Zhishuai, Hong, Daxing, Pin, & Xuejing, 2017), but these studies did not collect data on affective prosody perception. It is also unknown whether amusics can express and imitate emotions in speech to a similar degree as controls. Finally, future research may also reveal how amusia is linked to other developmental disorders in which socio-emotional processing is impaired, such as *autism spectrum disorders* (Griffiths et al., 2017). For example, current research seems to indicate that the ability to discriminate pitch direction may be correlated with melodic perception in individuals with autism spectrum disorders (Germain et al., 2018).

3.4 Syntax

Both language and music are auditory communication systems in which discrete elements are combined to form sequences of units that are arranged in a hierarchical structure. In both domains, these structures can be described by a set of *syntactic rules*. Patel (2003) hypothesized that language and music share syntactic resources in the brain, and abundant empirical evidence in support of this has accumulated since then. For example, a sharing of resources in Broca's area (see also Luef in this volume) in healthy individuals was demonstrated using fMRI (Kunert, Willems, Casasanto, Patel, & Hagoort, 2015). Studies using EEG have also shown that signatures of early and late processing stages are similar and overlap during language and music processing (Koelsch, Gunter, Wittfoth, & Sammler, 2005; Koelsch, Rohrmeier, Torrecuso, & Jentschke, 2013).

Sun et al. (2018) investigated English-speaking amusics and found that their brain responses to language- and music-syntactic irregularities (in speech and melodies) were abnormal compared to controls. Amusics were impaired in early, but not later, syntactic processing stages in both domains. To be specific, the authors could demonstrate that amusics showed no LAN (*Left Anterior Negativity*) and ERAN (*Early Right Anterior Negativity*), two event-related potential components associated with the processing of syntactic irregularities in language and music, respectively (Koelsch, 2005). Amusics' N5 and P600 components were similar to those of controls, reflecting intact syntactic processing at later stages. Furthermore, the authors did not observe any impairment regarding language-semantic processing among amusics, providing more evidence for different brain networks involved in syntactic and semantic processing.

3.5 Verbal short-term memory

Due to their dynamic nature, speech and music communication draw heavily on short-term memory. Amusia has not only been associated with perceptual deficits in pitch processing, but also with deficits in short-term memory for pitch (Tillmann, Schulze, & Foxton, 2009) and timbre (Marin, Gingras, & Stewart, 2012). However, it is important to state that short-term memory for the content of verbal materials, such as digits and words, is preserved in amusics (for a review see Tillmann et al., 2016). A recent review by Caclin and Tillmann (2018) suggests that the cognitive systems underlying short-term memory for musical and verbal materials only partly overlap. The authors also call for more systematic studies involving the comparison between short-term memory in music and speech, which could be accomplished by focusing only on one perceptual dimension in both domains (e.g., pitch) and by investigating individual components of short-term memory (i.e., encoding, retention and retrieval) in more detail.

4 SPEECH PROCESSING IN TONE LANGUAGE SPEAKERS WITH CONGENITAL AMUSIA

The substantial body of research on amusia in non-tone language speakers conducted during the first decade of the 21st century has provided a solid basis to initiate a similar research program in Asia. This move towards studying tone language speakers was partly motivated by the idea that being exposed to a tone language may compensate for being amusic (Peretz, 2008); however, as Jiang, Hamm, Lim, Kirk and Yang (2011) showed shortly after, Mandarin-speakers with amusia were impaired in two fine-grained pitch discrimination tasks. Moreover, around the same time, studies testing French-speaking amusics revealed impairment in processing *lexical tone* (Nguyen, Tillmann, Gosselin, & Peretz, 2009; Tillmann et al., 2011). Lexical tone refers to the alteration of word meaning by small pitch variations (i.e., a change in pitch contour). In Mandarin Chinese there are four tones, namely *level*, *mid-rising*, *dipping* and *high-falling*. If, for example, the syllable *ma* is pronounced with a level tone, it means 'mother', if it is pronounced with a dipping tone, it means 'horse'. Most research on amusia in tone language speakers has been conducted among Mandarin speakers, and more recently, some studies involved Cantonese speakers. Mandarin and Cantonese differ regarding pronunciation, grammar and lexicon. Since these two tone languages differ considerably, the research on speech processing in amusia is presented separately.

4.1 Speech processing in Mandarin speakers

4.1.1 *Lexical tone*

One of the first studies involving Mandarin-speaking amusics was published by Nan, Sun and Peretz (2010) who showed that the pitch impairment (i.e., the performance on the MBEA) in these amusics was comparable to the one reported for non-tone language speakers, which may suggest a common origin of the disorder. Moreover, they also assessed the perception and production of lexical tone (tone identification, tone discrimination and lexical tone production), and their results indicated that half of the amusics were impaired in lexical tone identification and discrimination while lexical tone production was intact (see section 3.2 for a description of a similar dissociation in non-tone language speakers). However, it was unclear whether the observed deficits were due to impaired phonological awareness or pitch processing because amusics were only impaired in tone perception when the syllables carrying the tone differed within a trial.

In a follow-up study on the same topic, Liu et al. (2012) used smaller tone contrasts in their word discrimination task, and they also minimized the demand for phonological awareness by adding Chinese characters to the tone identification task. They studied word identification and discrimination by using natural words as well as gliding tone analogues. Their results indicated impaired performance in the word discrimination task but intact performance in the word identification task. Surprisingly, Mandarin amusics performed better in the condition with gliding tones than in the natural tone condition. Differences in the experimental design may explain the discrepancy between Nan et al. (2010) and Liu et al. (2012) regarding the results of the word identification task. Altogether, these influential studies demonstrated that amusia is not specific to non-tone language speakers and, as previously shown, is not a music-specific disorder either.

