BIROn - Birkbeck Institutional Research Online

Loutrari, Ariadni and Lorch, Marjorie (2018) Music and language expressiveness: When emotional character does not suffice: the dimension of expressiveness in the cognitive processing of music and language. In: Gouk, P. and Kennaway, J. and Prins, J. and Thormahlen, W. (eds.) Routledge Companion to Music, Mind and Well-being. Routledge Music Companions. Abingdon, UK: Routledge. ISBN 9781138057760.

When emotional character does not suffice: the dimension of expressiveness in the cognitive processing of music and language

Ariadne Loutrari and Marjorie Lorch

Introduction

The power of music and its effect on mood and wellbeing have been considered throughout history from a variety of perspectives that have reflected contemporary interests at different times. Today, questions on how the brain derives pleasure from music and how one enjoys musical emotions, even when they evoke sadness, constitute part of contemporary explorations in cognitive science. While some of these questions are addressed in other chapters of this book, in the current chapter we investigate how the comparison of music with another cognitive faculty, namely language, can contribute to our understanding of emotions, mind, and brain.

In this chapter, we critically examine some of the contemporary areas of investigation in the parallel study of language and music in the domain of prosody. These areas mainly pertain to the expression of syntactical relationships or emotional tone. In later sections, we raise the question of whether the existing research on language and music prosody is sufficient to capture all possible prosodic aspects of speech and music streams. We go on to consider a novel way of examining "expressiveness" in relation to cognition and present some primary results of employing this approach that point to a dissociation between perception of "expressiveness" and other systematically studied aspects of auditory cognition, such as pitch. The additional aspects that we discuss later in this chapter revolve around the question of whether perception of aesthetic dimensions of prosody and analogous features in music are independent from more traditionally studied acoustic features. Our analysis will contribute to a more complete picture of the fundamental aspects of human communicative behaviours that have been hitherto overlooked in cognitive science.

Although scholars have reflected on the shared properties of music and language since at least the early eighteenth century, the parallel consideration of language and music within cognitive science has grown significantly during the last decades, with the new aim of understanding the neural mechanisms underlying these auditory experiences, and linking them thereby to perceptions of emotional states and other communicative content. This research has employed a range of methodological strategies employing behavioural and electrophysiological measures, functional neuroimaging techniques and the study of brain damaged individuals (Patel, 2012). Knowledge regarding music and speech processing acquired through these studies can potentially lead to new models for the neurophysiological and cognitive organisation of such processes. One dimension of interest is the emotional quality of these auditory experiences. Here, we add an additional dimension termed "expressiveness" which captures the aesthetic quality of the individual performance of speech and music beyond its emotional content. We propose to explore various prosodic aspects of these auditory experiences, including emotion and expressiveness, to provide a more nuanced understanding of the listener's perceptual experience of music and speech.

Before considering the empirical evidence regarding the relationship between auditory speech and music perceptual experiences, we review the formal properties of these two domains and compare the perceptual features of speech and music as acoustic events. Both speech and music encompass acoustic streams that are perceived in terms of pitch, timing, loudness and the space in which they occur (Griffiths et al., 1999) and variation across these characteristics bears information about their source of production. For example, a range of higher pitches can be indicative of smaller size, as in a child's voice rather than an adult one, or of a violin rather than a bass. Despite such shared acoustic components, they serve different functions in speech and music.

Researchers in the mid-twentieth century have determined that in speech the expression of emotion is conveyed through a number of interacting prosodic cues that involve changes, not only in duration as mentioned above, but also in fundamental frequency (F_0), intensity, and voice quality (Crystal, 1969). For example, Scherer described the rise of mean, range and variability of F_0 in "active" emotions such as anger, fear, and happiness, while "passive" emotions such as sadness were characterised by a decrease in fundamental frequency (Scherer, 1986). As such, some universal ability to identify emotions in speech does appear to exist, but great variability has been found among speakers of different languages (Scherer et al., 2001). The perception of emotional expression in music through variations in pitch, rhythm, duration, and tonal intensity has received much debate since the early eighteenth century (for historical perspectives on psychological approaches and emotional responses to music see chapters in Part I). Many of the twentieth century experimental investigations were largely limited to "everyday" emotions corresponding to experiences labelled as "happy", "sad", or "angry (e.g., Heilman et al., 1984; Ross, 1981; Shapiro and Danly 1985). However, there is new interest amongst empirical researchers in "aesthetic" emotions that may be experienced in regard to an artistic production or the performer's skill (Juslin, 2013). The references to the literature reviewed here mainly correspond to "everyday" emotions, while the theory of "expressiveness" we propose relates to "aesthetic" emotions.

Music and language display somewhat divergent acoustic properties and different biological and cultural foundations, rendering the study of the listener's experience a challenging task. Cross also highlights the differences in the state of research evidence in the two domains, drawing attention to the different sizes of historical and cross-cultural corpora that are available (Cross, 2010). While research into the universal aspects of language has been based on evidence arising from a wide range of cultures and time periods, Cross suggests that similar considerations of the possible universal properties of music is more problematic because the comparable body of evidence is much more limited in representing the cultural and historical diversity of musical forms, with most investigations focusing on the modern Western musical canon.

The language-music comparison also poses challenges if viewed from a biological perspective. Language acquisition appears to be constrained by maturational development which is sensitive to environmental input at different ages and stages. While there is evidence of a similar process of development for music cognition in infants (e.g., Trehub, 2001), a similar critical period sensitivity has not been so far been demonstrated (Trainor, 2005). From an evolutionary perspective, music has been hypothesised to relate to social cohesion and wellbeing, preceding language evolution. For example, Dunbar points to the endogenous opioid release effect on our ancestors during communal singing thought to promote social cohesion (Dunbar, 2003). At the same time, singing is argued to be cognitively simpler when compared to the complex computational devices found in language, leading to the proposal that music is an evolutionary precursor to language. Such considerations that blend biological and cultural factors appear to be indispensable in the understanding of these human behaviours.

