Music and Language: Exploring Evidence of Shared Neural Processing

A Literature Review and Original Experiments

David Cremins

Written in partial fulfillment of a B.A. in Cognitive Science

Advisor: Professor Deborah Burke

Pomona College, April 2018

Table of Contents

Acknowledgements

I am deeply indebted to my advisor, reader, professor, mentor, and friend Debby Burke for all her help, not just with this thesis but throughout my years at Pomona in and beyond the Cog Sci department. Invaluable help, ranging from detailed feedback on drafts to eye-opening conversations, came from Professors Robin Melnick, Michael Spezio, Mike Diercks, Tim Justus, Shlomi Sher, and Alfred Cramer. For assistance with the stimuli and general musical inspiration I credit Kieran Cremins and Sophia Sun. Thank you to everyone who participated in these experiments (especially Spencer Louie for help with Excel!). I also thank the researchers Mauricio Martins and Robert Slevc for their engaging communications with a stranger interested in their work. My interest in this work was first inspired by Micah Johnson when I worked for him in Martin Monti's lab at UCLA, and he has been a dependable collaborator and supporter throughout this process. Finally, I am forever grateful to my family and friends (<3 Sontag 160) for their support throughout the years and for help testing out different experimental iterations.

Abstract

The past few decades have seen a dramatic increase in the amount of research investigating the neural foundations of natural language and music processing. Researchers across disciplines and the world seek to determine to what degree there are shared neural and/or cognitive components between these two domains. This thesis first reviews the existing literature, focusing on theoretical perspectives and experimental results from the cognitive sciences relating to the processing of music and language, especially in Broca's area. Then, two original studies are presented, which aimed at further investigating the relationship between these expressive systems so central to the human experience. The main finding, that musicians and non-musicians have different interpretations of tonal music harmonic violations, forms the basis of some recommendations for further work in this field.

Literature Review

"I suspect music is [merely] auditory cheesecake, an exquisite confection crafted to tickle the sensitive spots of...our mental faculties"– Steven Pinker (*How the Mind Works,* 1997)

A perennial question in the study of the human brain and behavior goes something like this: how functionally specialized are the neural regions which enable cognitive capacities? Put more simply, how modular is the mind? The answers to these and related questions will shed light on the evolution of our species, provide insight into how to treat neurological diseases and deficits, and perhaps even inform the design of future intelligent systems. Yet, even when it comes to one of the most studied human cognitive systems – language – the debate rages on without broad consensus on some very fundamental issues: how dependent on domain-general cognitive abilities is our aptitude for parsing and producing language (Pinker & Jackendoff, 2005)? To what degree is our ability to think dependent on our ability to verbalize thought (Fedorenko & Varley, 2016)? If there are language-specific neural systems or computations, what are they (Besson & Schön, 2001)?

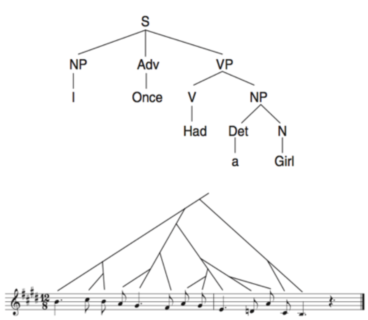
Increasingly over the past several decades, scholars have attempted to clarify these questions by comparing linguistic sub-systems with those of other cognitive domains, especially music. Comparing music and language is the topic of interest for myriad books and scientific articles. After all, these two acoustic communication systems appear to be uniquely human in the degree to which they allow for creative expressive. Research questions in this line of study are, most notably: How interconnected are the cognitive capacities for music and language? Are the same neural resources and computations recruited when we hear a rousing speech as when we hear a moving melody? What might these computations be? Did one domain inherit or descend from the other? While there are clear, cross-cultural constraints on the understanding of and production of both music and language, the answers to these questions relate to those more general questions posed in the preceding paragraph and may allow us to understand the true architecture of our neural system and contest the primacy of language in the human mind (Justus & Hustler, 2005).

Correlating Two Domains of Cognition

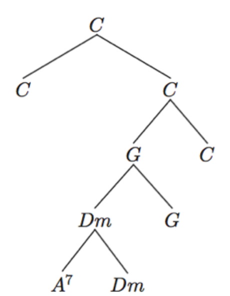
Increased interest in the connections between music and language as cognitive systems has inspired a wealth of research. Much of this work has sought to establish links between abilities in each domain. Although this largely behavioral, correlational evidence is not the focus of this review, it is helpful to begin by providing some examples of this type of work as a backdrop. Positive correlations have repeatedly been found between musical and reading abilities (Anvari, 2001) as well as musical and verbal abilities (Forgeard, 2008) in children, including in second-language verbal abilities (Slevc & Miyake, 2006). Additionally, highly musically trained individuals are shown to have increased grey matter density in regions sustaining higher cognitive abilities (James, 2014), improved performance in standard batteries of executive functioning (Zuk et al., 2014), and even improvements in speech processing (Patel, 2014). These results have sparked a debate as to how much transfer there is between musical entrainment and aptitude in other skills, with some researchers claiming music leads to improvement in a wide variety of capacities (Miendlarzewka & Trost, 2014). On the other hand, others are still skeptical that a consistent trend has been established, seeing that other factors such as general intelligence, memory capacity, and socioeconomic status are often not controlled for in the aforementioned studies and may entirely explain the observed variance (Sala & Gobet, 2017).

Hierarchical Cognition

This literature review will discuss much of the prominent experimental evidence and influential theories about the neural and cognitive overlap, or lack thereof, between music and language, largely focusing on the comprehension of each[1]. Before delving into this divisive debate, though, it is worthwhile to understand a key term which is used to link the two. Music and language, it is often claimed, are both examples of cognition which deals with hierarchical structures[2]. Music in particular has been described as hierarchical since at least the mid-1950s, and experimental evidence has long shown that highly-hierarchical musical sequences are processed more efficiently (Deutsch, 1980). Hierarchy here refers to a series of discrete elements (e.g. words in a sentence or notes in a musical phrase) structured in super-ordinate/sub-ordinate relations to one another such that an element from one part of a phrase can inform and be informed by subsequent or preceding elements (Fitch & Martins, 2014). Examples of these types of hierarchical structures can be seen in Figure 1 – though it is crucial to remember that structural representations do not necessarily imply anything about how information in either domain is processed. Care must be taken, additionally, in using the word 'hierarchy' with respect to cognitive processes because the term is used in a wide variety of ways, not least of which is in reference to the cortical organization of the cerebrum. For example, the frontal lobes are hierarchically organized when it comes to cognitive control operations in the sense that



Potential parsings of a phrase, both melodically and linguistically (from Figure 1, Jackendoff, 2009).



Tree-based representation of the dependency relationships in a common chord pattern (from Figure 2, Rohrmeier, 2011).

Figure 1. Examples of parsing musical and linguistic phrases from previous literature.

---

[1] This is not to suggest, of course, that there is not interesting work to be done into the generation of music as compared that of speech. Although relatively less examined, there have been investigations into the neural correlates of musical production, especially jazz improvisation (Limb & Braun, 2008).

[2] These are, of course, not the only such examples of hierarchically structured stimuli the brain has to deal with. Other examples include mathematics, logical inference, and, especially, the execution of complex action sequences, though a full exploration of these domains is beyond the scope of this review (Clerget et al., 2011).

rostral regions modulate the function of more caudal ones (Badre & Frank, 2011; Fuster, 2001).

Language is a highly-developed system in which the understanding and construction of long-distance and local dependency relationships (between words, clauses, and even sentences) are crucial. However, to what extent this is true in music is a crucial, unresolved research question. In addition to structural parallels, is there “also a substantial functional and neurophysiological overlap between aspects of language and non-linguistic sequential cognition”, such as musical cognition (Dominey et al., 2003)? Before directly addressing more nuanced questions of neural processing, though, the next two sections will review key literature concerning the study of syntax in language and music.

Syntax and Language

Syntax in language has been understood to be separable from other linguistic sub-systems such as semantics and phonology for decades (Ferreira & Clifton, 1986), a conclusion which has informed research in linguistics and the cognitive sciences alike. Research into the origins and functions of syntax has been largely driven by Noam Chomsky’s theory of universal grammar, i.e. the neuronal embeddedness of an innate, generative capability to learn and apply syntactic rules (Chomsky, 1965)[3]. But, exactly how and where linguistic syntax is enabled in the brain remains open for investigation. Clarifying these points will help our understanding of the neural foundation of human language, as well as potentially have implications for other cognitive domains such as music. As will be discussed later, there is increased interest in whether syntax should be considered a domain-specific capacity.

[3] While hugely influential, there are also a wealth of non-nativist approaches to understanding language structure and acquisition, so-called “constructionist grammars” (Cameron-Faulkner, Lieven, & Tomasello, 2003) – the scope of this debate is not strictly relevant to understanding neural processing in language and other domains as most of these linguistic debates center around the structure of language and syntax rather than on their cognitive comprehension or production.

Many neuroimaging and neurostimulation studies have linked the pars opercularis (Brodmann area 44, a.k.a BA 44, in the left frontal lobe) with syntactic processing and nearby regions such as pars triangularis (BA 45) with semantic processing (Heim, Opitz, & Friederici, 2003; Gough, Nobre, & Devlin, 2005; Friederici & Rueschemeyer, 2003; Goucha & Friederici, 2015). The next section will discuss more fully the complexity of localizing syntactic properties in language, and in other domains as well, to the left inferior frontal gyrus (LIFG), specifically in Broca's area which traditionally comprises BA 44 and 45.

