

Language, music, and the brain: a resource-sharing framework

Aniruddh D. Patel

The Neurosciences Institute

10640 John Jay Hopkins Drive

San Diego, CA 92121

apatel@nsi.edu

In press in: *Language and Music as Cognitive Systems*. P. Rebuschat, M. Rohrmeier, J.

Hawkins, & I. Cross (Eds.). Oxford: Oxford University Press.

Originally a keynote paper for “Language and Music as Cognitive Systems”

May 11-13, 2007

Cambridge University

England

Draft of Dec 10, 2009

1. Introduction

This paper discusses a theoretical framework for the cognitive study of music-language relations called resource sharing. Resource sharing makes a basic conceptual distinction between domain-specific knowledge and shared neural resources that operate upon this knowledge as part of cognitive processing. This framework was originally proposed as a way to reconcile contradictory evidence on music-language relations with respect to syntactic processing, since neuropsychology had pointed to independence and neuroimaging had pointed to overlap (Patel, 2003; 2008). This paper expands this framework, applying it to online processes of syntactic comprehension and to developmental processes involved in learning the phonemic structure of language. In both cases, resource sharing suggests that aspects of language and music which are very different in their structural organization can have deep connections in terms of cognitive processing.

The paper is organized as follows. Sections 2 and 3 provide a brief background on the neuropsychological and neuroimaging studies that gave rise to the resource sharing framework. Section 4 provides some details on the framework itself. Sections 5, which constitutes the body of the paper, discusses resource sharing with respect to musical syntax, while section 6 shows how this framework can be applied to developmental processes. Section 7 provides concluding remarks.

Throughout the paper, the focus is on comparing mechanisms involved in the comprehension of instrumental music and of ordinary spoken language. This reflects the greater amount of research on perception than on production in both domains.

Furthermore, the focus is on the Western European tonal tradition (which has received the greatest amount of theoretical and empirical research within music cognition), though the framework should apply to many other types of musical systems.

2. Neuropsychology: evidence for domain-specificity

Among the most exciting discoveries in music neuroscience have been cases of selective deficits of music cognition in normal individuals with focal brain damage (“acquired amusia”). While rare, such cases are significant because they imply that brain networks can be specialized for musical functions, without overlap with networks involved in language, environmental sound perception, etc. (Peretz & Coltheart, 2003). For example, Peretz and colleagues (Peretz, 1993; Peretz et al., 1994) studied a patient, G.L., who suffered an enduring loss of sensitivity to musical key structure after suffering bilateral temporal lobe damage due to strokes (those unfamiliar with the concept of a musical key may consult section 5.1 below). G.L.’s primary auditory cortex was spared, but there was damage to rostral superior temporal gyri, which encompasses several auditory association areas (cf. Tramo et al., 1990).

G.L. could discriminate changes between single pitches and was sensitive to differences in melodic contour in short melodies. However, he showed an absence of sensitivity to musical key. For example, G.L. was given a probe-tone task in which a few notes (which established a musical key) were followed by a target tone. The task was to rate the how well the target fit with the preceding context. Normal controls showed the standard effect, whereby tones from the key were rated higher than out-of-

key tones. G.L., in contrast, showed no such effect, and tended to base his judgments on the pitch distance between the penultimate and final tone. He also failed to show an advantage for melodies that did vs. did not adhere to a key (i.e., tonal vs. atonal melodies) in short term memory tasks, in contrast to controls. Additional experiments showed that his problems could not be accounted for by a general auditory memory deficit. Importantly for the current purposes, G.L. scored in the normal range on standardized aphasia tests, showing that he had “amusia without aphasia.”

While G.L.’s case concerns musical syntax (cf. section 5.1 below), other cases from neuropsychology demonstrate music-language dissociations at the phonological level. For example, individuals with “pure word deafness” (typically associated with bilateral lesions to the posterior superior temporal lobe) can no longer understand spoken material but retain sensitivity to other sounds, including music (Poehpel, 2001). Since these patients can speak and can understand and/or produce language in other modalities (e.g., writing), this disorder appears to be a selective problem with auditory phonemic encoding.

3. Neuroimaging: challenges to domain-specificity

If we only had evidence from neuropsychology, it would be tempting to conclude that music and language are largely independent cognitive functions (Marin & Perry, 1999). Neuroimaging has challenged this view, however, by revealing significant overlap in certain aspects of musical and linguistic processing in normal individuals. For example, the processing of one aspect of musical grammar (the harmonic structure or chords and

keys) appears to involve brain operations also involved in linguistic syntactic processing. This was first demonstrated in an ERP study which embedded an out-of-key chord in the middle of a tonal chord sequence (Patel et al., 1998). This out-of-key chord elicited a P600, an ERP component known to be associated with linguistic syntactic processing (see Gouvea et al., 2009, for a recent discussion of the linguistic processes underlying the P600). In fact, the P600 to out-of-key chords was statistically indistinguishable from the P600 generated by linguistic syntactic incongruities in sentences heard by the same participants. On this basis, it was suggested that the P600 reflected processes of structural integration shared by language and music.

Subsequent neuroimaging work provided further evidence of overlap between music and language. Using MEG, Maess et al. (2001) found that a different ERP component elicited by out-of-key chords, an early right anterior negativity (or ERAN), was generated in Broca's area and its right hemisphere homolog. Later research using fMRI and PET converged with electrophysiological measures in suggesting overlap in a number of brain regions involved in processing linguistic and musical structure, including (but not limited to) Broca's area (Koelsch et al., 2002; Tillmann et al., 2003; Schmithorst, 2005; Brown et al., 2006).

Of course, it is possible to raise a number of questions about these findings. For example, the musical P600 was elicited in a study where participants were explicitly asked to attend to (and make decisions about) the harmonic structure of a sequence. In other studies using out-of-key chords in which the participants' attention is not drawn to harmonic structure via task instructions, a P600 is not always elicited by out-of-key chords. Instead, one often observes a different, negative-going ERP component (the

N500), which is distinct from ERPs components elicited by language processing (e.g., Koelsch, 2005). In terms of the ERAN, while this component does appear to have a generator in left Broca's area, it also appears to have a stronger generator in the right hemisphere Broca's homolog, suggesting an opposite hemispheric bias for grammatical processing in music and language (cf. Tillmann et al., 2006). Finally, with regard to overlapping brain regions for music and language observed with fMRI and PET, it is possible that functionally separate networks are interdigitated within a common cortical region, and cannot be resolved given the spatial resolution of current neuroimaging techniques (Peretz & Zatorre, 2005).