Being interested in the *categorical perception* of lexical tones, Jiang, Hamm, Lim, Kirk and Yang (2012) studied the identification and discrimination of tones by generating continua of pitch patterns in 11 steps, ranging from a high-level tone to a mid-rising tone in one case, and in the other case from a high-level tone to a high-falling tone. In addition to the natural speech stimuli, non-linguistic pitch analogues were created. A classification task showed that amusics had shifted classification boundaries compared to controls and experienced a specific problem with classifying rising or falling pitches. The results of the discrimination task supported the view that amusics are impaired in lexical tone processing because their performances did not improve when the two stimuli crossed the classification boundary, which held true for both speech and non-linguistic pitch analogues.

Wang and Peng (2014) wanted to identify whether a deficit in pitch *or* phonological processing underlies the impairment found for lexical tone processing. Testing Mandarin amusics, the authors added familiarity as a potential factor to their research design by using fine-grained tone contrasts in Cantonese consonant-vowel combinations that were either similar or dissimilar to the Mandarin counterparts. Familiarity played a role when natural Cantonese level tones needed to be discriminated. Controls only performed better in the familiar condition, whereas in the unfamiliar condition controls and amusics performed similarly, indicating that the phonological deficit is the driving factor explaining amusics' performance. Amusics were also able to correctly imitate the Cantonese level tones (see section 2 for similar reports). This study also emphasizes that segmental and suprasegmental processing are closely intertwined in speech processing.

Motivated by the idea that there may be subgroups of amusics who are either impaired or unimpaired in their speech processing abilities (Nan et al., 2010; Patel et al., 2008; Yang, Feng, Huang, Zhang, & Nan, 2014), Huang, Liu, Dong and Nan (2015) studied categorical perception of lexical tones in three groups (controls, *pure amusics* and *amusics with lexical tone agnosia*) and compared their performances in a tone identification and discrimination task. Amusics with tone agnosia performed previously worse on a lexical tone identification and discrimination task than controls and pure amusics. For this study, the authors created a set of stimuli based on the vowel /a/ by covering a continuum of different frequency contours, ranging from rising to level tone, with a varying onset frequency and a fixed offset frequency. The results indicated that controls and pure amusics performed similarly well, and that only amusics with tone agnosia were impaired in categorical perception. The sample size was very small, but the data suggest that there are subsystems of pitch processing that are not shared in music and speech, which clearly reinvigorates the discussion about *modularity* (i.e., domain-specificity). The same research group also demonstrated that only amusics with tone agnosia were impaired in pitch contour processing (Huang, Nan, Dong, & Liu, 2015).

Neurophysiological evidence has recently corroborated this finding. Nan, Huang, Wang, Liu and Dong (2016) used EEG to record *mismatch negativities* (MMNs) in response to changes in lexical tones in a sequence of stimuli. The MMN is an event-related potential that occurs in response to an odd stimulus in a sequence of stimuli (Näätänen & Alho, 1995). Amusics with tone agnosia had reduced MMN responses in comparison to pure amusics and controls. Furthermore, the authors demonstrated that there was a positive association between the neural responses and the behavioral measures, showing that the MMN is a well-suited objective measure of performance. Future research will have to meet the challenge of determining whether the differentiation into pure amusics and amusics with tone agnosia is justified, and if so, to what degree genetic or environmental factors can explain the existence of these subgroups. It is plausible to assume that this debate will profit from our growing understanding of the underlying pitch processing deficit in amusia, in both tone and non-tone language speakers.

Research on lexical tone processing may also reveal insights into basic questions of pitch processing by investigating which aspects of pitch processing are impaired and how amusics compensate for them. In an elaborate experiment, Bones and Wong (2017) recently reported that amusics may use *temporal-envelope pitch cues* to process lexical tones. These are pitch cues that are inherent in the temporal envelope of the high-numbered unresolved harmonics. The authors used chimera stimuli in which ambiguous words contained pitch cues in the temporal fine-structure and temporal envelope. Amusics were more likely than controls to base their judgments of words on envelope cues, suggesting that amusics rely on a secondary mechanism of pitch processing.

4.1.2 *Speech intonation*

In the literature on linguistic prosody processing in tone language speakers, there seems to be a general agreement that amusics show an impairment in laboratory conditions, which is in line with previous studies involving non-tone language speakers (see section 3.2). Jiang, Hamm, Lim, Kirk and Yang (2010) were the first to demonstrate that Mandarin-speaking amusics are impaired in melodic contour processing as well as in speech intonation processing. By using discrimination and identification tasks, they showed that amusics performed worse than controls in differentiating between statements and questions in natural speech (based on two-syllable verb-object constructions whose pitch pattern of the final syllable was manipulated) and also in conditions with non-linguistic pitch-analogues. The data did not reveal subgroups of amusics. Liu et al. (2012) employed longer speech stimuli (statements and questions) that varied naturally in duration, intensity and pitch. The authors hypothesized that amusics should be able to differentiate between statements and questions by using acoustic cues other than pitch, which was indeed the case. This finding may explain why amusics generally do not report difficulties in speech perception in their daily lives.

Jiang, Hamm, Kim et al. (2012) studied the cognitive processing of linguistic prosody during listening to short discourses using EEG. In these discourses, a question (speaker A) was followed by a yes/no answer and a sentence consisting of two clauses (speaker B), of which the second clause was either a statement or question. Each discourse was recorded twice, once with the correct and once with the incorrect intonation at the end of the discourse (last syllable). This enabled the authors to have congruent and incongruent pairings between semantic content and prosody. Amusics performed worse than controls in a speech comprehension task, and the neurophysiological data revealed that the P600 and N100 effects differentiating between congruent and incongruent intonation in controls were not present in amusics. Among other things, the P600 effect can occur in the context of a syntax-prosody mismatch (Eckstein & Friederici, 2005), whereas the N100 effect has also been associated with intonation expectation (Näätänen & Picton, 1987). Therefore, the amusic brain seems to be impaired in processing the (in)congruency between semantic meaning and intonation.