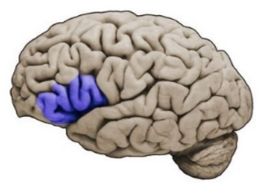


Figure 2. An approximate indication of the cortical surface of Broca's area as traditionally conceived (thanks to Micah Johnson for the image) comprising pars opercularis (BA44) and pars triangularis (BA 45).

Is Syntax Specific to Language?

An assumption of much empirical work on language is that its related processes, such as learning grammatical rules, are unique to that domain (Opitz & Friederici, 2004). Although some researchers leave open the possibility that other cognitive capacities or domains may make use of the same neural streams and regions, especially Broca's area (Rijntjes et al., 2012), the question remains then: Why do many researchers believe that there are language-specific syntactic operations in the LIFG, as opposed to there being more general mechanisms employed by the language system? At least some research has explicitly addressed this, postulating that although domain-general capabilities such as working memory (WM) or cognitive control (CC) also involve the LIFG, it is increased hierarchical complexity, not WM load, which leads to increased activation in this region during processing of syntactically-intricate language (Makuuchi et al., 2009). This work depends on a grammatical feature of the German language which allows researchers to vary the complexity of a sentence without changing its length or order of words.

Additionally, other work claims there is no demonstrable, consistent neural activation overlap between language and WM or CC (Fedorenko, Behr, & Kanwisher, 2011). Still, results such as these do not directly address whether or not syntax as a specific ability is necessarily confined to the language domain.

Syntax and General Mechanisms

Other work, however, does focus on the domain-general capacities associated with syntactic parsing. One review of the literature generally found that, at the least, "domain-general cognitive control and working memory resources are sometimes recruited during language comprehension" (Fedorenko et al., 2014). The implications of this idea, including the relationship between "cognitive control and working memory", will be considered later on in this review. Suffice for now, though, that perhaps the pars opercularis is crucial to linguistic syntax but not particular to it, and other domains may employ the area's abilities in a "general capacity" way (Grewe et al., 2005). One of the more domain-general ways syntax could be employed is in musical processing, a theory considered now before delving into to a wider debate about the function(s) of Broca's area.

Syntax and Music

Interest in connections between language, syntax and musical experiences go back to at least the late 19th century, according to Swain (1997), though this relationship was not widely puzzled over until 1973, thanks to a series of televised Harvard lectures, entitled "The Unanswered Question", given by the famed composer Leonard Bernstein. Bernstein's presentation was inspired by the Chomskyan style of linguistic analysis (arguing that music has generative structure with parallels to language, e.g. that musical phrases are like sentences), though it is widely considered more of an artistic than a scientific statement (Honing, 2009).

Then, starting around the 1980's, the scienctific community began to attempt to provide formal, rule-based accounts of the phrase-structure hierarchies which may govern diatonic – think 'Western' – harmonic progressions and their neural bases (Rohrmeier, 2011). One influential idea, the *generative syntax model* (GSM), describes Western tonal music as a mixture of short term, local relationships as well as a complex assortment of nested, nonlocal dependencies stretching over long ranges of music (Lerdahl & Jackendoff, 1985)[4]. GSM provides explicit rules for harmonic syntax creation and explains how, just like in natural language, most people neither possess nor require explicit structural knowledge of the underlying syntax of a phrase in order to implicitly know whether it is appropriately resolved or not given the preceding context (Koelsch et al., 2013). Put another way, from a young age we non-consciously build up a "long-term store" of a sort of "musical lexicon containing knowledge about timbres, melodic contours, phrases and musical pieces" which is necessary when processing the structure of novel music (Koelsch & Siebel, 2005; Koelsch et al., 2013).

Methods for Studying 'Musical Syntax'

Paradigms for how to empirically test models of tonal syntax have been around for over two decades: Bigand and Pineau in 1997 were the first to present a version of what will be called here the *harmonic expectancy violation paradigm* (HEVP). Working well within the Western tonal musical idiom, the stimuli from these initial experiments manipulated participants' expectations for the last chord of an 8-chord phrase by varying the preceding progression. When chords were less expected (e.g. a fourth harmonic degree following a full cadence (V-I-IV) as

---

[4] An interesting historical footnote, with credit to Alfred Cramer, is that Lerdahl and Jackendoff were in attendance at Chomsky's Norton lectures. Their issues with his lack of scientific rigor in large part prompted their magnum opus, 1985's *A Generative Theory of Tonal Music*.

opposed to simply the more expected V-I authentic cadence)[5], "subjects reported a lower degree of completion" and also "took longer to decide whether the last chord belonged to the sequence" and "whether it was consonant or dissonant" (Bigand & Pineau, 1997). These initial findings from the HEVP formed the basis for the vast majority of behavioral, neural activation, and neuroimaging studies from the last 20 or so years.

Theoretical Pushback

However, the assumptions underlying this approach are not without challenge. Even some who view harmonic structure as hierarchically-structured take issue with the leap to the term 'musical syntax', seeing as there are no strict syntactic categories "such as Noun or Adjective Phrase" in music (Jackendoff, 2009). Some researchers skeptically point out, additionally, that not even all Western tonal music subscribes to the same harmonic rules that are equated with a sort of idiomatic musical grammar (e.g. compare modern jazz with Mozart). Is it possible to provide a consistent account of musical processing, even within just the Western canon?

In an unpublished review paper written between 2008 and 2012 – aptly titled "Music is not a language" – researchers Julian Klein and Thomas Jacobsen argue that the structure of music is so "flexible and integrative" as to render approaches such as the HEVP vastly oversimplified. Although most of Klein and Jacobsen's work centers on the intricacies of music theory not fully accounted for by many researchers (including some cited later in this review), they also point to empirical work which calls into question the validity of saying music has a rigid syntax, per se. For instance, what are we to make of the fact that listener's experience of

---

[5] Expectation here refers to music theoretical harmonic expectation, which **Appendix B** discusses in more detail, and the roman numerals here are standard notation for referring to chords based off of different scale degrees within a key.

and expectations for music changes upon repeated listening (Szpunar, Schellenberg, & Pliner, 2004)? This suggests that HEVP-derived effects reflect predictive mechanisms based on frequency of occurrence more than strict, rule-dependent responses. Keeping the scientific literature entrenched in a classical theory of tonality which does not necessarily correspond with the actual perceptual experiences of many music listeners may limit our ability to understand the reality of music processing.

These points have led some to argue that focusing so much research energy on syntax specifically in music is a mistake. Perhaps truer overlap between music and language can be found by analyzing rhythm, which is also often analyzed as a hierarchical construct (Jackendoff & Lerdahl, 2006), or by comparing prosody and melody (Heffner & Slevc, 2015). With all this pushback in mind, though, a substantial amount of evidence has emerged in favor of music and language jointly depending on some form of cross-domain syntactic computation. Before presenting this evidence, however, it is important to understand the relevant theories and controversies surrounding one of the most controversial and theorized-about brain regions.

Broca's Area

Broca's area is most popularly recognized as comprising BA 44 and 45 and being the area responsible for speech production (Hagoort & Levelt, 2009) – although some researchers would like to expand this centuries-old understanding which may be overly focused on the cortical surface alone (Tremblay & Dick, 2016). Consequently, there is a verifiable "battle for Broca's", especially given its more recent implications in such a wide range of functions, from supporting action perception and goal-oriented planning to WM and syntactic movement (Grodzinsky & Santi, 2008). Also, it is often difficult to tell where Broca's area resides, cytoarchitectonically, in individuals (Amunts et al., 1999). Broca's area, as is true with all neural regions, must also be understood as part of a complex network of brain structures which together, in different combinations, accomplish a wide variety mental tasks.

Additionally, there is increasing focus on the subcortical white matter tracts, or fascicles, which support and connect Broca's, Wernicke's, and Geschwind's (in the inferior parietal cortex) areas (Lemaire et al., 2013; Amunts & Zilles, 2012). Due to this network-centered approach, it is now known that even the right hemisphere has a role in supporting some of the more general attentional, memory, prosodic, and executive needs of linguistic tasks (Vigneau et al., 2011; Friederici, 2011). Yet, even with a nuanced understanding of the limitations of viewing Broca's area as an isolable region, intense debate still focuses on what kinds of processing it may or may not be a hub for.

Broca's Area: A Parser Across Domains?

Some researchers have suggested that Broca's area acts as a *supramodal hierarchical parser* (SHP), in that it underpins domain-general abilities to parse hierarchical information ranging from linguistic syntax and action sequence plans to logical inference problems and musical processing (Tettamanti & Weniger, 2006). The SHP hypothesis is dependent on a conception of syntactic integration operations which process long-distance syntactic hierarchies, localized in Broca's area, and which are distinct from generalized mechanisms such as WM (Fiebach et al., 2005; Friederici, 2004; Bahlmann, Schubotz, & Friederici, 2008). But, is this conception that such syntactic operations operate in the same way, in the same location, outside of the language domain correct?

There is some evidence to support this view. One functional magnetic resonance imaging (fMRI) study found that the posterior portion of the prefrontal cortex, including Broca's area and its right-hemisphere homolog are involved in processing and making decisions about a variety of hierarchically structured sequences, regardless of temporal organization, potentially drawing on the same resources that enable this capability in language (Koechlin & Jubault, 2006). This could indicate that these regions modulate and parse hierarchical information across multiple domains. Additionally, a review of recent neuroimaging literature argued that the IFG (bilaterally) and premotor cortex may work together to keep track of and interpret hierarchical structures across the domains of language, action, and music via a “supramodal syntax” (Fadiga, Craighero, & D’Ausilio, 2009). Though this review focuses on the connections between music and language, it is useful to keep in mind the non-trivial amount of research which views these two systems as but subsets of a larger grouping of cognitive capabilities which are enabled via parsing centered in Broca’s area.