The considerations in this chapter are limited to evidence stemming from Western tonal and metred instrumental music, whose conventions do not necessarily apply to all existing musical systems. Whether our observations might hold more generally, and therefore reflect more universal properties of music, must be validated through cross-cultural investigation. Similarly, current understanding of speech perception research reflects evidence from a limited number of languages and may not pertain to all linguistic systems. It is yet to be determined how cultural influences shape biological foundations for such auditory perceptual processes. At present, the parallel study of language and music, even within a given musical and linguistic system, is subject to the tension between cultural relativism and the search for empirically testable universals.

Seeking the roots of a "contemporary" comparison

Although research investigating the relations between speech and music processing has intensified in recent decades, the link between the two human productions has been explored since ancient times. The intrinsic connection between the body and music in regard to emotion was considered in successive historical periods from the Ancient Greeks through Medieval European scholarship and Enlightenment philosophers and physicians (see chapters in Section I of this book). For example, with respect to rhythmic organisation of speech, the American experimental psychologist Thaddeus L. Bolton, in one of the earliest scientific studies in music education, suggested that specific rhythmic relationships did not arise naturally in speakers of Greek and Latin, but that they were developed as a result of the high value placed on the aesthetic aspects of theatre performance of poetry and music in those cultures (Bolton 1984).

There were also theoretical considerations of the comparisons between speech and music in ancient Greek scholarship. Aristoxenus, a fourth-century BC pupil of Aristotle, made observations of differences in how pitch is employed in speaking and singing. He noted that the movement from one pitch point to the next is continuous in spoken Greek (*syneches*), whereas, in singing, pitches are discrete (*diastematike*), becoming more stationary in comparison to speech (Anderson, 1973).This history reflects the musical presentation of verse in ancient Greece mentioned above. It is related to the ancient Greek "prosoidia", used to describe pitch variations in spoken poetry. This distinction is still relevant today, as individuals with neurogenic music deficits are shown to have greater difficulty in recognising pitch direction in discrete pitches even when these are artificially applied to speech stimuli (Liu et al., 2012).

Rhetoric encompasses the techniques employed by individuals trained to inform and to persuade, called *rhetors* in classical ancient Greece. The art of rhetoric was not viewed as an aspect of spontaneous production but, rather, had to be practised and mastered according to specific principles. The division of rhetoric into five subcomponents was described by ancient Greek and Roman writers, namely *invention* (finding the argument), *disposition* (ordering the argument), *elocution* (style), *memoria* (memory), and *pronuntiatio* (delivery) with the threefold aim to move, delight, and instruct the audience. Quintilian (Marcus Fabius Quintilianus) argued that a skilful orator must have a good understanding of music principles, thus explicitly drawing parallels to the properties of the two domains. Interestingly, technical terms describing speech prosody have etymological sources in music. For example, "accent" indicates the presence of acoustic stress and emphasis in speech. The English word is derived from the Latin *accentus*, meaning song added to speech; the stem *cantus* means singing.

In the Renaissance, the relation between speech and music was a topic of great interest. For example, the English philosopher Francis Bacon (1561–1626) emphasised the parallel nature of speech and music: “So again a man should be thought to dally, if he did note how the figures of rhetoric and music are many of them the same. The repetitions and traductions[1] in speech, and the reports and hauntings of sounds in music, are the very same things” (Bacon, [1603] 2000). Indeed, rhetorical practice had a great influence on music throughout the Baroque period. There was great interest in the qualities of successful communicative practice and the power to move the listener in both oration and musical performance-- *pronuntiatio* (delivery), as mentioned above. This debate about the relationship between music and rhetoric was extensive in the eighteenth century. Beyond the question of shared qualities of music and speech in performance, or in their essential compositional structure, there was also interest in considering their human origin from the seventeenth century onwards (Besson and Shon, 2003).

One very notable contribution to the understanding of the more physiological instantiations of music and speech as human behaviours was by the English physician Thomas Willis (1621-1675). In his book *Cerebri anatome* (1664), he developed an original neurological model for the physiological basis of music as motor expression, suggesting it was both natural and universal (Lorch, 2010). In the context of the Western Christian tradition in the seventeenth century, such a proposal regarding the mind would invoke an account of the human soul. Willis’ major innovation was to detail the corporeal nature of the “sensitive soul”, that along with the “vital soul”, controlled sensation and motion, knowledge and simple reasoning. Although earlier ideas about the sensitive soul had been put forward from Aristotle onwards, the sensitive soul had been constructed as an incorporeal entity that had effects on the

[1] The Oxford English Dictionary gives one definition of traduction (archaic) in relation to rhetoric as the successive repetitions of a word in various forms, or closely related words for rhetorical effect. "traduction, n.". OED Online. June 2017. Oxford University Press. http://www.oed.com.ezproxy.lib.bbk.ac.uk/view/Entry/204328?redirectedFrom=traductions (accessed January 15, 2018).

material world. Willis theorised that both animals and humans had a corporeal sensitive soul, while humans also had an (immaterial and immortal) "rational soul". Willis further described a material link between them: the intercostal nerves, which he termed the "reins of the soul". By this means, the "passions" of the body were connected to the mind, and through them, human actions could be directed by rational judgment.

Willis's work is exceptional with regard to an extensive neurological treatment of music as a human behaviour which encompassed the passions and the rational soul. While describing the anatomy of "hearing nerves" and their connection to the brainstem, cerebellum and cortex, he also considered the neurological storehouse of the ideas of sounds, and the process of remembering musical melodies. In his later book, *De anima brutorum* (1672) Willis goes on to offer his observation that some young children are skilled singers before they are able to speak fluently. Moreover, Willis suggested that musical ability is distinct from other aspects of memory, reasoning or intellect. He believed that a physiological difference between individuals would explain why some have "musical ears" and others are "wholly destitute of the faculty of Musick." Finally, he considered the direct and profound effect of music on the "passions" in humans, but not animals, and ascribed a physiological source to this difference. In his 1664 neuroanatomy book, he suggests "the Melody introduced to the Ears … does as it were inchant with a gentle breath the spirits there inhabiting, and composes them … and so appeases all tumults and inordinations therein excited" (Willis, [1664/1681] 1971, p. 97). Willis's treatment of the neurological basis of musical abilities appears to be unique in seventeenth-century neurological thinking. However, he did not develop parallel arguments for the neurological basis of language, nor did any others until the nineteenth century, although the nature of auditory perception and the effects of sound on the mind and body had been considered since the Enlightenment (Gouk, 2004).