Relatively little is understood about whether the pitch processing deficit underlying the impairment in intonation processing can be overcome by drawing on other linguistic information. In an ERP study, Lu, Ho, Liu, Wu and Thompson (2015) thus investigated whether the emotional valence of words can compensate for impaired intonation processing (statements vs. questions) in amusics. They were motivated by previous reports that emotional linguistic information can facilitate stimulus processing, but the authors did not find a positive effect of emotion. Amusics showed impaired performance in the intonation-matching task as well as a reduced N2 response to incongruent intonation pairs. The N2 response has previously been found to be larger in conflict than in non-conflict trials (Nieuwenhuis, Yeung, Van Den Wildenberg, & Ridderinkhof, 2003). Nevertheless, Lu et al.'s data indicated that early auditory processing is intact in amusics because the N1 response was similar in amusics and controls.

Speech intelligibility in noisy conditions was the center of interest in a study by Liu, Jiang, Wang, Xu and Patel (2015). Their stimuli comprised news-like sentences spoken in Mandarin, and the sentences' pitch contour was either natural or flattened. There were four conditions of babble noise and a quiet control condition. Participants were asked to listen to the sentences and to write down what they understood on a piece of paper. Amusics (young, highly-educated adults) were impaired in speech comprehension in all conditions regardless of the sentence type (natural or flattened). Moreover, the authors reported that this deficit among amusics was not related to their pitch processing deficit, which is in line with recent findings suggesting that amusics may be impaired in other aspects of auditory processing besides pitch processing (Whiteford & Oxenham, 2017). Liu et al. (2015) speculated that amusics may also be impaired in segmental processing, which may be relevant for the debate surrounding the rapid auditory processing hypothesis in amusia which has already been discussed in relation to non-tone language speakers (Albouy et al., 2016) and other neurodevelopmental disorders.

4.2 Evidence from Cantonese speakers with congenital amusia

Replication of results should be at the core of any empirical research field. Liu et al. (2016) recently replicated findings previously reported for Western (English and French) and Mandarin-speaking amusics in Cantonese-speaking amusics. Cantonese is characterized by six tones and thus exhibits a more complex tone system than Mandarin. Liu et al. employed a large set of tests, measuring pitch thresholds, lexical tone perception and production as well as singing ability. In brief, their results indicated that amusics showed severe impairments in music perception and production, mild impairments in speech perception and intact speech production. Therefore, this study nicely corroborates previous accounts and calls for similar studies involving other languages.

Since some evidence has accumulated that amusia may be more than a domain-general pitch processing deficit (Lehmann et al., 2015; Liu et al., 2015; Whiteford & Oxenham, 2017), Zhang, Shao and Huang (2017) investigated frequency processing beyond pitch processing. To be specific, they examined frequency/spectral processing in the context of segmental processing and the perception of vowels. Spectral frequency peaks (*formants*), in particular the first and second formants, determine the difference between vowels. Sound continua in three frequency-based stimulus types (lexical tone, pure tone and vowel) and one temporal stimulus type (voice onset time) were employed. The authors reported that amusics were impaired in the discrimination, but not in the identification, of all three frequency-based stimuli types. Amusics were unimpaired in the processing of temporal information. They concluded that the data suggest that amusics are impaired in phonological processing of speech sounds as well as in basic auditory pitch processing of pure tones, which clearly motivates further studies on how this disorder manifests itself in different auditory contexts.

To summarize, considerable research attention has been devoted to the study of speech intonation and lexical tone in amusia in tone language speakers during the last few years, including the employment of unique paradigms (Liu et al., 2015; Lu et al., 2015) that could be applied to the study of Western language speakers. In general, the prosodic impairment in tone language speakers with amusia seems to be similar to the one observed in non-tone language speakers, which suggests a culture-independent impairment. The next steps will probably include the study of affective prosody perception and production to be able to make similar claims regarding that aspect of speech processing.

5 CONCLUSION

In the last 15 years, research on the nature of amusia has made significant progress, which has led to the current notion that amusia involves several cognitive deficits related to pitch processing, namely those regarding pitch change detection, pitch direction discrimination as well as short-term memory for pitch. Importantly, future research will have to clarify whether other basic perceptual/cognitive deficits, going beyond pitch processing, need to be included in the characterization of amusia, and further, how these deficits overlap with other developmental learning disorders (Peretz, 2016).

In this light, research on speech processing in congenital amusia has considerably changed our view about this auditory neurodevelopmental disorder. Although amusia is mostly conceived of as a musical disorder, it is now generally accepted that individuals with amusia show a wide range of deficits in speech processing in laboratory conditions due to a domain-general impairment in pitch processing. Thus, the reported deficits regarding phonology, prosody and syntax are pertinent to our understanding of how music and speech are processed in the brain and support speculations about a common origin of music and language. This opens a door for further studies exploring cognitive and affective deficits in amusia across domains.

A representative body of research involving both tone and non-tone language speakers has dealt with speech perception and production, comprising a wide range of tasks and stimulus types as well as behavioral and neurophysiological measures. Only such a comprehensive approach involving researchers from different fields (linguists, music psychologists and neuroscientists) has made it possible to arrive at the current state of knowledge about amusia. Consequently, it is my fervent hope that research on amusia will remain an interdisciplinary playground hereafter.

Future research will have to meet the challenge of replicating the present results in language speakers other than English, French, Mandarin and Cantonese. This would necessitate that more research groups become interested in studying amusia, and ideally, that research groups collaborate and conduct the same experiments in cohorts from different cultures. Such an approach will also clarify whether subgroups of amusics exist, and if so, whether they are language-specific.