4. The resource-sharing framework

The preceding sections show that neuropsychology and neuroimaging yield rather different perspectives on music-language relations in the brain, the former favoring domain-specificity and the latter challenging this view. While there are certainly open questions about existing neuroimaging findings (as outlined above), these data at least raise the possibility that there are aspects of music and language (e.g., syntax) that exhibit domain-specificity *and* neural overlap. If this is the case, then one needs a conceptual framework for reconciling this contradiction and for guiding future work. Resource sharing is one such a framework, and is based on two basic principles:

1. Language and music involve domain-specific representations.

For example, knowledge of words and their syntactic properties involves a set of representations which are distinct from the representations of chords and their harmonic relations. The resource sharing framework posits that domain-specific representations are stored in long-term memory in distinct associative networks (cf. Tillmann et al., 2000). Hence these representations are susceptible to selective damage, leading to dissociations between linguistic and musical cognition (e.g., the case of G.L.).

2. When similar cognitive operations are conducted on domain-specific representations, the brain shares neural resources between the two domains.

Principle 2 posits that representational specificity (as posited by principle 1) should be conceptually distinguished from processing specificity. That is, there are circumstances where similar cognitive operations are conducted on domain-specific knowledge, as part of building coherent percepts. In this case, principle 2 suggests that these operations share neural resources. Hence it is possible to observe similar brain signatures for certain aspects of linguistic and musical processing.

The remainder of this paper will focus on evidence for resource sharing in online syntactic comprehension (section 5), and in learning linguistic and musical sound systems (section 6). Before proceeding, it is worth commenting on the motivation for these choices. In terms of online comprehension, there are several areas (other than syntax) where one could explore resource sharing, such as the affective appraisal of auditory signals (Juslin & Laukka, 2003; Thompson et al., 2004) and the encoding of melodic contour patterns (Patel et al., 1998b). Syntactic processing is chosen because it has been

the locus of strong claims for domain-specificity in both language and music (e.g., Fodor, 1983; Jackendoff & Lerdahl, 2006). It is also arguably unique to humans, and cross-domain research may yield insights into basic mechanisms underlying this special aspect of human cognition. Finally, linguistic and musical syntax have been well studied both theoretically and empirically, making this area attractive for research.

In terms of the second topic, a focus on the learning of sound categories is motivated by growing evidence for a link between musical ability and linguistic phonemic abilities (e.g., Slevc & Miyake, 2006; Wong et al., 2007). This finding has significant implications for educational and clinical issues, including the rehabilitation of certain language disorders (cf. Goswami, 2009). Understanding the nature of shared processing for sound category learning may thus have important practical consequences.

5. Syntax

Language is a syntactic system: it has discrete structural elements (e.g., words) which are combined in principled ways to form hierarchically organized sequences (e.g., sentences). These sequences have a rich relationship between their structure and the meaning they convey. For example, the sentence “The girl who kissed the boy opened the door” contains the word string “the boy opened the door”, but a speaker of English knows that the boy did not do the opening, because the relative pronoun “who” indicates an embedded relative clause that separates “girl” and its dependent verb, “opened” (Figure 1).

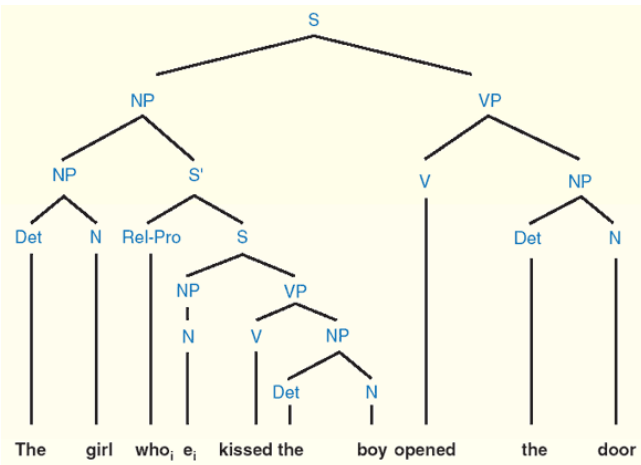


Figure 1. A sentence of English, showing its hierarchical phrase structure. Note that although the words “boy” and “opened” are immediately adjacent, they belong to different syntactic phrases. (S, sentence; NP, noun phrase, VP, verb phrase, S', sentence modifier [relative clause]; N, noun; V, verb; Det, determiner; Rel-Pro, relative pronoun.) Within the clause the relative pronoun “who” is referred to as a filler and is interpreted as the actor for the verb “kissed”. This relationship is identified by the presence of a co-indexed empty element e_i in the subject position of the relative clause. From Patel (2008).

Like language, music also has discrete structural elements (e.g. tones or chords) which are combined in principled ways into hierarchically structured sequences (Lerdahl & Jackendoff, 1983). The syntactic architecture of these sequences differs from linguistic syntax in a number of ways. For example, music does not have the same kind of dependency structure as language, where an incoming element demands a cognitive connection with a specific distant prior structural element (as in the relation between “opened” and “girl” in the above sentence). Instead, certain pitches in musical sequences are considered structurally more prominent than others, forming an “event hierarchy” that defines the structural skeleton of a piece (Bharucha, 1984), as illustrated in Figure 2.

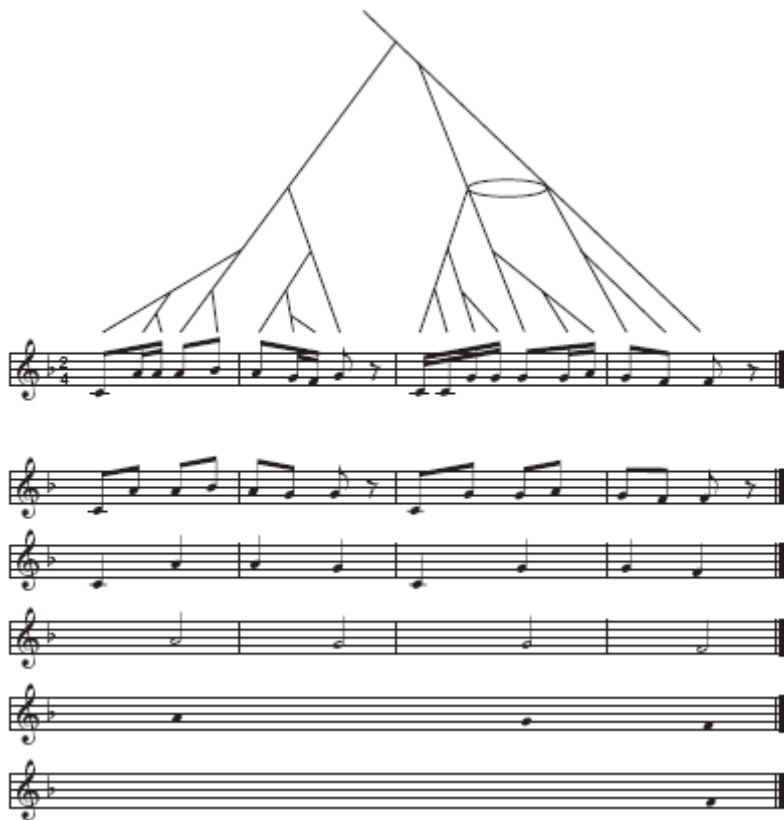


Figure 2. A time-span reduction of the first two phrases of the children's song "Hush little baby" (tree notation from Lerdahl & Jackendoff, 1983). Shorter branches terminate on less important pitches, while longer branches terminate on more important pitches. The lower staves show the dominant events at successively higher levels of tree structure. [The oval shape in the right part of the tree is meant to indicate that the short motif that projects upward to it from the left (the first four notes of the second phrase, C-C-G-G) is subordinate to both branches touched by the oval, not just the left branch (cf. Lerdahl & Jackendoff, 1983:138). This subtlety is not essential for the current discussion.] Modified from Large et al. (1995).