In the second half of the nineteenth century, interest in theoretical ideas on the evolution of, and parallels in, music and language focused specifically on communicative and emotional expression (Spencer, 1858; Darwin, 1872). This underpinned a renewed interest in determining the nature of the language and music faculties. The study of how brain damaged individuals responded to linguistic and musical tasks were seen as test cases for developing a general account of brain function and human behaviour in the nineteenth century (Johnson et al., 2010).

In the 1860s, the English neurologist John Hughlings Jackson (1835-1911) drew a neurological distinction between what he termed "propositional" language and emotional vocal expressions (Jackson, 1866), while his colleague Henry Charlton Bastian (1837-1915) began to routinely test musical abilities in patients with acquired disorders of language (Lorch, 2013). Furthermore, the ability to hum or sing was noted to be preserved in some individuals with little or no spoken verbal expression (Lorch and Greenblatt, 2015). Behavioural observations of such patients led to the construction of theoretical models that explained how language and music processing might be instantiated in different parts of the brain. The French clinician Jean-Martin Charcot (1825-1893) provided some of the earliest detail theories of how the brain processes music. In the 1880s, he developed an account of music processing drawing a parallel structure to his model for language, and described the dissociations of deficits for both the perception and production of these two domains in his neurological patients (Johnson, et al., 2013). The German neurologist August Knoblauch (1863-1919) (1888) coined the term "amusia" to describe individuals with an impairment in musical abilities, in an analogy to the term "aphasia" that had been defined two decades earlier to describe those with impairments in language.

By the end of the century, many cases of impairments dissociating music and language had been collected (e.g. Edgren, 1894) and there were growing numbers of observations of both acquired and congenital music disorders. Hence, the clinical research in the late nineteenth century displays a relatively systematic investigation of music alongside language. This suggests that they were viewed as having equal status as mental faculties, and were demonstrated to be overlapping in function and vulnerability to impairment through brain damage. Moreover, their parallel examination may also be viewed as signalling their equal significance as valued behavioural repertoires whose compromise would be taken as serious signs of ill health. The fact that music had such a privileged social status in the nineteenth century Western world is perhaps in contrast to the present day, where the importance of music to an individual's life experience is expected to be somewhat more marginal from a cognitive perspective.

In the second half of the twentieth century, there was a resurgence of interest in the cognitive basis of verbal and emotional expression with the advent of neuropsychological studies investigating laterality of function (e.g., Gorelick and Ross, 1987; Borod et al., 1985). Behavioural research with brain damaged patients indicated that in aphasic individuals with left hemisphere damage the linguistic content of their speech production was impaired while their prosodic expression was intact. However, the reverse picture was shown in patients with right hemisphere damage. Those with difficulties in expressive speech showed a range of abilities to express emotional content by employing a variety of non-linguistic vocal resources. These intact channels of expression were thought to contribute to the aphasic individuals' conversational success even though their ability to convey meaningful content was compromised (Lorch et al., 1999). Furthermore, renewed interest in the preservation of the ability to sing in aphasic individuals led to models of distributed hemispheric processing of language and music and the development of such speech rehabilitation techniques as Melodic Intonation Therapy (Albert et al., 1973). Developmental research investigating innate properties of language also considered the role of pitch and prosody in very young infants. This strand of research revived earlier nineteenth-century questions about the evolution and development of language with new models of the appreciation of melody as an underpinning for infant speech perception (Trehub, 1989; Trehub et al., 2013).

Cognitive investigation of auditory experiences of music and speech expanded in the 1980s through the methodological innovations offered by computer based manipulation and analysis of acoustic stimuli. However, the increasing use of more advanced technical manipulations in recent research may also be somewhat problematic. While the use of auditory stimuli with unnatural acoustic manipulations may be useful for investigating the auditory abilities of listeners, they may not be appropriate for research into individuals' music cognition (Bigand and Poulin-Charronnat, 2006). This observation calls for the use of more environmentally valid stimuli. In line with this view, examination of perception ~~of “~~of “expressiveness” in music and speech requires the use of naturally-occurring stimuli, as it involves the genuine aesthetic appreciation of real auditory objects in the human environment.

At the beginning of the twenty-first century, there have been renewed efforts to explore the relations between speech and music (Ayotte et al., 2002; Nicholson et al., 2003; Foxton et al., 2004; Patel et al., 2005). There is growing interest in the dimension of emotion with regard to acoustic expression. In a review of 104 studies of vocal expression and 41 studies of music performance, Juslin and Laukka found similarities in the accuracy of listener's ability to identify discrete emotions through emotion-specific patterns of acoustic cues in both vocal and musical stimuli (Juslin and Laukka, 2003).

With regard to the broader enterprise of determining the inter-connections between language and music beginning in the late twentieth century, Fodor's model of modularity has been instrumental in providing a conceptual foundation (Fodor, 1983). According to this model of cognitive processing, the mind consists of modules that are domain-specific, informationally encapsulated; they are, an innate endowment with neural specificity. Thus, different cognitive components function independently. Stimuli will engage a specific modular processor relative to their featural properties, while certain stimuli may be processed by different modules depending on contextual variables. Much of this work has focused on the perception of pitch, which has been shown to be processed as either music or language relative to a given acoustic context. The processing of acoustic experiences by infants, for example, can be conceptualised as either music or speech leading to different hypotheses about the development of such modules (McMullen and Saffran, 2004). A variety of investigations regarding the processing of pitch have also explored such distinctions in healthy adults and those with developmental or acquired difficulties of speech and music (see section 'Language, music, and cognitive impairments').

Pitch differences in speech and music

Important differences in the structuring of pitch sequences in speech and in music must be considered when assessing the auditory perception of these sound streams. For example, a melody in Western tonal music relates to specific sets of pitches (Steinke et al., 2001), and the insertion of pitches that do not belong to a given set is perceived as incongruous by the listener. The structural constraints for pitch and timing in music make it, in this sense more predictable than speech. In contrast, the organisation of speech into pitch events does not exclude the use of some pitches in favour of others.