ACKNOWLEDGEMENTS

I am very grateful to Fang Liu, Bruno Gingras and an anonymous reviewer for helpful comments, which have greatly strengthened the chapter.

REFERENCES

Albouy, P., Cousineau, M., Caclin, A., Tillmann, B., & Peretz, I. (2016). Impaired encoding of rapid pitch information underlies perception and memory deficits in congenital amusia. *Scientific Reports, 6*, 18861.

Albouy, P., Mattout, J., Bouet, R., Maby, E., Sanchez, G., Aguera, P. E., … Tillmann, B. (2013). Impaired pitch perception and memory in congenital amusia: The deficit starts in the auditory cortex. *Brain, 136*, 1639–1661.

Anderson, S., Himonides, E., Wise, K., Welch, G., & Stewart, L. (2012). Is there potential for learning in amusia? A study of the effect of singing intervention in congenital amusia. *Annals of the New York Academy of Sciences, 1252*, 345–353.

Anvari, S. H., Trainor, L. J., Woodside, J., & Levy, B. A. (2002). Relations among musical skills, phonological processing, and early reading ability in preschool children. *Journal of Experimental Child Psychology, 83*, 111–130.

Asaridou, S. S., & McQueen, J. M. (2013). Speech and music shape the listening brain: Evidence for shared domain-general mechanisms. *Frontiers in Psychology, 4*, 321.

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

Bachorowski, J.-A. (1999). Vocal expression and perception of emotion. *Current Directions in Psychological Science, 8*, 53–57.

Bänziger, T., Patel, S., & Scherer, K. R. (2014). The role of perceived voice and speech characteristics in vocal emotion communication. *Journal of Nonverbal Behavior, 38*, 31–52.

Bauer, J. J., & Larson, C. R. (2003). Audio-vocal responses to repetitive pitch-shift stimulation during a sustained vocalization: Improvements in methodology for the pitch-shifting technique. *The Journal of the Acoustical Society of America, 114*, 1048–1054.

Belyk, M., & Brown, S. (2014). Perception of affective and linguistic prosody: An ALE meta-analysis of neuroimaging studies. *Social Cognitive and Affective Neuroscience, 9*, 1395–1403.

Bernstein, J. G., & Oxenham, A. J. (2003). Pitch discrimination of diotic and dichotic tone complexes: Harmonic resolvability or harmonic number? *The Journal of the Acoustical Society of America, 113*, 3323–3334.

Besson, M., Barbaroux, M., & Dittinger, E. (2017). Music in the brain: Music and language processing. In R. Ashley & R. Timmers (Eds.), *The Routledge companion to music cognition* (pp. 37–48). New York: Routledge.

Blacking, J. (1973). *How musical is man?* Seattle: University of Washington Press.

Bones, O., & Wong, P. C. M. (2017). Congenital amusics use a secondary pitch mechanism to identify lexical tones. *Neuropsychologia, 104*, 48–53.

Brandt, A., Gebrian, M., & Slevc, L. R. (2012). Music and early language acquisition. *Frontiers in Psychology, 3*, 327.

Caclin, A., & Tillmann, B. (2018). Musical and verbal short-term memory: Insights from neurodevelopmental and neurological disorders. *Annals of the New York Academy of Sciences*. Advance online publication. https://doi.org/10.1111/nyas.13733

Chen, J. L., Kumar, S., Williamson, V. J., Scholz, J., Griffiths, T. D., & Stewart, L. (2015). Detection of the arcuate fasciculus in congenital amusia depends on the tractography algorithm. *Frontiers in Psychology, 6*, 9.

Cousineau, M., McDermott, J. H., & Peretz, I. (2012). The basis of musical consonance as revealed by congenital amusia. *Proceedings of the National Academy of Sciences USA, 109*, 19858–19863.

Cousineau, M., Oxenham, A. J., & Peretz, I. (2015). Congenital amusia: A cognitive disorder limited to resolved harmonics and with no peripheral basis. *Neuropsychologia, 66*, 293–301.

Cross, I., & Morley, I. (2009). The evolution of music: Theories, definitions and the nature of the evidence. In S. Malloch & C. Trevarthen (Eds.), *Communicative musicality: Exploring the basis of human companionship* (pp. 61–81). Oxford: Oxford University Press.

Crystal, D., & Quirk, R. (1964). *Systems of prosodic and paralinguistic features in English*. The Hague: Mouton.

Darwin, C. R. (1871). *The descent of man, and selection in relation to sex*. London: John Murray.

Dawson, C., Aalto, D., Simko, J., Vainio, M., & Tervaniemi, M. (2017). Musical sophistication and the effect of complexity on auditory discrimination in Finish speakers. *Frontiers in Neuroscience, 11*, 213.

Degé, F., & Schwarzer, G. (2016). The influence of music training on the development of phonological awareness in preschoolers. *Frühe Bildung, 5*, 150–156.

Don, A. J., Schellenberg, G. E., & Rourke, B. P. (1999). Music and language skills of children with Williams syndrome. *Child Neuropsychology, 5*, 154–170.

Doolittle, E., & Gingras, B. (2015). Zoomusicology. *Current Biology, 25*, R819-R820.

Eckstein, K., & Friederici, A. D. (2005). Late interaction of syntactic and prosodic processes in sentence comprehension as revealed by ERPs. *Cognitive Brain Research, 25*, 130–143.

Fisher, S. E. (2017). Evolution of language: Lessons from the genome. *Psychonomic Bulletin & Review, 24*, 34–40.

Fitch, W. T. (2006). The biology and evolution of music: A comparative perspective. *Cognition, 100*, 173–215.