There has been specific pushback against the SHP hypothesis, however, including a variety of neuroimaging studies which show functional dissociations between language tasks mediated by Broca’s area and tasks involving logical deduction and algebraic reasoning (Monti, 2012; Friederich & Friederici, 2009). And, as will be discussed in detail further on, some researchers believe that Broca’s area may not even serve a strictly syntactic function in language, instead more generally contributing “to sentence processing via verbal working memory”, that is, WM as it pertains to processing spoken language (Rogalsky, Matchin, & Hickok, 2008). First, though, this review will consider the literature linking this, increasingly complicated, notion of syntax with music.

Connections Between Music and Language

AD Patel is the most influential theorist and researcher arguing for an overlap in music and language processing. His chief contribution to that debate is the *shared syntactic integration resource hypothesis* (SSIRH) (Patel, 2003; Patel, 2009). The SSIRH outlines how even though cognitive representations of music and language are distinct, the hierarchically structured discrete elements which comprise each domain can still be processed in overlapping areas of the brain. Proof of this shared neural architecture could show that that language processes are not entirely isolable and modular. By extension, Patel and colleagues argue that even the disorders of aphasia and amusia are connected phenomena, with aphasics suffering often unnoticed issues with musical syntax processing which are not attributable to lower-level perceptual or WM problems but instead directly linked to their language deficits (Patel, 2005; Patel et al., 2008). According to this idea, the "two domains use a related functional computation", though they may still be impaired in unique ways (Patel, 2009). One potential objection to this assertion is that determining a common neural substrate for both domains does not imply that precisely the same computation is being executed both cases. Regardless, the SSIRH and its variants have motivated lots of research, both for and against this perspective, the most important of which will be presented in subsequent sections.

Behavioral Evidence

Some of the strongest evidence in support of, and inspired by, SSIRH comes from behavioral studies which draw on the HEVP from Bigand and Pineau, discussed earlier. In such an experiment, it was demonstrated that inserting unexpected chords within a musical progression playing alongside a complex sentence interfered with participant's ability to effectively process linguistic syntax, but, crucially, there was no similar effect on semantic representations (Hoch, Poulin-Charronnat, & Tillmann, 2011). In another study, one of Patel's protégées, Robert Slevc, asked participants to read and make decisions about complex sentences while listening to chord sequences, showing "substantially enhanced garden path effects" when structurally unexpected words were paired with unexpected chords. However, no such interaction followed given unexpected changes in musical timbre (i.e. the quality of the sound, e.g. piano vs. trumpet) or semanticity (e.g. “The dog chased the mailman/promotion”) (Slevc, Rosenberg, & Patel, 2009). Controls like changes in timbre or volume are often used because there is wide consensus that the cognitive organization of tonality in music and syntactic structure in language are separable from lower-level “psychoacoustic principles” such as sound quality or volume (Jackendoff & Lerdahl, 2006). Although, as will be discussed further on, Slevc and others have adopted a different view as to whether it is precisely a syntactic resource which is competed over by these two systems in moments of attentional conflict. In fact, a variation on this 2009 study indeed found the same significant interaction between harmonic violations and semantic garden path sentences (Perruchet & Poulin-Charronnat, 2013). There is still clearly some more informative behavioral work to be done to test the claims of the SSIRH.

Evidence from Neural Activation

Additionally, several event related potential (ERP) studies have shown neural activity overlap between music and language. Testing with magnetoencephalography, for instance, has demonstrated activation in Broca’s area and its right-hemisphere homolog following musical expectation violations within an established harmonic context that are similar to the activation patterns observed after linguistic syntax violations (Maess et al., 2001). Similar results were found in an electroencephalography (EEG) study which held the immediate local context of the critical chord constant and made the violation of musical dependent on the context established at the beginning of the phrase (Koelsch et al., 2013). Additionally, patients with lesions encompassing inferior BA 44 show considerably less such musically-caused EEG activation than those with other lesion-sites or controls (Sammler, Koelsch, & Friederici, 2011). Even children as young as two years’ old have demonstrable ERPs to musical violations, indicating that harmonic expectations are learned early on (Jentschke, Friederici, & Koelsch, 2014). And, in support of Patel's notion that language and music deficits are related, children with specific language impairments do not demonstrate such EEG responses to musical violations (Jentschke et al., 2008). These results are consistent with the idea that Broca's area processes syntax in a "domain-general way" (Sammler et al., 2009). An important caveat, however, is that these pro-SSIRH ERP studies average responses across electrodes, nor do they engage in time-frequency analysis, and therefore cannot comment on the specific neural signatures underlying their observations.

Neuroimaging Evidence

There are a handful of fMRI studies which have sought to more precisely localize where in the brain music and language overlap. One of the first of such experiments supporting this view claims that an entire cortical network (between Broca's and Wernicke's areas) subserves both language and music. This conclusion came from comparing differential activation patterns subsequent to unexpected musical harmonic events, again using changes in timbre as a control, in a sequence with the patterns found from similar experiments in language (Koelsch et al., 2002). This is an even stronger formulation than Patel's SSIRH, but, as will become clear in the next section, there are a wealth of contrary opinions to this 15-year old fMRI study which relied on now out-of-date, possibly invalid analytic techniques (Michael Spezio, personal communication, 11/17/2017).

More recently, it has been proposed that musical syntactic computations are carried out in part via the same dual-stream framework that language uses (i.e. ventral/dorsal tracts from pars opercularis/triangularis to tempoparietal regions), with language being largely left lateralized and music more bilateral (Musso et al., 2015). As a result of this functional neural activation work, there is a nearly canonical understanding in some research communities at this point that, when processing complex music and language stimuli, "high level syntactic integration resources in Broca's area" are drawn upon in both cases (Kunert et al., 2015). However, there is recent, initial evidence that hierarchical, non-locally nested dependencies in music are actually processed primarily, if not exclusively, in the right-IFG (Cheung et al., 2018).

As neuroimaging analytic techniques continue to improve in the coming years, it would benefit the field to renew efforts to untangle the neural regions which encode and process different types of musical and linguistic information. One particularly fruitful direction to explore would be the use of increasingly powerful multivariate decoding models for neural data (Holdgraf et al., 2017). These could allow, for example, researchers to know with greater certainty which specific neural networks are necessary and sufficient for various types of musical linguistic processing, as opposed to relying only on more traditional, univariate, correlational, and therefore inferential, fMRI techniques.

Distinguishing Between Music and Language Processing

Another influential music researcher, Isabelle Peretz, is skeptical that music is "the product of a general-purpose cognitive architecture" or even that it shares much at all in common with language (Peretz, 2006). She believes that the music system is highly modular and innate and that many of the studies cited above are too eager to view Broca's area as being an integrative center for music, language, and possibly other cognitive domains (Peretz & Coltheart, 2003)[6]. Other researchers are similarly skeptical, holding that Broca's area's syntactic abilities

[6] It is nonetheless true that many scholars think that music being modular does not mean it cannot be dependent on the LIFG (Bigand, Tillman, & Poulin-Charronnat, 2006), or, perhaps, made up of several modules, some of which overlap with language (Jackendoff & Lerdahl, 2006).

are reserved for language, a traditionally Chomskyan view, and that structural correspondences between domains are just that: superficial parallels (Friederici et al., 2010).

Much of Peretz's work focuses on neuropsychological, single patient cases which indicate a double dissociation between musical and linguistic productive and interpretative abilities. In contrast with Patel's view, her research suggests, for example, that impairment of "fine-grained pitch perception" is what most commonly leads to amusia, which is distinct from the impairments which lead to cases of aphasia (Peretz & Hyde, 2003)[7].

One fMRI study, directly testing whether Broca's area is a "common substrate...for hierarchical parsing" in music and language, demonstrated substantial non-overlap between the brain areas used to process complex sentences (more ventrolateral) and melodies (more dorsomedial and temporal) (Rogalsky et al., 2011). This stands in direct contrast with Koelsch and others who claim that music and language rely on the same neural pathways. Although this experiment focused on the processing of individual notes in sequence, rather than on harmonic structures, it still suggests that instead of the LIFG being necessary to integrate syntactic structure in music, a "structural violation could trigger" generalized, non-syntactic "working memory processes...or so-called 'cognitive control processes'" (Rogalsky et al., 2011). This idea has gained wider traction, with a recent review article stating: "evidence from the structural-violation paradigm [what is called here the HEVP] is at present largely consistent with an interpretation in which the effects arise within domain-general brain regions that respond to unexpected events across domains" (Fedorenko & Varley, 2016). This perspective, nonetheless, still allows for an interesting debate as to when, why, and which domain-general resources are necessary for music and language and processing.

---

[7] Patel's response would likely be that even if the causes are sometimes distinct there are still cross-domain effects in both amusia and aphasia that may go unobserved.

Shared Reliance on Domain-General Resources in Music and Language

In between more optimistic perspectives and those which deny any meaningful overlap between music and language in syntax processing lies a collection of more nuanced views. While not denying that there is shared neural recruitment across domains, these views emphasize that the cause for this overlap may be a domain-general dependency on cognitive architecture common to a multitude of processes, including music and language. Mauricio Martins and Tecumseh Fitch have advanced the notion that Broca's area functions as but one component in a "working-memory buffer" system, or scannable stack, which works to make sense of hierarchical stimuli across domains (Fitch & Martins, 2014)[8]. Martins has further theorized – empirical work is still pending – that a general executive function such as "passive sequential WM" could be used in such a system to accomplish domain-specific goals (Mauricio Martins, personal communication, 9/26/2017).