Before discussing evidence for resource sharing in the syntactic processing of language and music, it is worth reviewing some basic information about musical keys, since key structure is central to the experiments that follow.

5.1 Some basics of musical key

One important aspect of musical syntax concerns the notion of a musical "key". In Western tonal music, music within a given key selects a subset of 7 out of 12 available

pitches within the octave, which form a musical scale (such as the major scale). One of these pitches (the “tonic”, or “do” in the do-re-mi system of labeling pitches) acts as the most stable pitch, often serving as a point of repose. Other pitches vary in their degree of structural stability: for example the 7th tone of the scale (“ti”) is very unstable, and conveys a sense of restlessness that requires resolution. This “tonal hierarchy”, in combination with rhythmic factors, plays an important role in determining which pitches are perceived as structurally more prominent than others in musical sequences (the “event hierarchy” mentioned earlier).

An interesting feature of the tonal hierarchy is the contrast between physical and psychological distances between tones in a key. When listeners are asked to rate the perceived relatedness of tones in a key context, a subset of three tones (do, mi, and so, or c, e, and g in the key of C-major) are rated as closely related each other. The remaining scale tones are rated as less related to each other and to the tonic note, and out-of-key notes (e.g., c#) are rated as distantly related to notes within the key (in simple tonal melodies these out-of-key notes pop out as ‘sour notes’). Using multidimensional scaling, these relatedness judgments can be represented spatially, with the degree of perceived relatedness corresponding to the distance between the tones on the graph (Figure 3). A notable feature of this figure is the large distance which separates tones which are adjacent in frequency, such as c and c#. This contrast between the physical and psychological proximity of pitches is likely to be part of what animates tonal music.

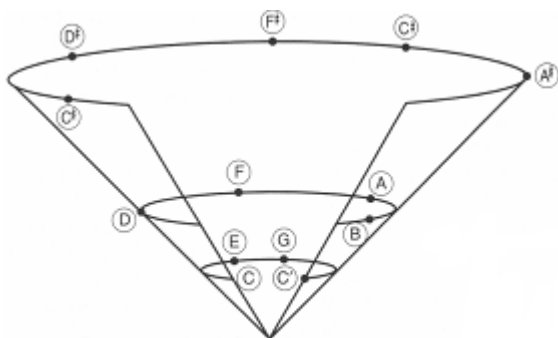


Figure 3. Geometrical representation of perceived similarity between musical pitches in a tonal context. The data are oriented toward the C major scale, where C serves as the tonic. C' is the pitch one octave above C. From Krumhansl (1979).

Tones from a musical key can be combined to form chords, which are then combined in sequences to form chord progressions. Both the “vertical” and “horizontal” organization of chords follows principles to which listeners are sensitive (Smith & Melara, 1990, Lhost & Ashley, 2006). At a higher level still, musical keys themselves have a systematic organization with respect to each other. Keys sharing more tones and chords are perceived as more closely related to each other, forming a “circle of fifths” (Figure 4) where increasing distance between two keys along the circle corresponds to a decrease in the perceived relatedness between these keys (Thompson & Cuddy, 1992).

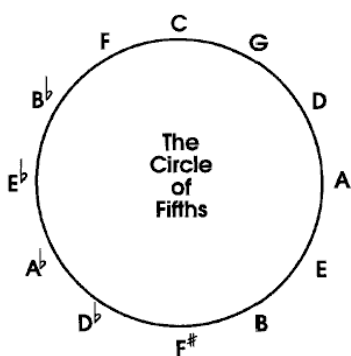


Figure 4: The circle of fifths for major keys. Each key is represented by a letter standing for its tonic.

In terms of relating structure and meaning in music, the key system has an important role to play. Composers use out-of-key notes and chords in systematic ways to create a sense of tension in musical pieces (Lerdahl & Krumhansl, 2007), and empirical studies indicate that the ebb and flow of tension is part of the emotional meaning of music (Meyer, 1956; Steinbeis et al., 2006).

5.2 Structural integration in language and music

In comprehending language and music, the structural relationship of incoming elements (such as words or chords) to preceding events must be determined in order to make sense of the sequence. On the basis of neuroimaging data and cognitive theory, Patel (2003) suggested that some aspect of this structural integration process is shared by language and music. Specifically, it was proposed that structural integration involves the rapid and selective activation of items in associative networks, and that language and music share the neural resources that provide this activation to the networks where domains-specific representation reside. This idea, termed the “shared syntactic integration resource hypothesis” (SSIRH) can be conceptually diagrammed as follows (Figure 5).

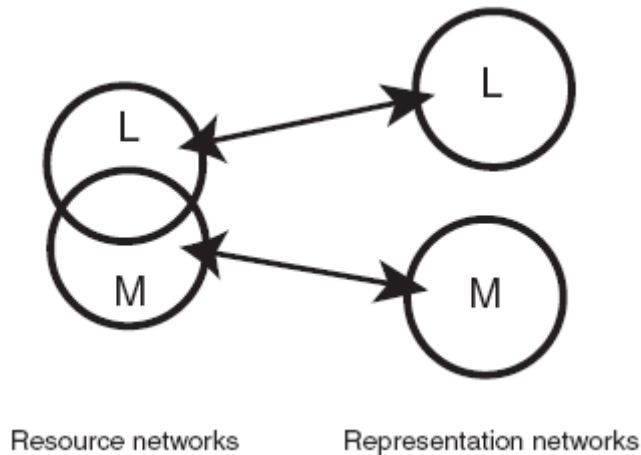


Figure 5. Schematic diagram of the functional relationship between linguistic and musical syntactic processing. L = language, M = music.