While pitch sequences in speech are continuous, they have a "stair-stepped" organisation in singing within the Western classical music system (Bidelman et al., 2009). The transition from one pitch event to the next in speech can be understood as a smoother process compared to singing where the voice seems to "pause" slightly in every interval. This is, of course, even more pronounced for musical instruments, although some variation exists depending on the type of instrument and the mechanism that generates its sound. Another important difference between pitch across these two domains is the distance between pitch events. That is, as a melody unfolds, the interval types between successive notes differ from those of speech. In the Western musical tradition (before the 1950s), a semitone is the smallest interval encountered in notated composition and, at the same time, it appears very frequently (Vos and Troost, 1989). The pitch organisation of speech differs substantially and the occurrence of these small intervals is not as common (Fitzsimons et al., 2001). Moreover, pitch sequences in speech are characteristically perceived as simple rises and falls.

The relativistic nature of pitch variation in speech can be juxtaposed with the determined nature of pitch representation in music. For example, a small manipulation of pitch can lead to a complete shift in the mood or emotional quality of a music piece, turning a tune perceived as happy into a sad one (McDermott and Oxenham, 2008), whereas in speech no such effect is common. Thus, the appreciation of musical stimuli highly depends on precision in pitch (Zatorre and Baum, 2012). This difference in terms of accuracy can be explained through a principle of pitch organisation that exists in music but not in language, that of tonality. Musical key can be understood as a probability distribution that determines the frequency of appearance and the duration of notes in relation to other notes (McDermott and Oxenham, 2008). Hence, the various pitches belonging to a set are organised in harmonic relationships that determine which sequences are allowed, more "appropriate", or more important in any given tonal context. As such, pitch sequences in music can be thought of as organised around a centre of gravity that determines which parts of the sequence are more stable and which are less stable (Tillmann et al., 2000).

Although juxtaposing music melody with the pitch contours of speech appears to be the most legitimate comparison to be made, evidence from Ross et al. suggests that the preference for specific tonal intervals may be linked to the formant properties of vowels in speech (Ross et al., 2007). Similar to the harmonics associated with musical instruments, formants constitute additional frequencies above the F_0 that, in the case of speech, vary according to different configurations of the vocal tract. Evidence presented in Ross et al. (2007) suggests that our preference for some intervals in music arises from the frequency relationship of the first two formants in vowels. This finding may shed light on the origins of musical preference in humans, but also indicates additional complexity to the parallel study of pitch across music and speech domains.

Prosody in language and music: parallels and differences

The role of pitch is also pivotal in understanding how syntax is organised in music. Despite some similarities found in cognitive experiments on the perception of syntactic violations in both music and language (e.g., Patel et al., 1998, Maess et al., 2001), there are structural differences in the two domains. Both language and music can be analysed as comprising small building blocks combined into larger structures governed by structural principles that define the relationship of each unit or phrase to adjacent ones. However, linguistic syntax has additional structural properties that have no equivalent in music, such as the morphological variants indicating lexical and grammatical functions for different parts of speech. For example, the English morphemes “-ing” or “-ly” have a very different componential status compared with elements such as “*girl*” or “*red*” in the phrase structure of an utterance. Another syntactic distinction is at the formal level, as music structure is not underpinned by predicate-argument relationships present in language. Parallel properties in musical phrasing are more difficult to identify.

Prosodic realisations in speech have been associated with functions for communicating syntactic, informational, interactive, attitudinal, and emotional functions (Vaissière, 2008). More specifically, prosodic information can convey grammatical distinctions, such as indicating a question with a rising tone. Other prosodic signals may reflect conversational conventions, such as turn taking or concluding an exchange, or speaker beliefs (for instance, truly believing in an idea or expressing disbelief and doubt). In addition, prosody may serve as a signal of speaker mood, emotion or affect with flattened prosody typically signalling sadness or depression (Alpert et al., 2001). Three main prosodic categories are typically described in speech—linguistic, emotional, and pragmatic. We will consider whether these prosodic labels provide a sufficient description of prosodic realisations in speech.

Beyond their basic grammatical functions, prosodic cues are also employed in order to emphasise syntactic structures or to disambiguate structures. Duration and pitch patterns as well as pausing help the listener to group words into constituent structures by identifying prosodic breaks (Wightman and Ostendorf, 1994). Manipulation of these features may also lead to varying degrees of emphasis or prominence (Skandera and Burleigh, 2005). Prosodic cues are also used to emphasise or disambiguate meanings. In contrastive stress prosody, the raising of pitch, duration, or loudness may be used to signal a topic/focus. For example, in the utterance “John gave her flowers.” the Subject “John” would typically be the topic and bear main prosodic stress. In other cases the Object “flowers” might be emphasised prosodically to indicate that this was unexpected, as in the context where the recipient was known to dislike receiving such a gift. Music prosody displays some similar use of prosodic force that facilitates the segmentation of an acoustic signal into smaller parts and places emphasis on important events (Palmer and Hutchins, 2006).

Prosodic features also reveal the emotional state of the speaker. As with linguistic prosody, emotional prosody also employs pitch and timing patterns to denote the speaker's emotional disposition (Pell, 2006). As speech uses emotional prosody in order to convey different emotions, a music passage can be executed with prosodic variation in order to communicate emotions (Juslin and Laukka, 2003). However, music does not normally have a semantic aspect. That is, in contrast to verbal symbols which have denotative meaning, musical sounds typically do not carry explicit denotations but may express ideas through context (Meyer, 1956). Admittedly, exceptions such as symphonic poems and operatic musical devices do exist (e.g., Prokofiev's 1936 "Peter and the Wolf"). These intend to convey description and narrative by employing sound symbolism. Music can also possess some general communicative force through its emotional character. In this respect, a parallel can be drawn between music and speech prosody, as prosody carries information that extends beyond word identity (Cowie et al., 2001), and emotion in music has an attitudinal rather than a semantic meaning.

A third type of prosody, pragmatic prosody, extends beyond basic syntactic and emotional signal functions but also by employing variation in pitch, duration, rate, and intensity. A tone may be considered as ironic when the prosodic features in a given utterance appear in sharp contrast to those that would be typically expected in a pragmatically neutral utterance matching its literal meaning (Attardo et al., 2003). An analogy may be drawn between language discourse and music discourse. Particular musical devices, such as rhythmic relationships or harmonic progressions, can turn into formulas that promote efficient communication, evoking customary responses (Meyer, 1956). However, this parallel fades when adding the pragmatic dimensions in speech of interlocutors' beliefs, social identities and relationships that have no exact counterparts in the musical domain.