Fitch and Martins are not alone in this more recent emphasis on finding an alternate, more general explanation for shared reliance between systems. The linguist Ray Jackendoff, for one, points out that the functional differences in language and music (which he takes to be the expression of propositional thought vs. the enhancement of emotional affect, though this is not a strictly scientific distinction) make it unlikely that their underlying structures have very much in common. In fact, he believes that specific musical-linguistic processing overlaps begin and end with rhythm, whereas harmonic hierarchies, to the extent they are processed in the same way as linguistic syntax, are dependent on "a more general, evolutionarily older function that could be

[8] There is considerable debate as to what such computation might be carried out in language and/or music in Broca's area, possibly via the scannable stack system mentioned here. Refer to **Appendix A** for a fuller discussion of this question, especially with respect to the potential role of recursion.

appropriated by both language and music" (Jackendoff, 2009)[9]. What might that general, older function be exactly? Slevc and Okada (2015) recently published work suggesting that a general cognitive control (CC) mechanism is what accounts for any joint reliance between music and language in Broca's area, rather than a shared need for syntax. This idea will be explored further, as it still unsettled which domain-general mechanism(s) may be implicated.

Accounts of Domain-General Executive Abilities

It is difficult to distinguish clearly between the different executive functions mediated by the prefrontal lobes. Take, for instance, evidence that executive functions such as "switching", "updating", and "inhibiting" are not neuroanatomically separable and may be linked far more than traditionally assumed (Miyake, 2000). In light of this, what should we make of potentially competing theories from Fitch and Martins, who focus on the role of WM, and Slevc and Okada, who focus on CC? Perhaps the dichotomy between WM and CC is a false one, as it has been claimed that these capacities specifically are highly interconnected (Fuster, 2001). That said, the focus here will be on WM as there is more promising theoretical and empirical work already linking WM, language, and music.

Alan Baddeley and Graham Hitch's (1974) influential conceptualization of WM (consisting of phonological loop, visuospatial sketchpad, and central executive components) has been both supported and challenged in subsequent decades[10]. Baddeley himself has argued for the addition of a "fourth component to the model, the episodic buffer...a limited capacity system

---

[9] This perspective differs from the SSIRH in that it takes a difference in structural representation between domains as evidence against shared processing dynamics as opposed to this being an unrelated observation. One could quibble that both Patel and Jackendoff focus too much on the structure of information rather than the nature of its processing.

[10] Recent work has challenged even the notion that short-term memory should be understood primarily in terms of components, a la Baddeley, and puts increasing emphasis on the role of dynamic neural states of attentional control in explaining how WM works (LaRocque, Lewis-Peacock, & Postle, 2014).

that provides temporary storage of information held in a multimodal code, which is capable of binding information" (Baddeley, 2000). This idea fits in well with Fitch and Martins suggestion that musical processing is dependent on such a WM buffer. Furthermore, Baddeley's suggestion that this buffer store may have a limited capacity of around four chunks may be consistent with evidence that even trained musicians, who can better chunk musical stimuli into longer sequences, have difficulty tracking hierarchical dependencies across musical structures longer than around 30 seconds (Tillman & Bigand, 2004).

Initial Evidence for Shared Reliance

There is growing scientific interest in investigating this idea that observed overlaps in language and music processing may be due to a shared reliance on more general mechanisms. This research program especially derives from an emerging literature explicitly linking WM and music. For example, it has been shown that musical experts who can pick out subtle, in-key harmonic transgressions also have increased use of a fronto-temporal network hosting WM compared to non-expert controls (Oechslin et al., 2013).

One study bolstering this perspective used a Stroop task (a classic test of response inhibition in which participants must name the text color of a color word, which may or may not be incongruent with one another). The Stroop task was used in parallel with a music task, showing that "manipulations of harmonic expectancy, but not of timbral expectancy, interacted with Stroop interference effects, suggesting that cognitive control is at least one specific process underlying shared syntactic processing in music and language" (Slevc, Reitman, & Okada, 2013). Again, timbre is used as a control because it is a perceptual but not structural component of music. This approach was inspired by work claiming that the LIFG is part of a CC network which detects incompatible stimuli across domains (Novick, Trueswell, & Thompson-Schill, 2005). Although this paper did not examine the interaction between linguistic syntax violations and Stroop interference as a comparison, these results constitute some of the first good evidence that the shared syntactic integration resource's key to Patel's SSIRH may in reality be shared, domain-general, executive resources.

Implications for Past and Future Work

As has been noted, much of the criticism leveled at the work I have presented in this literature review is that it is overly-focused on tonal violations as a proxy for eliciting what the neural processes are which enable 'musical syntax' processing. Partly in response to this, recent work has begun to shift away from the HEVP (see **Methods for Studying 'Musical Syntax'**), finding some evidence for interaction between syntactically complex sentences and closure judgments of musical sequences which transitioned between keys but didn't include theoretical harmonic violations (Kunert, Willem & Hagoort, 2016). Although the original experiments subsequently presented in this thesis are still steeped in the HEVP (while vigorously questioning some of its assumptions), expansions such as this one from Kunert's team are a positive step for the field as a whole and should be emulated. So, integrating this healthy skepticism that music processing involves domain-specific 'syntactic' resources, how should we understand past findings which have relied on this conception?

At least one research team, reviewing studies of infant music perception, reached the conclusion that early-stage musical abilities are not domain-specific but instead instances of general-purpose mechanisms being recruited by infant's environments (Trehub & Hannon, 2006). Could other previous work be similarly interpreted? It seems likely. For example, much of the ERP and neuroimaging evidence discussed previously could be considered consistent with the notion that overlapping processing could be either domain-specific syntactic processes or domain-general WM processes. It is difficult to interpret the exact character of neural signals, which is why complementary behavioral work is necessary.

However, the majority of the extant behavioral experimental work still takes as an underlying assumption that there is a specifically syntactic computation shared by music and language. This not need be the case. More work in the vein of Slevc et al., 2013 is necessary to reveal the true mechanism supporting observed behavioral and neural overlaps in music and language processing. To this effort we now turn.

***

Experiment 1

The *harmonic expectancy violation paradigm* (HEVP) literature lacks a robust, empirical justification for its principal assumption – that theoretical harmonic violations within tonal musical phrases are interpreted as such by all listeners in the Western musical idiom, regardless of their musical expertise. Further, most studies only make use of one specific violation, the Neapolitan sixth chord, which is based on the lowered second scale degree of major keys (thereby making it out of key and dissonant, e.g. a D flat major chord within a C major context). These two issues prompted Experiment 1, since Experiment 2 relies on musical stimuli in the vein of the HEVP. Experiment 1 examined both trained musicians' and non-musicians' interpretations of a variety of musical sequences which either did or did not contain theoretical harmonic violations. The main questions this experiment sought to address were: 1) What are the group (i.e. musician vs. non-musician) differences in evaluating the congruency of musical progressions which either contain or do not contain theoretical violations? 2) What are the group differences in evaluating various conditions of expected and violation-containing progressions? 3) Might any other variables – including the timbre of the presentation, length of the progression, and position of the error within the sequence – impact these observations?

Method

Participants

20 undergraduate students (9 males) at the Claremont Colleges in Claremont, California were recruited and paid $5 each for participation in this 15-minute experiment. 17 participants grew up primarily in the United States, 1 in Spain, 1 in Thailand, and 1 in Canada, and all were native English speakers or fluent in English as a second language. Their ages ranged from 18-27 (mean = 19.74 years).

Participant classification

As noted by others (e.g. Hanna-Pladdy & MacKay, 2011), classifying musicians and non-musicians is fraught, especially within a highly educated population where most have been given at least some musical training at some point in their life.

In her BA thesis, “The Effect of Lifelong Music Engagement on Executive Function in Older Adults”, Ellie Abrams (Pomona College, 2017) attempted to group subjects by musical ability. Her resulting musicians group had an average age of musical acquisition of 6.4 years, 14.6 years of formal training, 100% understood musical notation, 62.5% had a good grasp of music theory, they all played at least 2 instruments, and practiced on average 11.3 hours per week. Abrams classified participants as non-musicians if they had less than 5 years of formal training, acquired an instrument after the age of 10, and had no recent musical experience. While these metrics could not work precisely with a younger group, they informed Experiment 1’s participant classification scheme.

11 participants self-identified as musicians before testing, 7 said they were not musicians, and 2 responded “Not sure”. All participants were asked these questions about their musical background: “For how many years have you played music?” (m = 9.3 years), “For how many years have/did you receive musical training?” (m = 6.75 years), “If applicable, at what age did you begin music lessons?” (m = 9.84 years), “If applicable, what instruments, including voice, do you play?” (m number of instruments = 2.21), “If applicable, how many hours per week do you currently play/practice music?” (m = 3.85), “How well do you understand written music” on a Likert scale of 1 (“Not at all”) – 5 (“Perfect understanding of one or more clefs”) (m = 3.5), and “How well do you understand music theory?” on a Likert scale of 1 (“Not at all”) – 5 (“Advanced grasp of a wide range of topics”) (m = 2.79).

Upon analysis of participant responses to these questions about their musical background, it was decided that musicians would be considered those who fit 3 of the following 5 criteria: at least 10 years playing music, music lessons begun by age 10, a rating of 3 or higher on the written music question, a rating of 3 or higher on the music theory question, and currently playing at least 2 hours per week. Conversely, participants were classified as non-musicians if they fit 3 of the following 5 criteria: less than 10 years playing music, music lessons begun after the age of 10, a rating of less than 3 on the written music question, a rating of less than 3 on the music theory question, and not currently playing music. Following this method, all participants who self-identified as not being musicians and two who did (n = 9) were placed in the non-musician category. 9 of the participants who self-identified as musicians and the 2 who were not sure (n = 11) were classified as musicians. This demographic data is reported in Table 1.