The diagram in Figure 5 represents the hypothesis that linguistic and musical syntactic representations are stored in distinct brain networks (and hence can be selectively damaged), whereas there is overlap in the networks which provide neural resources for the activation of stored syntactic representations. Arrows indicate functional connections between networks. Note that the circles do not necessarily imply highly focal brain areas. For example, linguistic and musical representation networks could extend across a number of brain regions, or exist as functionally segregated networks within the same brain regions.

How does this proposal map onto neural architecture? At the moment the answer to this question is not known. In its original formulation, the SSIRH combined the functional proposal outlined above with a rough localizationist proposal, namely that that neural resources reside in frontal brain regions, while syntactic representations reside in posterior regions. This was inspired in part by a view of the inferior frontal cortex as providing activation to posterior regions for the purpose of selecting between competing

representations in those regions, both in language and in other domains (Thompson-Schill, 2006, cf. Hagoort, 2006). Patel (2003) noted that testing this proposal required localization techniques such as fMRI, applied to within-subjects comparisons of syntactic processing in language and music. Such work remains to be done. For the current purposes, the salient point is the functional and localization aspects of the SSIRH can be conceptually decoupled.

What sorts of predictions does this hypothesis make? One prediction, based on the idea of shared, limited resources for activation, is that simultaneous resource-intensive structural integrations in language and music should interfere with each other.

In language, resource-intensive structural integrations come in at least two forms. First, an incoming word can be distant from a prior word with which it shares a syntactic dependency. According to Dependency Locality Theory (Gibson, 1998; 2000), this integration is costly because it involves reactivating the representation of the prior dependent word, whose activation level has decayed since it was first encountered. This theory accounts for a number of language processing phenomena, including the difference in processing difficulty for sentences containing subject-relative vs. object-relative clauses. For example, by a number of measures, sentence a) below is easier to process than sentence b) (King & Just, 1991).

a) The cop that met the spy wrote a report about the case.

b) The cop that the spy met wrote a report about the case.

In a) the subject-relative clause “that met the spy” contains only local integrations, while in sentence b) the object relative clause “that the spy met” requires a distant integration between “met” and “that”.

Another type of resource-intensive structural integration in language involves the violation of a syntactic expectancy. There is growing evidence that the human parser continuously predicts the syntactic category of upcoming words (Gibson, 2006; Lau et al., 2006), and that violations of these predictions cause processing difficulty. For example, according to “constraint satisfaction” models, during incremental sentence processing different possible syntactic analyses of a sentence have different levels of activation, with the currently preferred analysis having the highest level of activation. An incoming word that is syntactically unexpected can force a re-ranking among the possible syntactic analyses, and this is costly because resources must be reallocated to boost the activation of a different structure (e.g. MacDonald, 1993, cf. Marslen-Wilson, 1975).

Expectancy effects can account for a range of processing phenomena in language (Levy, 2008), including the difficulty of “garden path” sentences such as:

c) The man accepted the prize was not going to him.

In c), when a person first processes “the prize”, s/he tends to interpret it as the direct object of “accepted”, and does not expect a following verb. The appearance of “was” violates this expectancy and forces a revision of the preferred syntactic structure (making “the prize” the subject of a reduced sentence complement), leading to processing difficulty (Trueswell & Kim, 1998).

Turning to music, one can conceive of resource-intensive structural integrations as those involving incoming elements which are harmonically distant from the current tonal context, where distance is measured in terms of tonal pitch space (Lerdahl, 2001, cf. Figures 3 and 4). For example, when in the key of C, a D-flat major chord is harmonically quite distant (5 steps away on the circle of fifths). Activating harmonically distant elements is costly because they have not been primed by spreading activation in associative networks that store knowledge of harmonic relations (Bhaurcha, 1984; Tillmann et al., 2000). Note that given the structural norms of tonal music, events that are harmonically distant from the current context (e.g., out-of-key notes and chords) are also unexpected (Huron, 2006). Hence manipulations of harmonic distance in music can also be thought of as manipulations of musical syntactic expectancy.

The foregoing discussion suggests that simultaneous resource-intensive structural integrations in language and music can be created by pairing resource-intensive sentences with resource-intensive musical sequences. More specifically, words that cause structural integration difficulty in sentences (due to distance or expectancy effects) can be aligned with out-of-key notes or chord in tonal music. According to the SSIRH, this should lead to interference between language and music processing (Patel, 2003). The following section describes three tests of this prediction, one using ERPs and two using behavioral methods.

5.3 Testing resource sharing with neural and behavioral methods

As background, it is worth noting that past studies combining language and music to study processing interactions have focused on the relationship between linguistic semantic processing and musical harmonic processing. For example, in an ERP study Besson et al. (1998) used opera-like vocal melodies to combine semantic incongruities in language with out-of-key notes in music. The authors found that the semantic anomalies gave rise to a N400, while the out-of-key notes gave rise to a late positive component (resembling a P600), and that a simple additive model predicted the data for combined semantic / harmonic violations quite well. Using similar stimuli in a behavioral study, Bonnel et al. (2001) had listeners either perform a single task (judge the incongruity of final word or note) or a dual-task (judge the incongruity of both), and found that the dual task did not result in a decrease in performance compared to the single-task conditions. These findings have been taken as evidence for the cognitive independence of musical and linguistic processing.¹ In contrast to these studies, the studies reported here combine a linguistic *syntactic* manipulation with a musical harmonic manipulation, which from the foregoing discussion of the SSIRH, is predicted to lead to an interaction. Each study below takes a different approach, but all three provide converging evidence which supports this prediction.

Koelsch et al. (2005) conducted an ERP study in which short sentences were presented visually in a word-by-word format simultaneously with musical chords, with one chord per word. In some sentences, the final word created a grammatical violation via a gender disagreement (the sentences were in German, in which many nouns are marked for gender). An example of a gender violation used in this study was: Er trinkt den kühlen Bier, “He drinks the_{masculine} cool_{masculine} beer_{neuter} . This final word violated a

syntactic expectancy in language. The chord sequences were designed to strongly invoke a particular key, and the final chord could either be the tonic chord of that key or an unexpected out-of-key chord from a distant key (e.g., a D-flat major chord at the end of a C-major sequence). The participants (all non-musicians) were instructed to ignore the music and simply judge if the last word of the sentence was linguistically correct.