Fitch, W. T. (August 5, 2010). Musical protolanguage: Darwin's theory of language evolution revisited. Retrieved from http://homepage.univie.ac.at/tecumseh.fitch/2010/08/05/musical-protolanguage-darwins-theory-of-language-evolution-revisited/

Flaugnacco, E., Lopez, L., Terribili, C., Zoia, S., Buda, S., Tilli, S., ... Schon, D. (2014). Rhythm perception and production predict reading abilities in developmental dyslexia. *Frontiers in Human Neuroscience, 8*, 392.

Fonagy, I. (1978). A new method of investigating the perception of prosodic features. *Language and Speech, 21*, 34–49.

Foxton, J. M., Dean, J. L., Gee, R., Peretz, I., & Griffiths, T. D. (2004). Characterization of deficits in pitch perception underlying 'tone deafness'. *Brain, 127*, 801–810.

Frühholz, S., Trost, W., & Kotz, S. A. (2016). The sound of emotions: Towards a unifying neural network perspective of affective sound processing. *Neuroscience & Biobehavioral Reviews, 68*, 96–110.

Germain, E., Foster, N. E., Sharda, M., Chowdhury, R., Tryfon, A., Doyle-Thomas, K. A., … & Hyde, K. L. (2018). Pitch direction ability predicts melodic perception in autism. *Child Neuropsychology*. Advance online publication. https://doi.org/10.1080/09297049.2018.1488954

Gillon, G. T. (2018). *Phonological awareness: From research to practice*. New York: Guilford.

Gingras, B., Honing, H., Peretz, I., Trainor, L. J., & Fisher, S. E. (2015). Defining the biological bases of individual differences in musicality. *Philosophical Transactions of the Royal Society B: Biological Sciences, 370*, 20140092.

Glushko, A., Steinhauer, K., DePriest, J., & Koelsch, S. (2016). Neurophysiological correlates of musical and prosodic phrasing: Shared processing mechanisms and effects of musical expertise. *PLoS ONE, 11*, e0155300.

Goodale, M. A., & Milner, A. D. (1992). Separate visual pathways for perception and action. *Trends in Neurosciences, 15*, 20–25.

Gorin, S., Kowialiewski, B., & Majerus, S. (2016). Domain-generality of timing-based serial order processes in short-term memory: New insights from musical and verbal domains. *PLoS ONE, 11*, 25.

Gosselin, N., Paquette, S., & Peretz, I. (2015). Sensitivity to musical emotions in congenital amusia. *Cortex, 71*, 171–182.

Goulet, G. M., Moreau, P., Robitaille, N., & Peretz, I. (2012). Congenital amusia persists in the developing brain after daily music listening. *PLoS ONE, 7*, e36860.

Griffiths, S., Jarrold, C., Penton-Voak, I. S., Woods, A. T., Skinner, A. L., & Munafò, M. R. (2017). Impaired recognition of basic emotions from facial expressions in young people with autism spectrum disorder: Assessing the importance of expression intensity. *Journal of Autism and Developmental Disorders*. Advance online publication. https://doi.org/10.1007/s10803–017–3091–7

Hauser, M. D., & Watumull, J. (2017). The Universal Generative Faculty: The source of our expressive power in language, mathematics, morality, and music. *Journal of Neurolinguistics, 43*, 78–94.

Henry, M. J., & McAuley, J. D. (2010). On the prevalence of congenital amusia. *Music Perception, 27*, 413–418.

Hickok, G., & Poeppel, D. (2004). Dorsal and ventral streams: A framework for understanding aspects of the functional anatomy of language. *Cognition, 92*, 67–99.

Higham, T., Basell, L., Jacobi, R., Wood, R., Ramsey, C. B., & Conard, N. J. (2012). Testing models for the beginnings of the Aurignacian and the advent of figurative art and music: The radiocarbon chronology of Geissenklosterle. *Journal of Human Evolution, 62*, 664–676.

Hockett, C. F. (1960). Logical considerations in the study of animal communication. In W. E. Lanyon & W. N. Tavolga (Eds.), *Animal sounds and communication* (pp. 392–430). Washington: American Institute of Biological Sciences.

Honing, H., ten Cate, C., Peretz, I., & Trehub, S. (2015). Without it no music: Cognition, biology and evolution of musicality. *Philosophical Transactions of the Royal Society B: Biological Sciences, 370*, 20140088.

Huang, W. T., Liu, C., Dong, Q., & Nan, Y. (2015). Categorical perception of lexical tones in Mandarin-speaking congenital amusics. *Frontiers in Psychology, 6*, 829.

Huang, W. T., Nan, Y., Dong, Q., & Liu, C. (2015). Just-noticeable difference of tone pitch contour change for Mandarin congenital amusics. *The Journal of the Acoustical Society of America, 138*, EL99-EL104.

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

Hutchins, S., & Peretz, I. (2012). Amusics can imitate what they cannot discriminate. *Brain and Language, 123*, 234–239.

Hutchins, S., & Peretz, I. (2013). Vocal pitch shift in congenital amusia (pitch deafness). *Brain and Language, 125*, 106–117.

Hyde, K. L., & Peretz, I. (2004). Brains that are out of tune but in time. *Psychological Science, 15*, 356–360.

Jiang, C. M., Hamm, J. P., Kim, V. K., Kirk, I. J., Chen, X. H., & Yang, Y. F. (2012). Amusia results in abnormal brain activity following inappropriate intonation during speech comprehension. *PLoS ONE, 7*, e41411.

Jiang, C. M., Hamm, J. P., Lim, V. K., Kirk, I. J., & Yang, Y. F. (2010). Processing melodic contour and speech intonation in congenital amusics with Mandarin Chinese. *Neuropsychologia, 48*, 2630–2639.

Jiang, C. M., Hamm, J. P., Lim, V. K., Kirk, I. J., & Yang, Y. F. (2011). Fine-grained pitch discrimination in congenital amusics with Mandarin chinese. *Music Perception, 28*, 519–526.