Language, music, and cognitive impairments

The study of music perception impairments is of particular interest for the understanding of the cognitive relationship between prosodic features in music and language (Patel, 2008). The following section reviews the evidence of how cognitive impairments can result in difficulties in the perception of speech and music sound streams. For example, pitch and its perceptual organisation from "low to high" is similar across speech and music, and brain damage in the right, but not left, hemisphere can lead to impaired judgement of the direction of pitch changes in both linguistic and music stimuli (Johnsrude et al., 2000; Meyer et al., 2002). In order to investigate prosodic features in larger speech streams, technical manipulations isolate the melodic and rhythmic content of a spoken utterance from its lexical semantic content. This precludes the engagement of cognitive mechanisms responsible for processing linguistic meaning. In one such investigation, Nicholson et al. reported that their participant with right fronto-parietal brain damage had equally compromised ability on both speech-like and music stimuli when lexical information had been removed (Nicholson et al., 2003).

In other studies, researchers have looked at the linguistic abilities of individuals with congenital amusia, a developmental difficulty with music appreciation and performance (Kalmus and Fry, 1980; Peretz and Hyde, 2003). These individuals feel less positive about musical experiences, often devote relatively little time to music listening and some report experiencing more limited psychological changes in response to musical stimuli (McDonald and Stewart, 2008). Individuals with congenital amusia have severe difficulties with basic pitch appreciation (e.g., Ayotte et al., 2002; Foxton et al., 2004; Hutchins et al., 2010). The diagnosis of amusia is typically based on poor performance on tasks requiring judgements of differences in melodic and/or rhythmic patterns (Peretz, Champod, & Hyde, 2003). This pitch difficulty has been shown to extend to the perception of speech patterns as well. However, Liu et al. suggest that congenital amusics' poor perception of intonation contours in de-lexicalised speech stimuli may be due to manipulations which caused pitch events to lose their gliding structure and acquired a more discrete nature causing the speech based stimuli to be perceived as music (Liu et al., 2012). Correlations between musical pitch processing impairments and acoustic impairments in amusics have been found across a large number of studies but further inspection of individual cases point to several exceptions (Vuvan et al., 2015). It therefore appears that some experimental techniques may require additional refinements in order to avoid artefacts such as these from technical acoustic manipulations that may confound the relationship between speech and music.

The assessment of the musical ability of amusics based on a limited set of tasks examining the appreciation of differences in pitch, contour, interval, rhythm, and meter does not take into consideration an important aspect of music engagement and appreciation. In the following section, we present a prosodic aspect present in both music and speech. We argue that this prosodic dimension has a dynamic quality extending beyond those isolable acoustic features that have been systematically investigated in the literature.

New avenues for assessing perception of "expressiveness"

As outlined earlier in this chapter, rhetoric devices in speech and music have been explicitly linked at various points in history. In the Baroque period, elements of rhetoric such as tone colour, phrasing and timing were ambiguous and their use was normally at the discretion of the performer and their ability. We argue that such additional prosodic aspects of speech and music should be included in the neuroscientific inquiry into the listener's experience of the performance. Despite the systematic investigation of the perception of basic prosodic elements in individuals with compromised cognitive abilities over the past fifty years, this shared aesthetic prosodic dimension of music and speech has not yet been addressed in the cognitive and clinical domains.

In the present research context, we argue that while componential analyses of acoustic features have contributed a great deal to the understanding of music and speech cognition, they may overlook more dynamic aspects of the appreciation of human acoustic streams. We suggest that there are additional dimensions of prosodic expression that are not accounted for in the existing prosodic elements that have been investigated. Aesthetic appreciation of speech and musical acoustic streams includes the perception of qualities that go beyond aspects of emotion or structural well-formedness. We characterise a useful additional aspect of the listener's experience which will further the understanding of the acoustic properties embodied in the appreciation of music and speech as communicative systems. This novel approach encompasses additional elements of expression related to the dynamic gestalt-like qualities of acoustic phrases that convey aesthetic properties in speech and music. Here we define this additional prosodic aspect as "expressiveness" and this term will be used throughout this section.

Studies on soloists' interpretation of pieces of music identify such aspects of prosodic expression (Repp, 1995). Each soloist will play in a distinctively different way from the same textual source. Even if their technical profile (fingering, muscle control etc.) is equal to other musicians with the same level of proficiency their performances will inevitably vary. These differences may be due the degree to which a particular acoustic manipulation or effect is employed by a given performer or how these acoustic cues are combined. Hence, choices in interpretation are not identical. This is evident, for example, in pianists who employ different styles of expression that can be perceived as differentiating their individual performance from others (Sloboda, 1983). This suggests that despite some generally accepted constraints that govern soloists' performance, a great deal of their rendition reflects personal choice. When studying soloists' performance of the unmeasured prelude, where durations are not notated, Gingras et al. found performers to exhibit large variation in expressive timing around parts of structural unexpectedness (Gingras et al., 2016). This finding suggests that when duration is not determined by a music score, performers are likely to produce a wide range of patterns of expression. Such choices cannot be deemed as right or wrong, but as more or less preferred (Palmer and Hutchins, 2006). Although identifying the particular component parts that comprise these variations in performance might seem hard to determine, this is easily contrasted with the sound of more "neutral" performances, such as those coming from music notation playback sound files which lack temporal deviations, loudness variation and timbre changes. This opens up new methodological opportunities for a systematic consideration of such differences, which are well recognized in performance studies, but overlooked in the cognitive science of music.

A "deadpan" performance can be thought of as the least possible expressive performance. In such a performance, all notes correspond to the exact durations imposed by the score while their loudness displays absolutely no variation across a piece of music (Sloboda, 1983). In speech, a "deadpan" style of speaking might include little variation of prosodic features in comparison to a dramatic performance. In a closer analogy to various renditions of a given music score, an inexperienced actor's literal and conventional rendering of the text will contain little prosodic variation in comparison to an experienced and talented actor playing the same role who will use a variety of prosodic manipulations to embellish the words of the playwright. In the case of a written text, loudness, duration and pausing patterns would not be as strictly prescribed as in a music score, but the acoustic properties of the inexperienced and the experienced performers' will differ in terms of prosodic richness. For example, consider the wide variation in actors' prosodic choices in renditions of Hamlet's "To be or not to be…" speech in Shakespeare's play. This captures the aesthetic dimension of prosodic variation in speech and music which is perceived by the listener. The question is what perceptual mechanisms are involved in appreciating such differences that seem to be fundamental to the qualitative experiences of music and speech streams.