Participant Pre-Testing Survey Descriptive Summary

	Years of training		**Number of instruments**		**Years playing**		**Hours per week**		**Grasp of written music**		**Grasp of music theory**	
	0	**1**	**0**	**1**	**0**	**1**	**0**	**1**	**0**	**1**	**0**	**1**
Number	9	11	9	11	9	11	9	11	9	11	9	11
Mean	3.67	9.27	1.78	2.36	4.78	13.00	1.78	5.55	2.44	4.36	1.56	3.82
Std. Deviation	2.12	3.66	1.09	0.67	2.64	2.28	3.03	6.11	1.01	0.67	0.53	1.17

Table 1. Group descriptive statistics following participant classification from the pre-testing survey. Non-musicians are coded as 0 and musicians are coded as 1.

Materials

Musical stimuli were divided into 6 conditions: 3 contained a theoretical harmonic violation and 3 did not. Each condition contained 12 musical chord sequences (1 sequence in each major key * 6 groups of sequences = 72 total stimuli), 4 of which were 6 chords long, 4 were 7 long, and 4 were 8 long. For the sequences containing violations, the position of the violation was counterbalanced such that occurred with equal frequency in the final, penultimate, third-to-last, or fourth-to-last serial position of the sequence.

The 3 conditions of chord sequences with no violation (Expected I, Expected II, and Expected III) were structured as such: Expected I followed a standard I V vi IV V I progression ending on an authentic cadence (e.g. C G Am F G C) when it was of length 6, I V vi IV I V I at length 7, and I V vi IV V I V I at length 8; Expected II was slightly less standard in that it contained an in-key ii chord (I vi IV ii V I, I vi IV ii I V I, and I vi IV ii V I V I); Expected III varied on Expected I by containing a V7 chord as part of the closing authentic cadence (I IV V vi V7 I, I IV V vi I V7 I, I IV V vi V I V7 I). Those sequences with a theoretical violation (Violation I, Violation II, and Violation III) were designed such that 1/3 of each of them followed the patterns in Expected I, Expected II, and Expected III. All of the violation sequences varied from the no violation sequences in that they included an out-of-key chord: Violation I had a Neapolitan sixth (flat II) chord (e.g. G Em C G# D G), Violation II had a flat vii (e.g. E C#m A F#m B Dm), and Violation III had a sharp IV (e.g. D G A G# A D).

The stimuli were composed in the program MuseScore 2.1 (https://musescore.org/en) and exported as chord audio files with a duration of 1 second each using the grand piano and orchestral (cello and violin) timbre effects. The Python library Expyriment 0.9 (http://www.expyriment.org/; Krause & Lindermann, 2014) was used to code the stimuli presentations and collect data. Statistical analysis was done using the software programs Jasp 0.8.4 (https://jasp-stats.org) and SPSS 25 (https://www.ibm.com/products/spss-statistics).

Procedure

Upon arriving at the testing room, participants completed a pre-test survey on a computer, detailed in **Participants**. Upon completion, they moved to a second computer, put on headphones, and began the testing session, which lasted about 12 minutes. The instructions explained that they would be presented with a series of musical chord sequences, half played on a piano and half with strings, and after each one they would be asked whether or not the entire sequence sounded congruous or whether it contained an incongruous acoustic event. Example sequences with and without a theoretical violation in them (1 of each) were played for reference. Participants were then pseudo-randomly presented with the 72 phrases detailed in **Materials**. After each phrase, participants were given 3 seconds to respond whether or not they heard a violation. Upon completion, each participant was paid and debriefed.

Results

Demographic predictors of overall performance

A linear regression was performed using participant data from the pre-testing survey. The continuous criterion variable was overall mean accuracy scores in the experiment. Pearson's r correlations were calculated to choose predictor variables to put into the regression model. The accuracy criterion was statistically significantly correlated with participants' self-classification as a musician or not ($r = .644$, $p = .002$ where coding scheme was 1 = musician and 0 = not), the number of years they were playing ($r = .618$, $p = .004$), hours per week spent playing ($r = .502$, $p = .024$), and grasp of music theory on a 1-5 scale where 5 is better ($r = .661$, $p = .002$). This left out participant age, gender, number of years getting formal lessons, number of instruments, and grasp of sheet music as potential predictors. Of the selected predictors, there were internal correlations between self-classification and number years playing ($r = .870$, $p < .001$), self-classification and grasp of theory ($r = .784$, $p < .001$), grasp of theory and number years playing ($r = .853$, $p < .001$), and grasp of theory and number of hours per week ($r = .510$, $p = .022$).

A general entry model was statistically significant [R = .721, R^2 = .520, adjusted = .392, F (4, 19) = 4.061, p = .02] but, due to issues of multicollinearity, no individual model coefficient was a statistically significant predictor. So, a stepwise approach was applied with the same model terms, which was also statistically significant [R = .661, R^2 = .437, adjusted = .406, F (1, 19) = 13.97, p = .002] with only grasp of theory as the remaining contributing coefficient [B = .049, SE = .013, beta = .661, t (19) = 3.737, p = .002], accounting for 43.7% of variance. For each one unit increase in self-reported grasp of theory on a 1-5 scale, mean accuracy scores increased 4.9%. Though this predictor is the best of the bunch, this is not to say that no other variable from the pre-testing survey accounts for some of the variance in this experiment's dependent measure[11].

Group differences within and across type of progression

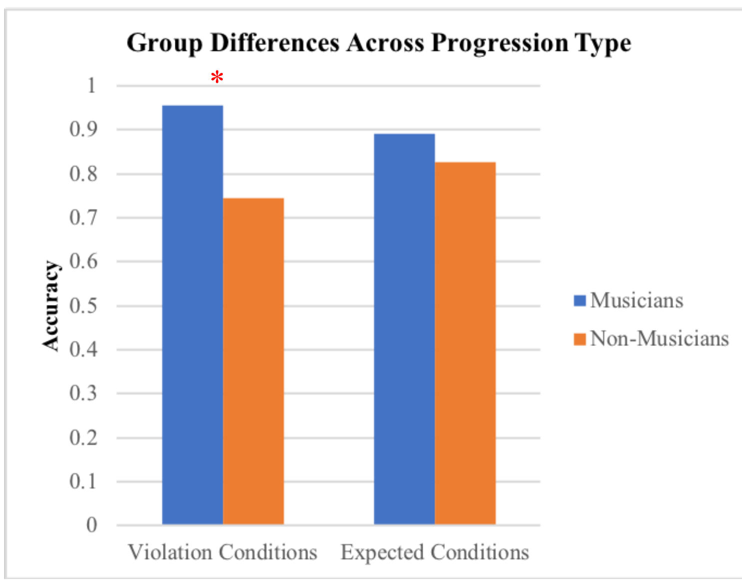


Figure 3. Mean accuracy for musician and non-musician groups for the Violation and Expected conditions collapsed across the three conditions of each. Significant group differences are indicated with a red star.

[11] Post-hoc examination of the 20 individual accuracy scores (indicated as proportion correct) across all conditions of the experiment (min = .60, max = .99, m = .86, SD = .11) revealed two potential participant outliers whose means were unexpected, according to both their self-identification and group classification, above or below (though still within one standard deviation of) the population mean: a non-musician with an overall accuracy of .94 and a musician with an overall accuracy of .76. To retain power – and acknowledging that no classification scheme for musicianship can be perfect – these participants were retained for data analysis.

A mixed 2 (Violation Conditions, Expected Conditions) x 2 (Musicians, Non-Musicians) repeated measures ANOVA revealed a statistically significant main effect of group on mean accuracy scores [F (1, 18) = 12.78, p = .002, η^2 = .415], but no main effect of progression type. There was additionally a marginal interaction approaching significance between group and progression type [F (1, 18) = 4.344, p = .052, η^2 = .194]. As indicated in Figure 3, there was a statistically significant group difference between musicians (m = .955, SD = .057) and non-musicians (m = .744, SD = .124) accuracy within the Violation conditions [t (18) = 5.037, p < .001], but not within the Expected conditions. Additionally, paired sample t-tests indicated there were no statistically significant differences within either the non-musician or musician groups' scores due to progression type. Overall, this pattern reflects musicians generally superior performance in this experiment, especially when asked to identify whether a sequence contained an incongruent chord.