Jiang, C. M., Hamm, J. P., Lim, V. K., Kirk, I. J., & Yang, Y. F. (2012). Impaired categorical perception of lexical tones in Mandarin-speaking congenital amusics. *Memory & Cognition, 40*, 1109–1121.

Jiang, C., Lim, V. K., Wang, H., & Hamm, J. P. (2013). Difficulties with pitch discrimination influences pitch memory performance: Evidence from congenital amusia. *PLoS ONE*, *8*, e79216.

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

Juslin, P. N., & Laukka, P. (2003). Communication of emotions in vocal expression and music performance: Different channels, same code? *Psychological Bulletin, 129*, 770–814.

Kalmus, H., & Fry, D. (1980). On tune deafness (dysmelodia): Frequency, development, genetics and musical background. *Annals of Human Genetics, 43*, 369–382.

Koelsch, S. (2005). Neural substrates of processing syntax and semantics in music. *Current Opinion in Neurobiology, 15*, 207–212.

Koelsch, S., Gunter, T. C., Wittfoth, M., & Sammler, D. (2005). Interaction between syntax processing in language and in music: An ERP study. *Journal of Cognitive Neuroscience, 17*, 1565–1577.

Koelsch, S., Rohrmeier, M., Torrecuso, R., & Jentschke, S. (2013). Processing of hierarchical syntactic structure in music. *Proceedings of the National Academy of Sciences USA, 110*, 15443–15448.

Kunert, R., Willems, R. M., Casasanto, D., Patel, A. D., & Hagoort, P. (2015). Music and language syntax interact in Broca's area: An fMRI study. *PLoS ONE, 10*, e0141069.

Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science, 207*, 203–205.

Lamb, S. J., & Gregory, A. H. (1993). The relationship between music and reading in beginning readers. *Educational Psychology, 13*, 19–27.

Lee, H. Y., Sie, Y. S., Chen, S. C., & Cheng, M. C. (2015). The music perception performance of children with and without dyslexia in Taiwan. *Psychological Reports, 116*, 13–22.

Lehmann, A., Skoe, E., Moreau, P., Peretz, I., & Kraus, N. (2015). Impairments in musical abilities reflected in the auditory brainstem: Evidence from congenital amusia. *European Journal of Neuroscience, 42*, 1644–1650.

Lennon, J., & Slesinski, C. (2003). Comprehensive test of phonological processing (CTOPP): Cognitive linguistic assessment of severe reading problems. Retrieved from http://alpha.fdu.edu/psychology/comprehensive_test_of_phonologic.htm

Lerdahl, F., & Jackendoff, R. (1985). *A generative theory of tonal music*. Cambridge: MIT press.

Lima, C. F., Brancatisano, O., Fancourt, A., Müllensiefen, D., Scott, S. K., Warren, J. D., & Stewart, L. (2016). Impaired socio-emotional processing in a developmental music disorder. *Scientific Reports, 6*, 34911.

Lima, C. F., & Castro, S. L. (2011). Speaking to the trained ear: Musical expertise enhances the recognition of emotions in speech prosody. *Emotion, 11*, 1021–1031.

Linnavalli, T., Putkinen, V., Lipsanen, J., Huotilainen, M., & Tervaniemi, M. (2018). Music playschool enhances children's linguistic skills. *Scientific Reports, 8*, 8767.

Liu, F., Chan, A. H., Ciocca, V., Roquet, C., Peretz, I., & Wong, P. C. M. (2016). Pitch perception and production in congenital amusia: Evidence from Cantonese speakers. *The Journal of the Acoustical Society of America, 140*, 563–575.

Liu, F., Jiang, C., Francart, T., Chan, A. H. D., & Wong, P. C. M. (2017). Perceptual learning of pitch direction in congenital amusia. Evidence from Chinese speakers. *Music Perception, 34*, 335–351.

Liu, F., Jiang, C., Thompson, W. F., Xu, Y., Yang, Y., & Stewart, L. (2012). The mechanism of speech processing in congenital amusia: Evidence from Mandarin speakers. *PLoS ONE, 7*, e30374.

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*, 1682–1693.

Lochy, A., Hyde, K. L., Parisel, S., Vanhyfte, S., & Peretz, I. (2004). Discrimination of speech prosody in congenital amusia. Poster session presented at the Meeting of the Cognitive Neuroscience Society, San Francisco.

Lolli, S. L., 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.

Loui, P., Guenther, F. H., Mathys, C., & Schlaug, G. (2008). Action-perception mismatch in tone-deafness. *Current Biology, 18*, R331-R332.

Loutrari, A., & Lorch, M. P. (2017). Preserved appreciation of aesthetic elements of speech and music prosody in an amusic individual: A holistic approach. *Brain and Cognition, 115*, 1–11.

Lu, X., Ho, H. T., Liu, F., Wu, D. X., & Thompson, W. F. (2015). Intonation processing deficits of emotional words among Mandarin Chinese speakers with congenital amusia: An ERP study. *Frontiers in Psychology, 6*, 385.

Lu, X., Sun, Y., Ho, H. T., & Thompson, W. F. (2017). Pitch contour impairment in congenital amusia: New insights from the Self-paced Audio-visual Contour Task (SACT). *PLoS ONE, 12*, e0179252.

Mariath, L. M., da Silva, A. M., Kowalski, T. W., Gattino, G. S., de Araujo, G. A., Figueiredo, F. G., … Schuch, J. B. (2017). Music genetics research: Association with musicality of a polymorphism in the AVPR1A gene. *Genetics and Molecular Biology, 40*, 421–429.

Marin, M. M. (2009). Effects of early musical training on musical and linguistic syntactic abilities. *Annals of the New York Academy of Sciences, 1169*, 187–190.