Koelsch et al. focused on early ERP negativities elicited by syntactically incongruous words and chords. Previous research on language or music alone had shown that the linguistic syntactic incongruities were associated with a left anterior negativity (LAN), while the musical incongruities were associated with an early right anterior negativity (ERAN) (Gunter et al., 2000; Koelsch et al., 2000; Friederici, 2002). For their combined language-music stimuli, Koelsch et al. found that when sentences ended grammatically but were accompanied by an out of key-chord, a normal ERAN was produced. Similarly, when chord sequences ended normally but were accompanied by a syntactically incongruous word, a normal LAN was produced. The question of interest was how these brain responses would interact when a sequence had simultaneous syntactic incongruities in language and music. The main finding was that the brain responses were not simply additive. Instead, there was an interaction: the LAN to syntactically incongruous words was significantly *smaller* when these words were accompanied by an out-of-key chord, as if the processes underlying the LAN and ERAN were competing for similar neural resources. In a control experiment, Koelsch et al. showed that this was not due to general attentional effects because the LAN was not influenced by a simple auditory oddball paradigm involving physically deviant tones on the last word in a sentence. Thus the study supported the prediction that tasks which

combine linguistic and musical syntactic integration will show interference between the two processes.

Turning to recent behavioral research, Fedorenko et al. (2009) manipulated syntactic integration difficulty via the distance between dependent words, rather than via expectancy violations. These researchers used fully grammatical sentences of the type shown in a) and b) in section 5.2 above, reproduced here:

a) The cop that met the spy wrote a book about the case.

b) The cop that the spy met wrote a book about the case.

The sentences were sung to melodies (one note per word) that did or did not contain an out-of-key note on the last word of the relative clause, underlined above. (Recall that this word is associated with a distant structural integration in b) but not in a)). A control condition was included for an attention-getting but non-harmonically deviant musical event: a 10 dB increase in volume on the last word of the relative clause. After each sentence, participants were asked a comprehension question, and accuracy was assumed to reflect processing difficulty.

The results revealed an interaction between musical and linguistic processing: comprehension accuracy was lower for sentences with distant vs. local syntactic integrations (as expected), but crucially, this difference was larger when melodies contained an out-of-key note. The control condition (loud note) did not produce this effect: the difference between the two sentence types was of the same size as that in the

conditions that did not contain an out-of-key note. These results suggest that some aspect of structural integration in language and music relies on shared processing resources.

The final study described here, by Slevc et al. (2009), manipulated linguistic structural integration difficulty via syntactic expectancies, and also directly compared the influence of musical harmonic manipulations on linguistic syntactic vs. semantic processing (cf. Steinbeis & Koelsch, 2008). However, unlike the Koelsch et al. (2005) study, linguistic expectancy was manipulated via garden-path sentences rather than via grammatical incongruities.

In Slevc et al.'s study, participants read sentences phrase by phrase on a computer screen. They controlled the timing of phrases by pushing a button to get the next phrase. In such studies, the amount of time spent viewing a phrase is assumed to reflect the amount of processing difficulty associated with that phrase. This "self-paced reading" paradigm has been used often in psycholinguistic research. The novel aspect of Slevc et al.'s study was that each phrase was accompanied by a chord so that the entire sentence made a coherent, chorale-style chord progression.

The sentences contained either a linguistic syntactic or semantic manipulation. In the syntactic manipulation, sentences like d) included either a full or reduced sentence complement clause, achieved by including or omitting the word "that". (Note: the vertical slashes below indicate the individual phrases used in the self-paced reading experiment):

d) The scientist | wearing | thick glasses | confirmed (that) | the hypothesis | was | being | studied | in his lab.

In such sentences, the omission of “that” results in the reduced complement clause “the hypothesis was being studied in his lab”. In this case, readers tend to interpret “the hypothesis” as the direct object of “confirmed”, which causes syntactic integration difficulty when “was” is encountered, as this signals that “the hypothesis” is actually the subject of an embedded clause. In other words, the simple omission of “that” leads to a downstream localized processing difficulty (on “was”) due to violation of a syntactic expectancy.

In the semantic manipulation, sentences like e) included either a semantically consistent or anomalous word, thereby confirming or violating a semantic expectancy.

e) The boss | warned | the mailman | to watch | for angry | dogs / pigs | when | delivering | the mail.

The chord played during the critical word (underlined) was either harmonically in-key or out-of-key. (Out-of-key chords were drawn from keys 3-5 steps away on the circle of fifths from the key of the phrase. For example, a chord progression in C major might have an E-major chord as the out-of-key chord, cf. Figure 4). Since out-of-key chords are harmonically unexpected, the experiment crossed syntactic or semantic expectancy in language with harmonic expectancy in music. The dependent variable of interest was the reading time for the critical word.

The main finding was a significant three-way interaction between linguistic manipulation type (syntactic or semantic), linguistic expectancy, and musical expectancy.

That is, syntactically and semantically unexpected words were read more slowly than their expected counterparts, but a simultaneous out-of-key chord caused substantial additional slowdown for syntactically unexpected words, but not for semantically unexpected words. Thus, processing a harmonically unexpected chord interfered with the processing of syntactic, but not semantic, relations in language. (A control experiment showed that when a chord of unexpected timbre, which created an attention-getting psychoacoustic event, was aligned with the target word, it did not interfere with either type of linguistic processing.) Once again, these results support the claim that neural resources are shared between linguistic and musical structural integration.

The three studies described above provide converging evidence for interactions between musical and linguistic syntactic processing. Interestingly, there appears to be a specific interaction between musical harmonic processing and syntactic (vs. semantic) processing in language, suggesting some level of separation between the brain systems that handle linguistic syntax vs. semantics (cf. Osterhout et al., in press; Tyler & Marslen Wilson, 2008). Furthermore, the work points to a specific point of convergence between linguistic and musical syntactic processing, namely, a sharing of neural resources for activating items in associative networks as part of a process of structural integration. This shows how linguistic and musical syntax can have an important point of contact despite many differences in the formal organization of syntactic structures in the two domains.

In terms of connections to broader issues, studies combining language and music could help refine our understanding of the brain mechanisms involved in generating syntactic expectations. It seems that in both language and music the brain continuously

predicts the structural categories of upcoming events, perhaps because this leads to more efficient processing of predicted events (which constitute the majority of events).

Expectation-based syntactic processing is an emerging research area in psycholinguistics (Levy, 2008), and can perhaps benefit from a conceptual synthesis with the study of expectation in music processing, a topic of active research with a long intellectual history (Meyer, 1956; Huron, 2006). Using music and language together could help identify the more abstract cognitive operations that are being computed by brain circuits involved in generating syntactic expectancies.

6. Phonology

The remainder of this paper briefly discusses evidence for shared resources involved in learning the sound systems of language and music. The focus is on the relationship between musical abilities and phonemic abilities in language. There is growing evidence that either pitch-related or rhythm-related musical skills are related to phonemic abilities in language, such as the segmentation, categorization, or discrimination of phonemes. However, as noted earlier in this paper, there is also evidence from neuropsychology that auditory phonemic encoding of sounds can be selectively disrupted by brain damage, leaving the perception of musical sounds intact (i.e., in pure word deafness). Hence the relationship between musical and phonemic skills fits nicely into a resource-sharing framework, under the assumption that the end *products* of phonemic development are unique, but that some of the *processes* that give rise to these representations are shared by music and language (cf. McMullen & Saffran, 2004). If this is the case, the key issue is to

specify the nature of the shared neural resources linking music and language. This issue is taken up in section 6.2 below, after describing evidence for a relationship between musical and phonemic abilities.