It is of interest to examine possible reasons why some elements or "techniques" which are used in speech are perceived as expressive. For instance, deviations from temporal regularity in music can be perceived as musically expressive (Large and Palmer, 2002). In language, although there are minimal constraints on pitch, intensity and duration variation in speech beyond which utterances will sound distorted, there is no such thing as temporal regularity or a default rhythmic structure. Rather, those who attempted to capture isochronic dimensions of speech in typological classification systems have often based their determinations on listeners' impressions rather than objective acoustic measures of speech (Roach, 1982). However, even in music Fitch notes that isochronicity is a relative feature, as no music is shown to be completely isochronous (Fitch, 2006). One could argue that deviation from temporal regularity in music might be perceived as expressive, due to its speech-like quality. That is, the intentional deviation from the assigned rhythm would perceptually resemble speech where no such restrictions exist. As speech rhythm does not have prescribed organisation–in contrast to meter in Western tonal music, it might represent a more free form of expression. For this reason, a musical interpretation that shifts away from faithfully following its measure might represent the performer's intuitive choice to mimic speech prosody.

In language, it is also useful to consider the written representation of the sentence and its acoustic realisation in a spoken utterance. If a text is read aloud, there is the potential for a wide range of variation in suprasegmental features of the oral production, for example, in the renditions of audiobooks. The acoustic realisations of a written sentence are only limited to a very minor extent by syntactic and discourse pragmatic constraints. These are signalled to a small degree by punctuation. However, a wide range of emotional and social information is also conveyed by individual utterances. There has been only limited exploration of such aspects of the acoustic signal within the domain of linguistics, although this has been long recognised to be a significant factor in the dramatic arts. Various studies have documented the unpredictability of acoustic realisation of written language in terms of durational patterns of syllables or pauses (e.g., Brown and Miron, 1971; Ferreira, 1991; Gee and Grosjean, 1983). Hence, although syntactic rules are specific and describable, the acoustic properties in the oral rendition of sentences displays richer acoustic elements than those that can be predicted based on the prosodic-syntactic relationship.

It does appear that listeners can and do appreciate and distinguish between various expressive renditions on an aesthetic level, which is not simply to do with emotional content. This suggests that comparisons at the level of expressiveness need to be sought across domains, but also in relation to other perceptual processes that have already been systematically explored with regard to prosody. In a recent study we showed that congenital amusia does not necessarily deprive an individual of the ability to appreciate aesthetic aspects of music (Loutrari and Lorch, 2017). We tested one individual diagnosed as congenital amusic and found that she was able to perceive expressive features of music and speech prosody. This amusic participant was able to differentiate expressive from non-expressive melodies. The two melody categories differed simultaneously in several acoustic features: temporal regularity versus slight deviation from it, variation in dynamics versus uniform dynamics, and connected transitions between notes versus heavily accented notes. Our participant scored normally on this task despite severe difficulties with pitch perception. Such dissociations in performance raise questions regarding the evaluation of components of music cognition in current amusia research. We suggest that examining the perception of expressiveness in individuals with congenital and acquired music impairments can contribute substantially to our picture of music cognition and its disturbances.

Oratory style (elocution) and rhetorical practice have been valued as an instrument to move the listener in various historical periods. However, present day cognitive studies of language and music do not seem to view these elements as integral parts of our perceptual mechanisms. By including consideration of expressiveness in the investigation of music and language cognition, we suggest that new associations and dissociations in individuals with amusia may be identified. This approach can cast some light on the relationship of expressiveness to the perception of other aspects of the acoustic sound stream that have previously been the focus of research. The future objective will be to develop a clearer determination of the nature of expressiveness and to explore whether it should be considered a single perceptual entity or depends on a collection of acoustic features yet to be determined. This will lead to a more nuanced picture of the individual listener experience of such socially mediated auditory stimuli as comprised in music and speech. Further questions regarding the extent to which expressiveness is a construct determined by culturally specific traits and/or has underlying universal principles can add to our understanding of the nature and origin of this channel of human expression.

Works cited

Albert, M. L., Sparks, R. W., Helm, N. A. (1973). Melodic intonation therapy for aphasia. *Archives of neurology*, 29 (2), pp. 130-1.

Alpert, M., Pouget, E. R., Silva, R. R. (2001). Reflections of depression in acoustic measures of the patient's speech. *Journal of affective disorders*, 66 (1), pp. 59-69.

Anderson, W. D. (1973). Word accent and melody in ancient Greek musical texts. *Journal of music theory*, 17, pp. 186–203.

Attardo, S., Eisterhold, J., Hay, J., Poggi, I. (2003). Multimodal markers of irony and sarcasm. *Humor*, 16, pp. 243-60.

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

Bacon, F. ([1603] 2002). *Valerius terminus: of the interpretation of nature.* Project Gutenberg, Internet resource. http://www.gutenberg.org/cache/epub/3290/pg3290-images.html (Accessed 16.10.2015).

Besson, M., Schoen, D. (2003). Comparison between language and music. In: I. Peretz and R. Zatorre, eds., *The cognitive neuroscience of music*. Oxford: Oxford University Press, pp. 232-58.

Bidelman, G. M., Krishnan, A. (2009). Neural correlates of consonance, dissonance, and the hierarchy of musical pitch in the human brainstem. *The journal of neuroscience*, 29, pp. 13165-71.

Bigand, E., Poulin-Charronnat, B. (2006). Are we “experienced listeners”? A review of the musical capacities that do not depend on formal musical training. *Cognition*, 100, pp. 100-130.

Bolton, T. (1894). Rhythm. *American journal of psychology*, 6, pp. 145–238.

Bonds, M. E. (1991). *Wordless rhetoric: musical form and the metaphor of the oration*. Cambridge: Harvard University Press.