Group differences across and within individual conditions

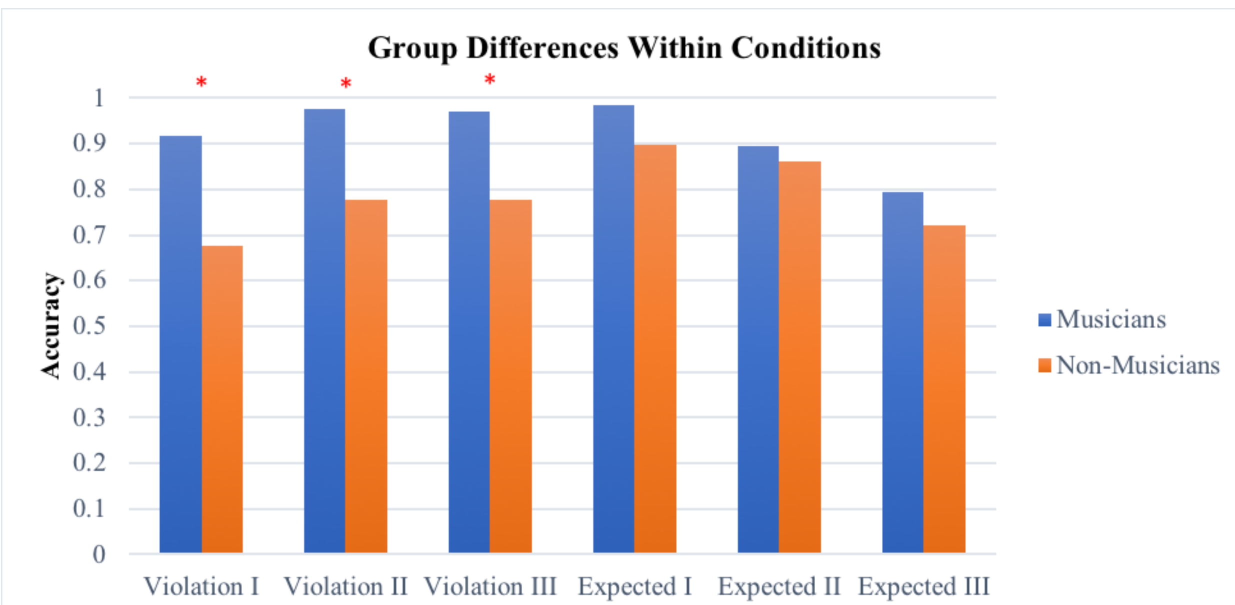


Figure 4. Mean accuracy for musician and non-musician groups for across each of the six experimental conditions. Significant group differences are indicated with a red star.

To test the patterns of mean accuracy scores across groups and conditions, two mixed 3 x 2 (Condition x Group) repeated measures ANOVAs were performed: one for the three Violation conditions and one for the three Expected conditions. The Violation conditions ANOVA revealed a statistically significant main effect of condition [F (2, 36) = 4.129, p = .024, ${\eta_p}^2$ = .187] and of group [F (1, 18) = 25.37, p < .001, η^2 = .585] but no interaction between condition and group. Post-hoc analysis showed that the main effect of group reflects a mean accuracy difference of .211 in favor of musicians [t = -5.037, p_{bonf} < .001] and the main effect of condition is reflective of statistically significant differences between Violation I and Violation III [t = -2.781, p_{bonf} = .036]. Using a series of t-tests to look further into the within-group differences, musicians did better [t (10) = 2.390, p = .038] responding to Violation II (m = .977, SD = .039) than Violation I (m = .917, SD = .112), with Violation III (m = .970, SD = .056) not differing from either. On the other hand, non-musicians only approached doing statistically significantly better (p = .056) responding to Violation III (m = .778, SD = .167) than Violation I (m = .676, SD = .169), with Violation II (m = .778, SD = .161) also not differing from either. And, as is made clear in Figure 4, t-tests revealed statistically significant group accuracy differences between musicians and non-musicians in the Violation I [t (18) = 3.823, p = .001], Violation II [t (18) = 3.983, p < .001], and Violation III [t (18) = 3.596, p = .002] conditions, with non-musicians performing worse across all conditions.

The Expected conditions ANOVA revealed a statistically significant main effect of condition [F (2, 36) = 10.938, p < .001, ${\eta_p}^2$ = .378] but no main effect or group or interaction between condition and group. Post-hoc analysis showed that the main effect of condition is reflective of differences between Expected I and Expected III [t = 4.142, p_{bonf} = .002] as well as Expected II and Expected III [t = 2.746, p_{bonf} = .039]. Using a series of t-tests to look further into the within-group differences, musicians did better responding to Expected I (m = .985, SD = .034) than to either Expected II (m = .894, SD = .163) [t (10) = 2.292, p = .045] or Expected III (m = .795, SD = .254) [t (10) = 2.731, p =.021]. On the other hand, non-musicians did better [t (8) = 3.23, p = .012] responding to Expected I (m = .898, SD = .166) than Expected III (m = .722, SD = .177), with Expected II (m = .861, SD = .138) not differing from either.

Overall, these results reveal both musicians' and non-musician's relative deficits in identifying sequences containing the theoretical tonal violations embedded in Violation I sequences (which contained Neapolitan sixth chords) and non-musicians consistent deficit in identifying violations compared to musicians. Additionally, across groups participants were relatively superior in identifying a lack of violation in Expected I sequences (containing only common I IV V vi chords) and struggled with Expected III sequences containing dominant seventh chords in the final cadence.

Other factors

Two additional mixed repeated measures ANOVAs (Extra Factor x Group) were performed to test if either the additional variables of timbre of presentation (either Piano or Strings) or length of presentation (either 6, 7, or 8 chords long) impacted participant performance between groups. The former revealed an expected main effect of group [F (1, 18) = 12.78, p = .002, η^2 = .415], reflective of musicians generally superior performance, but no main effect of timbre on mean accuracy scores in the experiment or any interaction between group and timbre. Similarly, the latter ANOVA revealed the same main effect of group but no main effect of length or interaction between length and group.

With respect to the position of the error in the violation sequences, the data is more difficult to unpack. As a result of the experimental design, out of the 12 sequences in each violation condition, each participant was presented with 1 position 3 error and 1 position 8 error (3 total of each), 2 each of position 4 and 7 errors (6 total), 3 each of position 5 and 6 errors (9 total). This imbalance, especially within a classification design, can lead to distortions in analyses of covariance (Schneider, Avivi-Reich, & Mozuraitis, 2015). However, it is less fraught and still informative to analyze the potential impact of error position on performance across groups and within the individual violation conditions.

Another repeated measures ANOVA proved impossible due to sphericity issues (Mauchly's W = .098, $p < .001$), even after applying Greenhouse-Geisser and Huynh-Feldt corrections. Therefore, a 6 (Error Position: 3, 4, 5, 6, 7, 8) x 2 (Group) multivariate ANOVA was performed. There were statistically significant multivariate main effects of group on mean accuracy at the various points of violation insertion [$\lambda = .212$, F (6, 13) = 8.048, p = .001, ${\eta_p}^2 = .788$]. A series of follow-up univariate ANOVAs on each of the dependent variables revealed statistically significant group differences for error positions 4 [F (1, 18) = 22.810, $p < .001$, ${\eta_p}^2 = .559$] (such that musicians [m = .939, SD = .112] performed better than non-musicians [m =.574, SD = .222]), 5 [F (1, 18) = 15.707, p = .001, ${\eta_p}^2 = .446$] (such that musicians [m = .949, SD = .091] performed better than non-musicians [m =.679, SD = .204]), and 6 [F (1, 18) = 13.290, p = .002, ${\eta_p}^2 = .425$] (such that musicians [m = 1.00, SD = .0] performed better than non-musicians [m =.827, SD = .158]). It should be noted that error positions 4, 5, 6, and 8 each failed Levene's Test of Equality of Error Variances [F (1, 18) = 4.562; 5.810; 34.957;12.989 and p = .047; .027; .000.; .002 respectively]. There were no group differences at error positions 3, 7, and 8 between musicians [(m = .848, SD = .229), (m = .955, SD = .078), (m = .97, SD = .101) respectively] and non-musicians [(m = .704, SD = .261), (m = .87, SD = .111), (m = .815, SD = .242) respectively].

As a final point of investigation, errors by position were grouped into two variables – early (error positions 3, 4, 5) and late (error positions 6, 7, 8) violation. A 2 (Group) x 2 (Time) repeated measures ANOVA revealed a main effect both of group [F (1, 18) = 22.99, p < .001, η^2 = .561] and time [F (1, 18) = 28.327, p < .001, ${\eta_p}^2$ = .611] as well as a statistically significant interaction between these factors [F (1, 18) = 5.911, p = .025, ${\eta_p}^2$ = .250]. The main effect of time reflects generally superior performance in responding to late (m = .898, SD = .101) than to early (m = .792, SD = .170) violation [t (19) = 4.541, p < .001]. Meanwhile, the interaction reflects that, while performance between early and late violations improved in both groups, musicians had less room to improve (from m = .901, SD = .109 for early violations to m = .961, SD = .045 for late violations) than non-musicians did (from m = .659, SD = .135 in early to m = .821, SD = .097 in late).

Discussion

This is one of the first analyses of how both musicians and non-musicians judge the congruency of a variety of Western tonal chordal music sequences which either do or do not contain theoretical violations of their harmonic context[12]. Previous experiments (e.g. Slevc et al., 2009; Rogalsky et al., 2011) which have used variations of the HEVP have either not investigated or not reported significant group differences on their critical measures based on individual participants' musical background. Further, the pre-existing literature seems to have exclusively depended on using sequences which simply are or are not violations (e.g. just presenting sequences with or without Neapolitan sixth chords included) without seriously

[12] It would be incorrect to suggest this is the first attempt to relate musical experience with musical cognitive capacities – quite the contrary is true. **Appendix B** presents a portion of a final project completed for Alfred Cramer's Spring 2018 "Linguistics in Musical Analysis" class which provides further context for this inquiry which, while interesting and relevant to Experiment 1's findings, goes beyond the scope of the literature review or discussion here.

considering more broadly what is considered a 'violation' by their participants or whether this blanket categorization (violation or not) may be too simple given the complex flexibilities of any musical tradition.

The most notable result here is that non-musicians consistently performed worse at identifying theoretical violations than those participants with more musical training. This is especially interesting in the Violation I condition, in which the violation was a Neapolitan sixth chord, since this is the most common violation assumed to be universally understood as such in the HEVP literature (80.83% correct identification across all participants, the second worse of any condition). And, although there were no apparent group differences in the Expected conditions, participants did universally poorly identifying sequences of the Expected III variety – which contained in-key dominant V7 chords in the final cadence – as being congruous in context (76.25% correct identification, the worst of any condition).