6.1 Evidence for a link between musical and phonemic skills

This section describes two studies which support a link between pitch-related musical abilities and linguistic phonemic abilities. For studies examining a link between musical rhythmic abilities and linguistic abilities, see Overy (2003) and Corriveau & Goswami (2009).

Anvari et al. (2002) studied the relation between early reading skills and musical development in a large sample of English-speaking 4 and 5 year-olds. Learning to read English requires mapping visual symbols onto phonemic contrasts, and thus taps into linguistic sound categorization skills. The children were given an extensive battery of tasks, which included tests of reading, phonemic awareness, vocabulary, auditory memory, and mathematics. (Phonemic awareness refers to the ability to identify the sound components of a word, and a large body of research indicates that children with greater phonemic awareness have advantages in learning to read.) On the musical side, both musical pitch and rhythm discrimination were tested. The pitch tasks included same/different discrimination of short melodies and chords, and the rhythm tasks involved same/different discrimination of short rhythmic patterns and reproduction of rhythms by singing. The most interesting findings concerned the 5 year olds. For this group, performance on musical pitch (but not rhythm) tasks predicted unique variance in

reading abilities, even when phonemic awareness was controlled for. Furthermore, statistical analysis showed that this relation could not be accounted for via the indirect influence of other variables such as auditory memory.

Turning to research on adults, Slevc & Miyake (2006) examined the relationship between proficiency in a second language and musical ability. Several prior studies on this topic had not found any significant relationship between these variables. Crucially, however, these studies had always relied on participant's self reports of their own musical ability. Slevc and Miyake went beyond self-report and tested both linguistic and musical skills in a quantitative fashion. Working with a group of 50 Japanese adult learners of English living in the USA, they administered language tests which examined receptive and productive phonology, syntax, and lexical knowledge. (The tests of receptive phonology included identifying words which differed by a single phoneme, e.g. "clown" vs. "crown"). The musical tests examined pitch pattern perception, e.g., via the detection of an altered note in a chord or in a short melody, as well as accuracy in singing back short melodies. In addition to these tests, the researchers also measured a number of variables known to be associated with second language proficiency, such as age of arrival in the foreign country, number of years spent living there, amount of time spent speaking the second language (L2), and phonological short term memory in the native language. The question of interest was whether musical ability could account for variance in L2 ability beyond that accounted for by these other variables. As in Anvari et al. (2002), the authors used hierarchical regression to tease apart the influence of different variables. The results were clear: musical ability did in fact predict unique variance in L2 skills.

Most relevant for the current discussion, this predictive relationship was confined to L2 receptive and productive phonology, i.e., to phonemic skills.

6.2 What is the nature of the shared neural resources?

What neural mechanisms could mediate associations between musical abilities and phonemic skills? Overy (2003) has suggested that musical training improves temporal processing abilities, which are relevant to phonological segmentation skills (cf. Goswami, 2009). In an update of this idea, Tallal and Gaab (2006) suggested that musical training improves auditory rapid spectrotemporal processing, which is used in the processing of linguistic phonemic components. Both hypotheses suggest that musical experience improves sensory encoding of dynamically changing sounds, and that this improved sensory processing benefits the perception of speech. However, they do not specify at what level of the nervous system this “sensory tuning” acts.

A recent study by Wong et al. (2007) provides intriguing evidence that such tuning occurs at a very early stage of brain processing. The study used scalp-recorded EEG to examine an oscillatory brainstem neural response to sound known as the frequency-following response (FFR), which is thought to be generated in the inferior colliculus. The FFR has been shown to have an interesting relationship to voice pitch, in that its oscillation contains considerable energy at the fundamental frequency (F0) of the voice, and can dynamically track linguistically-relevant F0 changes over short time scales (such as a single syllable, approximately 250 ms in duration). Previous research had examined the FFR during perception of syllables of Mandarin Chinese. (Mandarin uses

pitch to distinguish between words, so that a single syllable can have different lexical meanings depending on its pitch pattern.) Specifically, Krishnan et al. (2005) examined the FFR during perception of Mandarin monosyllables and found that the quality of F0 tracking was superior in native speakers of Mandarin than in native speakers of English. This suggested that auditory experience could tune subcortical sound processing mechanisms.

Wong et al. (2007) extended this work by examining F0 tracking of Mandarin monosyllables in musicians vs. nonmusicians, neither of whom had prior familiarity with Mandarin. Participants heard the syllable “mi”, spoken with different lexical tones, in the background as they watched a movie. Wong et al. found that that the quality of F0 tracking was superior in musicians, and found positive correlations between F0 tracking quality and amount of musical training, and between F0 tracking quality and performance on identification and discrimination tasks using Mandarin syllables. While prior evidence had indicated that musicians are more sensitive to subtle pitch variations in speech than non-musicians (as reflected in behavioral measures and cortical evoked potentials, e.g., Schön et al., 2004; Magne et al., 2006) Wong et al.’s findings were surprising because they suggest that musical experience influences speech processing at a very basic neural level. However, the correlational nature of the findings does raise questions about causality, as discussed in the next section (cf. Patel & Iversen, 2007).

6.3 Avenues for future research

The hypotheses of Overy (2003) and Tallal and Gaab (2006) suggest that musical training actively shapes brain mechanisms that impact linguistic ability. However, most of the behavioral and neural data supporting an association between musical skills and phonemic skills (including the study of Wong et al., 2007) are correlational in nature. Hence it is not clear if the neural resources for musical and phonemic processing are genuinely intertwined in the brain, or inborn differences in brain anatomy or physiology influence distinct neural resources involved in musical vs. phonemic abilities. The only way to address this issue is via experimental studies where groups of individuals are matched at the outset in terms of neural and behavioral measures of musical ability, and then exposed to different amounts of musical training (cf. Norton et al., 2005). If musical training improves phonemic abilities, this would be strong evidence for shared neural resources, and a variety of neural techniques could then be brought to bear to study the nature of these resources (e.g., ERPs, fMRI, MEG, TMS).

Another important issue for future research concerns the acoustic features used in musical training vs. phonemic testing. The study of Wong et al. (2007) focused on linguistic pitch processing, which was sensible since pitch is a highly structured aspect of music. However, it is of considerable interest to know whether musical training that focuses on one acoustic dimension (e.g., pitch) can benefit phonemic skills that rely on a different acoustic dimension (e.g., timbre). If this is the case, it would suggest that language and music share more abstract cognitive processes involved in sound categorization, possibly involving common cortical mechanisms.