Bonds, M. E. (2014). *Absolute music: the history of an idea*. Oxford: Oxford University Press.

Borod, J., Koff, E., Perlman, M., Nicholas, M. (1985). Channels of emotional expression in patients with unilateral brain damage. *Archives of neurology,* 42, pp. 345-8.

Brown, E., Miron, M. S. (1971). Lexical and syntactic predictors of the distribution of pause time in reading. *Journal of verbal learning and verbal behavior*, 10, pp. 658-67.

Cowie, R., Douglas-Cowie, E., Tsapatsoulis, N., Votsis G., Kollias, S., Fellenz, W., Taylor, J. (2001). Emotion recognition in human-computer interaction. *IEEE signal processing magazine*, 18, pp. 32-80.

Cross, I. (2010). The evolutionary basis of meaning in music: some neurological and neuroscientific implications. In: F. C. Rose, ed., *The neurology of music*. London: Imperial College Press, pp. 1-15.

Crystal, D. *Prosodic Systems and Intonation in English*, 1969. Cambridge: Cambridge University Press.

Darwin, C. (1872). *The expression of the emotions in man and animals*. London: John Murray.

Dunbar, R. I. (2003). The origin and subsequent evolution of language. In: M. H. Christiansen and S. Kirby, eds., *Studies in the evolution of language*. Oxford: Oxford University Press, pp. 219-34.

Edgren, J. G. (1895). Amusie (musikalische Aphasie). *Deutsche Zeitschrift für Nervenheilkunde*, 6, pp. 1-64.

Ferreira, F. (1991). Effects of length and syntactic complexity on initiation times for prepared utterances. *Journal of memory and language*, 30, pp. 210-33.

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

Fitzsimons, M., Sheahan, N., Staunton, H. (2001). Gender and the integration of acoustic dimensions of prosody: implications for clinical studies. *Brain and language*, 78, pp. 94-108.

Fodor, J. A. (1983). *The modularity of mind: an essay on faculty psychology*. Cambridge: MIT Press.

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, pp. 801-10.

Gee, J. P., Grosjean, F. (1983). Performance structures: a psycholinguistic and linguistic appraisal. *Cognitive psychology*, 15, pp. 411-58.

Gorelick, P. B., Ross, E. D. (1987). The aprosodias: further functional-anatomical evidence for the organisation of affective language in the right hemisphere. *Journal of neurology,* 50 (5), pp. 553-60.

Gouk, P. (2004). Some English theories of hearing in the seventeenth century. In: M. M. Smith, ed., *Hearing history: a reader.* Athens: The University of Georgia Press, pp. 136-50.

Griffiths, T. D., Flees, A., Green, G. G. (1999). Disorders of human complex sound processing. *Neurocase*, 5, pp. 365-78.

Heilman, K. M., Bowers, D., Speedie, L., & Coslett, H. B. (1984). Comprehension of affective and nonaffective prosody. *Neurology*, *34*(7), pp. 917-917.

Hutchins, S., Zarate, J. M., Zatorre, R. J., Peretz, I. (2010). An acoustical study of vocal pitch matching in congenital amusia. *The journal of the acoustical society of America*, 127, pp. 504-12.

Jackson, J. H. (1866). Clinical remarks on emotional and intellectual language in some cases of disease of the nervous system. *Lancet*, 1, pp. 174-6.

Johnson, J. K., Graziano, A. B., Hayward, J. (2010). Historical perspectives on the study of music in neurology. In: F. C. Rose, ed., *The neurology of music.* London: Imperial College Press, pp. 17-30.

Johnson, J. K., Lorch, M., Nicolas, S., Graziano, A. (2013). Jean-Martin Charcot's role in the nineteenth-century study of music agraphia and music aphasia. *Brain*, 136 (5), pp. 1662-870.

Johnsrude, I. S., Penhune, V. B., Zatorre, R. J. (2000). Functional specificity in the right human auditory cortex for perceiving pitch direction. *Brain,* 123, pp. 155–63.

Juslin, P. N. (2013). From everyday emotions to aesthetic emotions: towards a unified theory of musical emotions. *Physics of life reviews*, 10, pp. 235-66.

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

Kalmus, H., Fry, D. B. (1980). On tune deafness (dysmelodia): frequency, development, genetics and musical background. *Annals of human genetics*, 43, pp. 369–82.

Knoblauch A. (1888). Uber Störungen der musikalischen Leistungsfähigkeit infolge von Gehirnläsionen. *Deutsches für Archiv Klinische medecin*, 43, pp. 331-52.

Large, E. W., Palmer, C. (2002). Perceiving temporal regularity in music. *Cognitive science*, 26, pp. 1-37.

Liu, F., Xu, Y., Patel, A. D., Francart, T., Jiang, C. (2012). Differential recognition of pitch patterns in discrete and gliding stimuli in congenital amusia: evidence from Mandarin speakers. *Brain and cognition*, 79, pp. 209-15.

Lorch, M. (2010). Fools at musick: Thomas Willis (1621-1675) on congenital amusia. In: F. C. Rose, ed., *The neurology of music*. London: Imperial College Press, pp. 151-71.

Lorch, M. (2013). Examining language functions: a reassessment of Bastian's contribution to aphasia assessment. *Brain*, 136 (8), pp. 2629-37.

Lorch, M., Borod, J. C., Koff, E. (1998). The role of emotion in the linguistic and pragmatic aspects of aphasic performance. *Journal of neurolinguistics,* 11, pp. 103-18.

Lorch, M. Greenblatt, S. (2015) Singing in speechless children. In: E. Altenmuller, F. Boller and S. Finger, eds., *Music, neurology, and neuroscience: historical connections and perspectives.* Progress in brain research, vol. 216. New York: Elsevier, pp. 53-72.

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, pp. 1-11.

Maess, B., Koelsch, S., Gunter, T. C., Friederici, A. D. (2001). Musical syntax is processed in Broca's area: an MEG study. *Nature neuroscience*, 4, pp. 540-45.

McDermott, J. H., Oxenham, A. J. (2008). Music perception, pitch, and the auditory system. *Current opinion in neurobiology*, 18, pp. 452-63.