Marin, M. M., Gingras, B., & Stewart, L. (2012). Perception of musical timbre in congenital amusia: Categorization, discrimination and short-term memory. *Neuropsychologia, 50*, 367–378.

Marin, M. M., & Peltzer-Karpf, A. (2009). Towards a dynamic systems approach to the development of language and music: Theoretical foundations and methodological issues. In J. Louhivuori, T. Eerola, S. Saarikallio, T. Himberg, & P.-S. Eerola (Eds.), *Proceedings of the 7th Triennial Conference of the European Society for the Cognitive Sciences of Music (ESCOM 2009)* (pp. 284–292). Jyväskylä: University of Jyväskylä.

Marin, M. M., Thompson, W. F., Gingras, B., & Stewart, L. (2015). Affective evaluation of simultaneous tone combinations in congenital amusia. *Neuropsychologia, 78*, 207–220.

Martin, J. G. (1972). Rhythmic (hierarchical) versus serial structure in speech and other behavior. *Psychological Review, 79*, 487–509.

Martins, M. D., Gingras, B., Puig-Waldmueller, E., & Fitch, W. T. (2017). Cognitive representation of "musical fractals": Processing hierarchy and recursion in the auditory domain. *Cognition, 161*, 31–45.

McDermott, J., & Hauser, M. (2005). The origins of music: Innateness, uniqueness, and evolution. *Music Perception, 23*, 29–59.

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

Merker, B., Morley, I., & Zuidema, W. (2015). Five fundamental constraints on theories of the origins of music. *Philosophical Transactions of the Royal Society B: Biological Sciences, 370*, 20140095.

Moreno, S., Friesen, D., & Bialystok, E. (2011). Effect of music training on promoting pleliteracy skills: Preliminary causal evidence. *Music Perception, 29*, 165–172.

Morrill, T. H., McAuley, J. D., Dilley, L. C., & Hambrick, D. Z. (2015). Individual differences in the perception of melodic contours and pitch-accent timing in speech: Support for domain-generality of pitch processing. *Journal of Experimental Psychology-General, 144*, 730–736.

Näätänen, R., & Alho, K. (1995). Mismatch negativity-a unique measure of sensory processing in audition. *International Journal of Neuroscience, 80*, 317–337.

Näätänen, R., & Picton, T. (1987). The N1 wave of the human electric and magnetic response to sound: A review and an analysis of the component structure. *Psychophysiology, 24*, 375–425.

Nan, Y., Huang, W. T., Wang, W. J., Liu, C., & Dong, Q. (2016). Subgroup differences in the lexical tone mismatch negativity (MMN) among Mandarin speakers with congenital amusia. *Biological Psychology, 113*, 59–67.

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

Nguyen, S., Tillmann, B., Gosselin, N., & Peretz, I. (2009). Tonal language processing in congenital amusia. *Annals of the New York Academy of Sciences, 1169*, 490–493.

Nieuwenhuis, S., Yeung, N., Van Den Wildenberg, W., & Ridderinkhof, K. R. (2003). Electrophysiological correlates of anterior cingulate function in a go/no-go task: Effects of response conflict and trial type frequency. *Cognitive, Affective, & Behavioral Neuroscience, 3*, 17–26.

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*, 2986–2994.

Omigie, D., Müllensiefen, D., & Stewart, L. (2012). The experience of music in congenital amusia. *Music Perception, 30*, 1–18.

Ozernov-Palchik, O., & Gaab, N. (2016). Tackling the 'dyslexia paradox': Reading brain and behavior for early markers of developmental dyslexiax. *Wiley Interdisciplinary Reviews: Cognitive Science, 7*, 156–176.

Patel, A. D. (2003). Language, music, syntax and the brain. *Nature Neuroscience, 6*, 674–680.

Patel, A. D. (2017). Using music to study the evolution of cognitive mechanisms relevant to language. *Psychonomic Bulletin & Review, 24*, 177–180.

Patel, A. D., Foxton, J. M., & Griffiths, T. D. (2005). Musically tone-deaf individuals have difficulty discriminating intonation contours extracted from speech. *Brain and Cognition, 59*, 310–313.

Patel, A. D., Wong, M., Foxton, J., Lochy, A., & Peretz, I. (2008). Speech intonation perception deficits in musical tone deafness (congenital amusia). *Music Perception, 25*, 357–368.

Peretz, I. (2008). Musical disorders: From behavior to genes. *Current Directions in Psychological Science, 17*, 329–333.

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

Peretz, I., Ayotte, J., Zatorre, R. J., Mehler, J., Ahad, P., Penhune, V. B., & Jutras, B. (2002). Congenital amusia: A disorder of fine-grained pitch discrimination. *Neuron, 33*, 185–191.

Peretz, I., Champod, A. S., & Hyde, K. (2003). Varieties of musical disorders. The Montreal Battery of Evaluation of Amusia. *Annals of the New York Academy of Sciences, 999*, 58–75.

Peretz, I., & Coltheart, M. (2003). Modularity of music processing. *Nature Neuroscience, 6*, 688–691.

Peretz, I., Cummings, S., & Dubé, M.-P. (2007). The genetics of congenital amusia (tone deafness): A family-aggregation study. *The American Journal of Human Genetics, 81*, 582–588.

Pfeifer, S., & Hamann, S. (2018). The nature and nurture of congenital amusia: A twin case study. *Frontiers in Behavioral Neuroscience, 12*, 120.

Pfordresher, P. Q., & Brown, S. (2007). Poor-pitch singing in the absence of 'tone deafness'. *Music Perception, 25*, 95–115.

Prosody. (2017). *Oxford English Dictionary*. Retrieved from http://www.oed.com/view/Entry/152998?redirectedFrom=prosody&

Ramus, F., Rosen, S., Dakin, S. C., Day, B. L., Castellote, J. M., White, S., & Frith, U. (2003). Theories of developmental dyslexia: Insights from a multiple case study of dyslexic adults. *Brain, 126*, 841–865.