7. Conclusion

This paper has presented a framework for music-language studies which draws a basic distinction between domain-specific representations and shared neural resources which act upon these representations, either in online processing or as part of cognitive development. Of course, language and music are unlikely to be unique in terms of resource sharing. When the two domains draw on common brain resources, they do so because these resources provide a particular processing function needed in both domains. If other mental faculties also require these resources, they will likely share them with both language and music.

While the resource sharing framework is rooted in cognitive neuroscience, testing this framework requires an integrated approach involving cognitive theory and behavioral methods. The goal of such work is to generate hypothesis-driven research which will illuminate the shared cognitive foundations of linguistic and musical abilities.

Acknowledgements

I thank John Iversen and Bob Slevc for helpful comments. Supported by Neurosciences Research Foundation as part of its program on music and the brain at The Neurosciences Institute, where ADP is the Esther J. Burnham Senior Fellow.

Notes

1. A study by Poulin-Charronat et al. (2005) combined semantic anomalies in language with a harmonic manipulation of chord sequence structure in music, and found an interaction between semantic and musical

processing. However, subsequent research by Escoffier & Tillmann (2006) suggests that this is due to a non-specific effect of the musical manipulation on attention, rather than representing a specific cognitive link between semantic and harmonic processing.

References

Anvari, S., 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.

Besson, M., Faita, F., Peretz I., Bonnel, A.J.R. (1998). Singing in the brain: independence of lyrics and tunes. *Psychological Science*, 9, 494–498.

Bharucha, J.J. (1984). Event Hierarchies, Tonal Hierarchies and Assimilation: A Reply to Deutsch and Dowling. *Journal of Experimental Psychology, General*, 113, 421-5.

Bonnel, A.-M., Faita, F., Peretz, I., & Besson, M. (2001). Divided attention between lyrics and tunes of operatic songs : Evidence for independent processing. *Perception and Psychophysics*, 63, 1201-1213.

Brown, S., Martinez, M.J. & Parsons, L.M. (2006). Music and language side by side in the brain: a PET study of the generation of melodies and sentences. *European J. Neuroscience*, 23, 2791-2803.

Corriveau, K.H. & Goswami, U. (2009). Rhythmic motor entrainment in children with speech and language impairments: Tapping to the beat. *Cortex*, 45, 119-130.

Escoffier, N. & Tillmann, B. (2006). Tonal function modulates speed of visual processing. In: M. Baroni et al. (Eds.) *Proceedings of the 9th International Conference on Music Perception & Cognition (ICMPC9), Bologna/Italy* (pp. 1878).

Fedorenko, E., Patel, A.D., Casasanto, D., Winawer, J. & Gibson, E. (2009). Structural integration in language and music: Evidence for a shared system. *Memory & Cognition*, 37, 1-9.

Fodor, J. A. (1983). *Modularity of mind*. Cambridge, MA: MIT Press.

Friederici, A. D. (2002). Towards a neural basis of auditory sentence processing. *Trends in Cognitive Science*, 6, 78-84.

Gibson, E. (1998). Linguistic complexity: Locality of syntactic dependencies. *Cognition*, 68, 1-76.

Gibson, E. (2000). The dependency locality theory: A distance-based theory of linguistic complexity. In: Y. Miyashita, A. Marantz, & W. O'Neil (Eds.), *Image, Language, Brain* (pp. 95-126). Cambridge, MA: MIT Press.

Gibson, E. (2006). The interaction of top-down and bottom-up statistics in the resolution of syntactic category ambiguity. *Journal of Memory and Language*, 54, 363-388.

Goswami, U. (2009). Mind, brain, and literacy. Biomarkers as usable knowledge for education. *Mind, Brain, and Education*, 3, 176-184.

Gouvea, A., Phillips, C., Kazanina, N. & Poeppel, D. (in press). The linguistic processes underlying the P600. *Language and Cognitive Processes*. DOI: 10.1080/01690960902965951

Gunter, T. C., Friederici, A. D., & Schriefers, H. (2000). Syntactic gender and semantic expectancy: ERPs reveal early autonomy and late interaction. *Journal of Cognitive Neuroscience*, 12, 556-568.

Hagoort, P. (2006). On Broca, brain and binding. In: Y. Grodzinsky & A. Amunts. *Broca's Region* (pp. 242-253). New York: Oxford Univ. Press.

Huron, D. (2006) *Sweet Anticipation: Music and the Psychology of Expectation*. Cambridge, MA: MIT Press.

Jackendoff, R. & Lerdahl, F. (2006). The capacity for music: What is it, and what's special about it? *Cognition*, 100, 33-72.

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

King & Just (1991). Individual differences in syntactic processing: The role of working memory. *Journal of Memory and Language*, *30*, 580-602.

Koelsch, S. (2005). Neural substrates of processing syntax and semantics in music. *Curr. Opin. Neurobiol.*, *15*, 207-212.

Koelsch, S., Gunter, T. C., Friederici, A. D., & Schröger, E. (2000). Brain indices of music processing: 'non-musicians' are musical. *Journal of Cognitive Neuroscience*, *12*, 520-541.

Koelsch S., Gunter T.C., von Cramon D.Y., Zysset S., Lohmann G., & Friederici A.D. (2002). Bach speaks: A cortical 'language-network' serves the processing of music. *Neuroimage*, *17*, 956-966.

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

Krishnan, A., Xu, Y., Gandour, J. & Cariani, P. (2005). Encoding of pitch in the human brainstem is sensitive to language experience. *Cogn Brain Res.*, *25*, 161-168.

- Krumhansl, C.L. (1979). The psychological representation of musical pitch in a tonal context. *Cognitive Psychology*, *11*, 346-374.
- Large, E. W., Palmer, C., & Pollack, J. B. (1995). Reduced memory representations for music. *Cognitive Science*, *19*, 53-96.
- Lau, E., Stroud, C., Plesch, S., & Phillips, C. (2006). The role of structural prediction in rapid syntactic analysis. *Brain & Language*, *98*, 74-88.
- Lerdahl, F. (2001). *Tonal Pitch Space*. Oxford, Oxford Univ. Press.
- Lerdahl, F., & Jackendoff, R. (1983). *A Generative Theory of Tonal Music*. Cambridge, MA: MIT Press.
- Lerdahl, F. & Krumhansl, C. (2007). Modeling tonal tension. *Music Perception*, *24*, 329-366.
- Levy, R. (2008). Expectation-based syntactic comprehension. *Cognition*, *106*, 1126-1177.
- Lhost, E. & Ashley, R. (2006). Jazz, blues and the language of harmony: Flexibility in online harmonic processing. In: M. Baroni et al. (Eds.) *Proceedings of the 9th*

International Conference on Music Perception & Cognition (ICMPC9), Bologna/Italy
(pp. 1282-1288).