These data are informative in the creation of useful chordal sequences for further experiments needing harmonic musical stimuli. While it seems the timbre or length of the sequences are not important factors, the position within the sequence in which the error occurs is. Musicians significantly outperformed non-musicians in error identification in the 4^{th}, 5^{th}, and 6^{th} positions, though by the 7^{th} and 8^{th} positions performance was comparable. There was also a trend towards better identification of theoretical violations near the end of sequences (i.e. in the 6^{th}, 7^{th}, and 8^{th} positions). It may be informative that between-group performance also didn't significantly differ in the 3^{rd} error position, though interpreting these data is hampered by a low number of such trials. A further experiment could more systematically assess the impact of serial error position on measures of tonal music congruity.

If the goal of a given study is to get close to 100% accuracy of behavioral identification of theoretical violations (as is the case in **Experiment 2**), then there is good evidence here that the Expected I (which contained just canonical I IV V vi chords) condition should be emulated (94.58% accuracy, the highest of any condition). Choosing which violation to use will still be tricky, though, especially if HEVP approaches are fundamentally flawed. Nonetheless, both Violation II (an insertion of a flat vii chord, 88.75% accuracy) and Violation III (sharp IV chord, 88.33% accuracy) both seem like good candidates due to their relatively high accuracy rates. However, both of these conditions still proved significantly more difficult for non-musicians than musicians.

Even though most participants in this study had some sort of musical training background, clear group differences with strong effect sizes still arose. What accounts for this? The best predictor of performance on this task, as shown via a regression analysis, was self-assessed knowledge of music theory, i.e. explicit familiarity with the 'rules' governing Western tonal music. This is at least one point of major difference, then, between music and language – even native English speakers who have never taken a class on English grammar can easily 'intuit' when a grammatical rule is being violated. While the current results should be further tested and validated, both brain and music scholars alike should take note of the implications (see **Appendix B**) of these findings in their experimental designs and theories of 'musical syntax' going forward.

Experiment 2

To date, no other study has directly compared the relative dependence of language and music processing on general working memory (WM) capacities. Previous research in support of Patel's *shared syntactic integration resource hypothesis* (SSIRH) using the HEVP has identified neural responses to 'incongruent' musical chords as syntactic computations, identical or related to those necessary for the language domain. It is possible that, instead, such responses may be attributable to music's reliance on domain-general WM resources used to track and integrate a variety of sequential information. Therefore, Experiment 2 tested effects of WM load on the processing of linguistic and musical phrases (as measured by response times (RT) and accuracy) with and without violations. To engage WM resources, a numerical 2-back task was used. In wide use since 1958 as a measure of WM, the 'n-back' task is "a complex measure involving multiple processes that seem to be largely stimulus and material independent" (recall from Miyake et al., 2000 that "executive functions are not unitary"). Nevertheless, it reliably correlates with other tests of WM and general fluid intelligence (Jaeggi et al., 2010).

It was hypothesized that interference while processing structural violations in music and a simultaneous WM task would lead to worse relative processing efficiency than when there was no such violation. Further, it was hypothesized that this effect would be more pronounced in the music conditions than in comparable linguistic conditions. Such findings would constitute further evidence against previous interpretations of interference between language and music processing as demonstrating a joint reliance on syntactic mechanisms (e.g. Hoch, Poulin-Charronnat, & Tillmann, 2011). Instead, in line with more recent proposals (e.g. Slevc & Okada, 2015; Perruchet & Poulin-Charronnat, 2013), such results would be consistent with a theory of shared use across domains of more general cognitive capacities, especially WM.

Method

Participants

23 undergraduate students (15 males) at the Claremont Colleges in Claremont, California were recruited and paid $10 each for participation in this 30-minute experiment. 20 participants grew up primarily in the United States, 1 in Canada, 1 in Mexico, and 1 in Ecuador. 21 were native English speakers and 2 were Spanish native speakers now fluent in English. Their ages ranged from 18-22 (mean = 20.78 years). The same classification scheme for musicianship as described in Experiment 1 was used here. This resulted in 12 participants being classified as musicians and 11 as non-musicians.

Materials

This investigation employed a Domain Type (Linguistic or Musical) x Structure[13] (Expected, Violation, or Control) factorial design with performance on a simultaneous WM task as the primary dependent measure. Alongside presentation of auditory material, in the form of sentences or chord progressions which either contained a theoretical structural violation or did not, numbers comprising the 2-back task were visually presented. Each trial was 7 units long (i.e. 7 words or 7 chords along with 7 corresponding numbers). The lists of numbers were pseudo-randomly generated and then edited to create 2-back sequences, with the 2-back repetition stimulus (hereafter referred to as a 'go') appearing at a critical position (e.g. 4 5 7 2 9 2 3 for a critical position of 6). There were 16 critical trials each for the musical and language sequences, 8 of which contained a structural violation in the same critical position as the target 2-back number. These critical positions were, in equal proportion, in the 7th, 6th, 5th, and 4th serial

---

[13] Structure from here onwards refers to syntactic structure in the language condition and harmonic structure in the music condition. The term syntax is not used because, as discussed previously, there is still considerable controversy over whether it is an appropriate term to use outside of the language domain.

positions. Therefore, for the critical trials, there were 8 each of the No Violation Language, Violation Language, No Violation Music, and Violation Music conditions.

In addition to the 32 critical trials, there were 16 each of practice, control, and catch trials. The practice trials consisted of 2 practice trials each for the linguistic and musical task in isolation, 4 practice trials for the 2-back task in isolation, and 2 practice trials for each condition of the critical trials. The catch trials (not used for analysis) contained an equal mixture of trials without a 2-back repetition 'go' stimulus, multiple such 'go' moments within a trial, and trials with 1-back and 3-back sequences to ensure the trials did not become predictable. The control trials (8 each of music and language) were constructed similarly to the critical violation trials, except they contained a change in voice or timbre (i.e. a female instead of a male voice and strings instead of piano) instead of a structural violation at the critical 'go' moment. This type of control selection is in line with previous research (e.g. Slevc, Rosenberg, & Patel, 2009) and is used to ensure that any observed interference effects are due specifically to the structural nature of the violation and are not merely elicited by any salient, unexpected acoustic event.

Musical stimuli

The musical stimuli were again created MuseScore 2.1. They consisted of a series of simple tonal chord progressions with or without an out-of-key chord (e.g. C G Am F C G C vs. C G Am F G <u>F#</u> C). The chords were recorded using a grand piano timbre and lasted for 1 second each. They were also recorded using an orchestral string timbre for use in the control trials. The chord patterns followed standard harmonic progressions (e.g. I V vi IV V I), cycling through the 12 major keys. For the critical 'go' moments within the violation conditions, a sharp IV chord (e.g. an F# chord within a C major context) was inserted (refer to **Experiment 1** for a discussion of this choice).

Linguistic stimuli

The linguistic stimuli, consisting of a series of simple English sentences with or without a grammatical subject-verb or number agreement error (e.g. "The traveler has packed for her journey." vs. "The child will have been plays alone.") were recorded via the Amazon Polly™ (https://aws.amazon.com/polly) text to speech software using the Matthew voice. Words used in the control trials were recorded using the Joanna voice. The sentences were adapted from stimuli used by Micah Johnson for work in Martin Monti's UCLA lab and were similar in terms of syntactic complexity and semantic arousal levels. 5 participants from Experiment 1 (3 female, 1 non-native English speaker) were recruited and paid $5 to test and validate the linguistic stimuli in isolation, listening for and responding to grammaticality. Sentences which were mis-interpreted by more than one participant were edited for clarity.

Stimuli presentation

Auditory stimuli were edited using the Audacity™ 2.2.0 (https://www.audacityteam.org) audio editing software. The experiment was coded, presented and the data accumulated using the experimental software application PsychoPy 1.82 (http://www.psychopy.org/; Peirce, 2009). Due to issues with millisecond (ms) accuracy from commercial keyboards, participants responded using an external dedicated timer box from ioLabs (http://www.iolab.science/) which has negligible jitter variability (< .03 ms) and correspondingly excellent RT precision. Additionally, to account for variation in frame presentation rates on the Mac computer used for testing, all stimuli and button box responses were synchronized based on the frame of presentation rather than via a clock. Because the frame duration on this computer was consistently slightly less than 60 Hz, each stimulus was presented for 60 frames, or approximately 1 second, further ensuring accuracy in RT collection. The stimuli used in this experiment can be found in **Appendix C**.

Procedure

Upon arriving at the testing room, participants completed a pre-test survey on a computer and gave their signed consent to the experiment. The survey collected demographic information on participants as well as their linguistic and musical backgrounds, similar to what was done in Experiment 1. Upon completion, they moved to the testing room, sat at the testing computer, put on headphones, and began the practice and instruction session. Participants were guided through the practice trials to ensure they were responding in the correct way and given a chance to ask any questions before beginning the test trials. They were trained to respond with the ‘no’ button every time the number presented did not match the 2-back criterion and the ‘yes’ button when it did. They were further instructed to answer as quickly and accurately as possible. Participants were told to listen for grammatical violations in the language conditions and for incongruent chords in context for the musical conditions. This instruction and practice session lasted about 6 minutes.