McDonald, C., Stewart, L. (2008). Uses and functions of music in congenital amusia. *Music perception: an interdisciplinary journal*, 25, pp. 345-55.

McMullen, E., Saffran, J. R. (2004). Music and language: a developmental comparison. *Music perception: an interdisciplinary journal*, 21, pp. 289-31.

Meyer, L. B. (1956). *Emotion and meaning in music*. Chicago: University of Chicago Press.

Meyer, M., Alter, K., Friederici, A. D., Lohmann, G. von Cramon, D. Y. (2002). FMRI reveals brain regions mediating slow prosodic modulations in spoken sentences. *Human brain mapping*, 17, pp. 73–88.

Michel, L. (2013). The role of rhetoric in music of the Baroque.
Conflict, transformation Creolization conference National Collegiate Honors Council, November 6-10, New Orleans, LA.

Nicholson, K. G., Baum, S., Kilgour, A., Koh, C. K., Munhall, K. G., Cuddy, L. L (2003). Impaired processing of prosodic and musical patterns after right hemisphere damage. *Brain and cognition*, 52, pp. 382–9.

Palmer, C., Hutchins, S. (2006). What is musical prosody? *Psychology of learning and motivation*, 46, pp. 245-78.

Patel, A. D. (2012). Language, music, and the brain: a resource-sharing framework. In: P. Rebuschat, M. Rohrmeier, J. Hawkins and I. Cross, eds., *Language and music as cognitive systems*. Oxford: Oxford University Press, pp. 204-23.

Patel, A.D. (2008). *Language, music, and the brain*. New York: Oxford University Press.

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, pp. 310-13.

Patel, A. D., Peretz I., Tramo, M., Labreque, R. (1998). Processing prosodic and musical patterns: a neuropsychological investigation. *Brain and language*, 61, pp. 123-44.

Patel, A. X., Gibson, E., Ratner, J., Besson, M., Holcomb, P. J. (1998). Processing syntactic relations in language and music: an event-related potential study. *Journal of cognitive neuroscience*, 10, pp. 717-33.

Pell, M. D. (2006). Cerebral mechanisms for understanding emotional prosody in speech. *Brain and language*, 96, pp. 221-34.

Peretz, I., Hyde, K. L. (2003). What is specific to music processing? Insights from congenital amusia. *Trends in cognitive sciences*, 7, pp. 362–7.

Peretz, I., Champod, A. S., Hyde, K. (2003). Varieties of musical disorders. *Annals of the New York Academy of Sciences*, 999, pp. 58–75.

Ralston, J. T. (1991). Some influences of accent structure on melody recognition. *Memory & cognition*, 19, pp. 8–20.

Repp, B. H. (1995). Expressive timing in Schumann's ''Träumerei:''an analysis of performances by graduate student pianists. *The journal of the Acoustical Society of America*, 98, pp. 2413-27.

Roach, P. (1982). On the distinction between ''stress-timed'' and ''syllable-timed'' languages. In D. Crystal, ed*., Linguistic controversies: Essays in linguistic theory and practice in honour of F. R. Palme*r. London: Arnold.

Ross, D., Choi, J., Purves, D. (2007). Musical intervals in speech. *Proceedings of the National Academy of Sciences*, 104, pp. 9852-7.

Ross, E. D. (1981). The aprosodias: Functional-anatomic organization of the affective components of language in the right hemisphere. *Archives of Neurology*, *38*(9), pp. 561-569.

Scherer, K. R. (1986). Vocal affect expression: a review and a model for future research. *Psychological Bulletin,* 99*,* pp. 143-65.

Scherer, K. R., Banse, R., Wallbott, H. G. (2001). Emotion inferences from vocal expression correlate across languages and cultures. *Journal of cross-cultural psychology*, 32 (1), pp. 76-92.

Shapiro, B. E., & Danly, M. (1985). The role of the right hemisphere in the control of speech prosody in propositional and affective contexts. *Brain and language*, *25*(1), pp. 19-36.

Spencer, H. (1858). *Essays: scientific, political and speculative*. London: Longman, et al., pp. 359-84.

Skandera, P., Burleigh, P. (2005). *A manual of English phonetics and phonology.* Tübingen: Gunter.

Sloboda, J. A. (1983). The communication of musical metre in piano performance. *The quarterly journal of experimental psychology*, 35, pp. 377-96.

Steinke, W. R., Cuddy, L. L., Jakobson, L. S. (2001). Dissociations among functional subsystems governing melody recognition after right-hemisphere damage. *Cognitive neuropsychology*, 18, pp. 411–37.

Tillmann, B., Bharucha, J. J., Bigand, E. (2000). Implicit learning of tonality: a self-organising approach. *Psychological review*, 107, pp. 885-915.

Trainor, L. J. (2005). Are there critical periods for musical development? *Developmental psychobiology*, 46(3), pp. 262-278.

Trehub, S. E. (2001). Musical predispositions in infancy. Annals of the New York Academy of Sciences, 930(1), pp. 1-16.

Trehub, S. E. (1989). Infants' perception of musical sequences: Implications for language acquisition. *Journal of Speech-Language Pathology and Audiology*, 13, 3-11.

Trehub, S., Plantinga, J., Brcic, J., Nowicki, M. (2013). Cross-modal signatures in maternal speech and singing. *Frontiers in psychology*, 4, pp. 1-9.

Vaissière, J. (2008). Perception of intonation. In: D. Pisoni and R.Ramez, eds., *The handbook of speech perception.* London: John Wiley & Sons, pp. 236-63.

Vos, P. G., & Troost, J. M. (1989). Ascending and descending melodic intervals: Statistical findings and their perceptual relevance. *Music Perception: An Interdisciplinary Journal*, *6*(4), pp. 383-396.

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

Wightman, C. W., Ostendorf, M. (1994). Automatic labelling of prosodic patterns. *IEEE transactions on speech and audio processing*, 2, pp. 469–81.

Willis, T. ([1664/1681] 1971). The anatomy of the brain. Reprint of the English translation by S. Pordage of the Latin text *Cerebri Anatome,*1664. Tackahoe, NY, USV Pharmaceutical Corp.

Zatorre, R. J., Baum, S. R. (2012). Musical melody and speech intonation: singing a different tune. *PLoS biology*, 10, e1001372.