MacDonald, M. C. (1993). The interaction of lexical and syntactic ambiguity. *Journal of Memory and Language*, 32, 692-715.

Maess, B., Koelsch, S., Gunter, T., & Friederici, A.D. (2001). Musical syntax is processed in Broca's area: an MEG study. *Nature Neuroscience*, 4, 540-545.

Magne, C., Schön, M., & Besson, M. (2006). Musician children detect pitch violations in both music and language better than nonmusician children: Behavioral and electrophysiological approaches. *J. Cognitive Neuroscience*, 18, 199–211.

Marin, O.S.M. & Perry, D.W. (1999). Neurological aspects of music perception and performance. In: D. Deutsch (Ed.), *The Psychology of Music, 2nd Edition* (pp. 653-724). San Diego, CA: Academic Press.

Marslen-Wilson, W. D. (1975). Sentence perception as an interactive parallel process. *Science*, 189, 226-228.

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

Meyer, L.B. (1956). *Emotion and Meaning in Music*. Chicago: Univ. of Chicago Press.

Norton, A., Winner, E., Cronin, K., Overy, K., Lee, D.J. & Schlaug, G. (2005) Are there neural, cognitive, or motoric markers for musical ability prior to instrumental training? *Brain and Cognition*, 59, 124-34.

Osterhout, L., Kim, A., & Kuperberg, G. (in press). The neurobiology of sentence comprehension. In: M. Spivey, M. Joannisse, & K. McRae (Eds), *The Cambridge Handbook of Psycholinguistics*. Cambridge: Cambridge University Press

Overy K. (2003). Dyslexia and Music: From Timing Deficits to Musical Intervention. *Annals of the New York Academy of Sciences*, 999, 497-505.

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

Patel, A.D. (2008). *Music, Language, and the Brain*. New York: Oxford Univ. Press.

Patel, A.D., Gibson, E., Ratner, J., Besson, M. & Holcomb, P. (1998). Processing syntactic relations in language and music: An event-related potential study. *Journal of Cognitive Neuroscience*, 10, 717-733.

Patel, A.D., & Iversen, J.R. (2007). The linguistic benefits of musical abilities. *Trends in Cognitive Sciences*, 11, 369-372.

Patel, A.D., Peretz, I., Tramo, M. & Labrecque, R. (1998b). Processing prosodic and musical patterns: a neuropsychological investigation. *Brain and Language*, 61, 123-144.

Peretz, I. (1993). Auditory atonalia for melodies. *Cognitive Neuropsychology*, 10, 21-56.

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

Peretz, I., Kolinsky, R., Tramo, M., Labrecque, R., Hublet, C., Demeurisse, G., & Belleville, S. (1994). Functional dissociations following bilateral lesions of auditory cortex. *Brain*, 117, 1283-1302 .

Peretz I, & Zatorre, R.J. (2005). Brain organization for music processing. *Annual Review of Psychology*, 56, 89–114.

Poeppel, D. (2001). Pure word deafness and the bilateral processing of the speech code. *Cognitive Science*, 21, 679-693.

Poulin-Charronnat, B., Bigand, E., Madurell, F., & Peereman, R. (2005). Musical structure modulates semantic priming in vocal music. *Cognition*, 94, B67–B78.

- Schmithorst, V.J. (2005). Separate cortical networks involved in music perception: Preliminary functional MRI evidence for modularity of music processing. *Neuroimage*, 25, 444-451.
- Schön, D., Magne, C., & Besson, M. (2004). The music of speech: Electrophysiological study of pitch perception in language and music. *Psychophysiology*, 41, 341–349.
- Slevc, L.R. & Miyake, A. (2006). Individual differences in second language proficiency: Does musical ability matter? *Psychological Science*, 17, 675-681.
- Slevc, L.R., Rosenberg, J.C., & Patel, A.D. (2009). Making psycholinguistics musical: Self-paced reading time evidence for shared processing of linguistic and musical syntax. *Psychonomic Bulletin and Review*, 16, 374-381.
- Smith, J.D., & Melara, R.J. (1990). Aesthetic preference and syntactic prototypicality in music: ‘Tis the gift to be simple.’ *Cognition*, 34, 279-298.
- Steinbeis, N., Koelsch, S. & Sloboda, J.A. (2006). The role of harmonic expectancy violations in musical emotions: Evidence from subjective, physiological, and neural responses. *Journal of Cognitive Neuroscience*, 18, 1380–1393.

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.

Tallal, P. & Gaab, N. (2006). Dynamic auditory processing, musical experience and language development. *Trends in Neurosciences*, *29*, 382-370.

Thompson, W.F., & Cuddy, L.L. (1992). Perceived key movement in four-voice harmony and single voices. *Music Perception*, *9*, 427-438.

Thompson, W.F., Schellenberg, E.G., & Husain, G. (2004). Decoding speech prosody: Do music lessons help? *Emotion*, *4*, 46-64.

Thompson-Schill, S. L. (2006). Dissecting the language organ: A new look at the role of Broca' area in language processing. In: A. Cutler, (Ed.) *Twenty-First Century Psycholinguistics: Four Cornerstones* (pp. 173-189). Mahwah, NJ: Lawrence Erlbaum.

Tillmann, B., Bharucha, J.J. & Bigand, E. (2000). Implicit learning of tonality: a self-organizing approach. *Psychological Review*, *107*, 885-913.

Tillmann, B., Janata, P. & Bharucha, J.J. (2003). Activation of the inferior frontal cortex in musical priming. *Cognitive Brain Research*, *16*, 145-161.

Tillmann, B., Koelsch, S., Escoffier, N., Bigand, E., Lalitte, P., Friederici, A.D. & vonCramon, D.Y. (2006). Cognitive priming in sung and instrumental music: Activation of inferior frontal cortex. *NeuroImage*, 31, 1771-1782.

Tramo, M.J., Bharucha, J.J. & Musiek, F.E. (1990) Music perception and cognition following bilateral lesions of auditory cortex. *Journal of Cognitive Neuroscience*, 2, 195-212.

Trueswell, J. C., & Kim, A. E. (1998). How to prune a garden path by nipping it in the bud: fast priming of verb argument structure. *Journal of Memory & Language*, 39, 102–123.

Tyler, L. K. & Marslen-Wilson, W. (2008). Fronto-temporal brain systems supporting spoken language comprehension. *Philosophical Transactions of the Royal Society*, 363, 1037-1054.

Wong, P.C.M., Skoe, E., Russo, N.M., Dees, T., & Kraus, N. (2007). Musical experience shapes human brainstem encoding of linguistic pitch patterns. *Nature Neuroscience*, 10, 420-422.