Raschle, N. M., Stering, P. L., Meissner, S. N., & Gaab, N. (2014). Altered neuronal response during rapid auditory processing and its relation to phonological processing in prereading children at familial risk for dyslexia. *Cerebral Cortex, 24*, 2489–2501.

Rohrmeier, M., Zuidema, W., Wiggins, G. A., & Scharff, C. (2015). Principles of structure building in music, language and animal song. *Philosophical Transactions of the Royal Society B: Biological Sciences, 370*, 20140097.

Sauter, D. A., Eisner, F., Ekman, P., & Scott, S. K. (2010). Cross-cultural recognition of basic emotions through nonverbal emotional vocalizations. *Proceedings of the National Academy of Sciences USA, 107*, 2408–2412.

Slevc, L. R., & Miyake, A. (2006). Individual differences in second-language proficiency: Does musical ability matter? *Psychological Science, 17*, 675–681.

Spencer, H. (1911). *Essays on education and kindred subjects*. London: Dent & Sons.

Spinelli, M., Fasolo, M., & Mesman, J. (2017). Does prosody make the difference? A meta-analysis on relations between prosodic aspects of infant-directed speech and infant outcomes. *Developmental Review, 44* (Supplement C), 1–18.

Steinbeis, N., & Koelsch, S. (2008). Shared neural resources between music and language indicate semantic processing of musical tension-resolution patterns. *Cerebral Cortex, 18*, 1169–1178.

Sun, Y. A., Lu, X., Ho, H. T., Johnson, B. W., Sammler, D., & Thompson, W. F. (2018). Syntactic processing in music and language: Parallel abnormalities observed in congenital amusia. *NeuroImage: Clinical*. Advance online publication. https://www.sciencedirect.com/science/article/pii/S2213158218301785

Sun, Y. A., Lu, X., Ho, H. T., & Thompson, W. F. (2017). Pitch discrimination associated with phonological awareness: Evidence from congenital amusia. *Scientific Reports, 7*, 44285.

ten Cate, C. (2017). Assessing the uniqueness of language: Animal grammatical abilities take center stage. *Psychonomic Bulletin & Review, 24*, 91–96.

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 USA, 109*, 19027–19032.

Tillmann, B. (2012). Music and language perception: Expectations, structural integration, and cognitive sequencing. *Topics in Cognitive Science, 4*, 568–584.

Tillmann, B., Burnham, D., Nguyen, S., Grimault, N., Gosselin, N., & Peretz, I. (2011). Congenital amusia (or tone-deafness) interferes with pitch processing in tone languages. *Frontiers in Psychology, 2,* 120.

Tillmann, B., Leveque, Y., Fornoni, L., Albouy, P., & Caclin, A. (2016). Impaired short-term memory for pitch in congenital amusia. *Brain Research, 1640*, 251–263.

Tillmann, B., Schulze, K., & Foxton, J. M. (2009). Congenital amusia: A short-term memory deficit for non-verbal, but not verbal sounds. *Brain and Cognition, 71*, 259–264.

Trehub, S. E., Trainor, L. J., & Unyk, A. M. (1993). Music and speech processing in the first year of life. *Advances in Child Development and Behavior, 24*, 1–35.

Vaiouli, P., & Andreou, G. (2018). Communication and language development of young children with autism: A review of research in music. *Communication Disorders Quarterly, 39*, 323–329.

Vuvan, D. T., Nunes-Silva, M., & Peretz, I. (2015). Meta-analytic evidence for the non-modularity of pitch processing in congenital amusia. *Cortex, 69*, 186–200.

Wagner, R. K., & Torgesen, J. K. (1987). The nature of phonological processing and its causal role in the acquisition of reading skills. *Psychological Bulletin, 101*, 192–212.

Wang, J., Zhang, C., Wan, S., & Peng, G. (2017). Is congenital amusia a disconnection syndrome? A study combining tract- and network-based analysis. *Frontiers in Human Neuroscience, 11*, 473.

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

Wennerstrom, A. (2001). *The music of everyday speech: Prosody and discourse analysis*: Oxford: Oxford University Press.

Whiteford, K. L., & Oxenham, A. J. (2017). Auditory deficits in amusia extend beyond poor pitch perception. *Neuropsychologia, 99*, 213–224.

Wilbiks, J. M. P., Vuvan, D. T., Girard, P.-Y., Peretz, I., & Russo, F. A. (2016). Effects of vocal training in a musicophile with congenital amusia. *Neurocase, 22*, 526–537.

Yang, W.-X., Feng, J., Huang, W.-T., Zhang, C.-X., & Nan, Y. (2014). Perceptual pitch deficits coexist with pitch production difficulties in music but not Mandarin speech. *Frontiers in Psychology, 4*, 1024.

Zhang, C., Peng, G., Shao, J., & Wang, W. S. Y. (2017). Neural bases of congenital amusia in tonal language speakers. *Neuropsychologia, 97*, 18–28.

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*, e0183151.

Zhao, Y. X., Chen, X. Z., Zhong, S. Y., Cui, Z. X., Gong, G. L., Dong, Q., & Nan, Y. (2016). Abnormal topological organization of the white matter network in Mandarin speakers with congenital amusia. *Scientific Reports, 6*, 26505.

Zhishuai, J., Hong, L., Daxing, W., Pin, Z., & Xuejing, L. (2017). Processing of emotional faces in congenital amusia: An emotional music priming event-related potential study. *NeuroImage: Clinical, 14*, 602–609.

Zhou, L. S., Liu, F., Jing, X. Y., & Jiang, C. M. (2017). Neural differences between the processing of musical meaning conveyed by direction of pitch change and natural music in congenital amusia. *Neuropsychologia, 96*, 29–38.