The experimenter then left the testing room and participants were left to complete the rest of the experiment, which lasted about 22 minutes. Each of the experimental trials went as follows: a 1-second slide said: “Get ready for the next presentation”. Then, the 7 audio items in a music or language sequence were presented, 1 item at a time, simultaneously with the numeric sequence, with 1-second fixation crosses between each presentation. At the end of this sequence, a screen asked: “Did the preceding presentation contain a violation”? A ‘yes’ or ‘no’ response (or 5 seconds elapsing with no response) triggered the next trial to begin. Upon completion of all trials, each participant was paid and debriefed.

Results

The following data were collected for each critical and control trial for each participant: accuracy and RT for the numeric 'go' presentation in each trial as well as for the violation question at the end of each trial. There were several types of errors participants could commit in this experiment. A false positive occurred when they responded 'yes' to a number that was not a 2-back number, but these were not used for analysis. Conversely, a false negative was recorded to any 'no' response to a 2-back number. Participants could also miss entirely by not responding in time to a numeric presentation. This was a common occurrence, especially in the beginning of sequences. Participants were excluded who committed 10 or more (> 20%) misses in the critical 'go' 2-back moments. This led to 4 participants being excluded, and, therefore, 19 participants were included in the final analyses (11 musicians and 8 non-musicians). These remaining participants committed an average of 3.74 critical misses of the 2-back target (7.79% of trials) and 3.21 false positives (6.69%). Analyses focused on true positive (correct) responses and false negative (incorrect) responses to the critical 2-back 'go' moment in each trial (that is, only looking at recorded button responses). The analyses were completed using Jasp.

Analysis of directly measured variables

A series of within-subject 2 x 2 [Domain (Linguistic vs. Musical) x Structure (No Violation vs. Violation or Control)] repeated measure ANOVAs were performed for each of the following 4 dependent variables (DVs): accuracy in responding to the critical 2-back number (Number Accuracy), 2-back RT for correct responses (Number RT), post-trial violation question response accuracy (Question Accuracy), and RT for correct responses to this question (Question RT). For each DV, two ANOVAs were performed because the second level of the structure factor (Violation or Control) changed in order to compare the effects of both structural violations (the critical Violation trials) and changes in timbre or voice (the Control trials) relative to the No Violation trials.

For Number Accuracy, there was a main effect of structure on responses in both the Musical and Linguistic Violation analysis [F (1, 18) = 6.848, p = .017, η^2 = .276 with post-hoc tests revealing a mean difference of .079 (No Violation > Violation), t = 2.27, p_{bonf} = .029] and in the Control analysis [F (1, 18) = 8.355, p = .01, η^2 = .217 with post-hoc tests revealing a mean difference of .105 (No Violation > Control), t = 2.622 , p_{bonf} = .013] (Figures 5 and 6). There were no statistically significant main effects or interactions for either of the Number RT analyses (Figures 7 and 8).

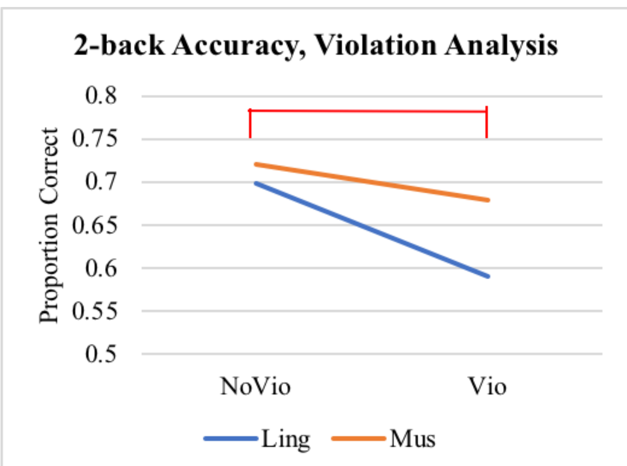


Figure 5. Mean proportions correct for the 2-back task for the Linguistic and Musical Violation and No Violation conditions. The red bracket denotes a main effect of structure (No Violation > Violation).

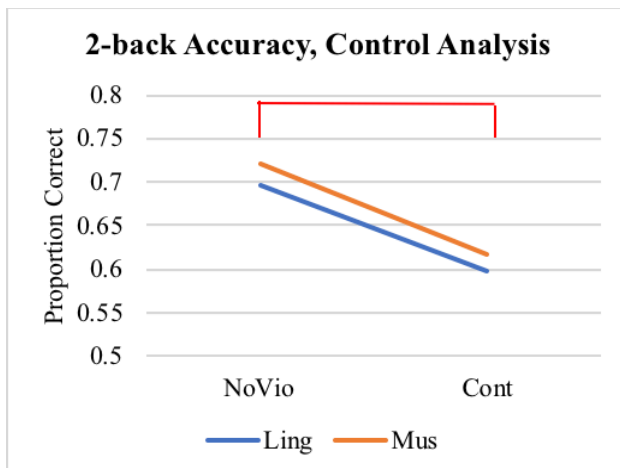


Figure 6. Mean proportions correct for the 2-back task for the Linguistic and Musical Violation and Control conditions. The red bracket denotes a main effect of structure (No Violation > Control).

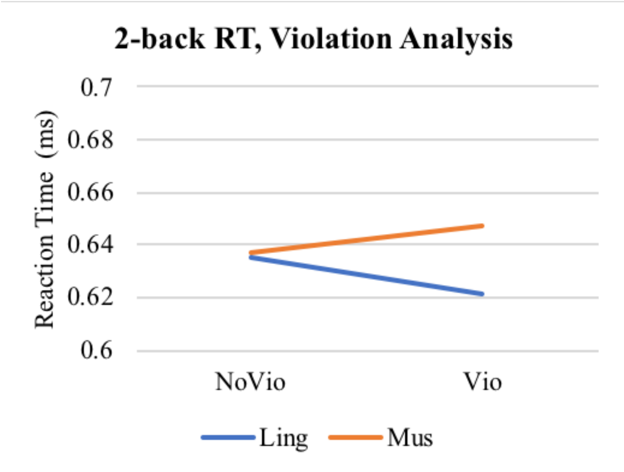


Figure 7. Mean reaction times for the 2-back task for the Linguistic and Musical No Violation and Violation conditions.

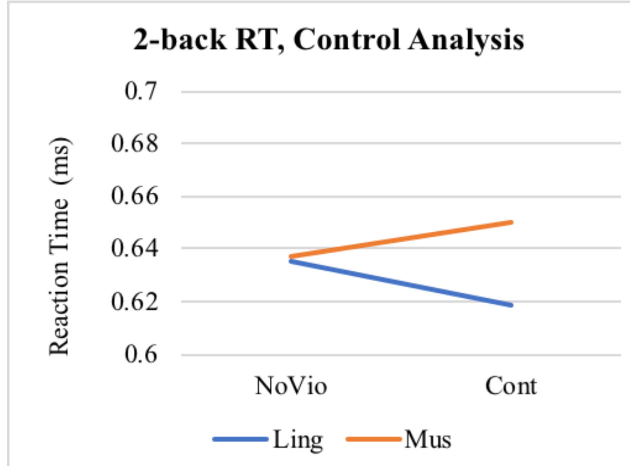


Figure 8. Mean reaction times for the 2-back task for the Linguistic and Musical No Violation and Control conditions.

The Question Accuracy Violation analysis revealed a main effect of domain [F (1, 18) = 5.1, p = .037, η^2 = .221 with post-hoc tests revealing a mean difference of .082 (Linguistic > Musical), t = 1.972, p_{bonf} = .056] and a statistically significant interaction between structure and domain [F (1, 18) = 14.674, p = .001, η^2 = .449]. Follow-up t-tests showed a statistically significant difference here between the Musical No Violation (m = .921, SD = .095) and Violation (.691, SD = .296) conditions [t (18) = 3.005, p = .008] (Figure 9). The Question Accuracy Control analysis also revealed a main effect of domain [F (1, 18) = 6.886, p = .017, η^2 = .227 with post-hoc tests revealing a mean difference of .115 (Linguistic > Musical), t = 2.19, p_{bonf} < .035], an additional main effect of structure [F (1, 18) = 12.734, p = .002, η^2 = .414 with post-hoc tests revealing a mean difference of .161 (No Violation > Control), t = 3.035, p_{bonf} = .004], and a similar interaction between structure and domain [F (1, 18) = 12.827, p = .002, η^2 = .416] (Figure 10). Follow-up t-tests similarly showed a statistically significant difference between the Musical No Violation (m = .921, SD = .095) and Control (m = .592, SD = .386) conditions [t (18) = 1.769, p = .001].

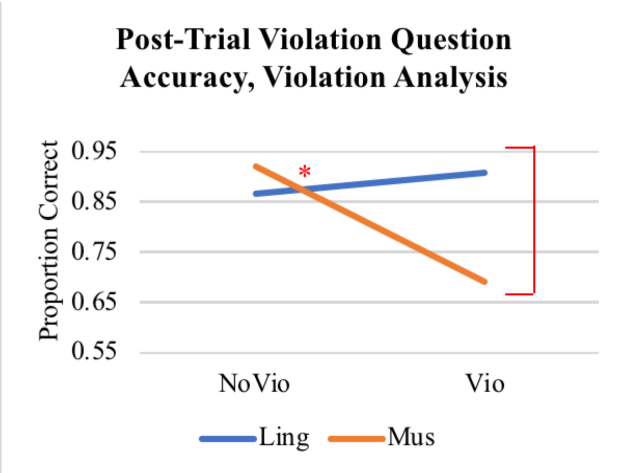


Figure 9. Mean proportions correct for the post-trial violation question for the Linguistic and Musical No Violation and Violation conditions. The red bracket denotes a main effect of domain (Linguistic > Musical) and the red star an interaction between structure and domain.

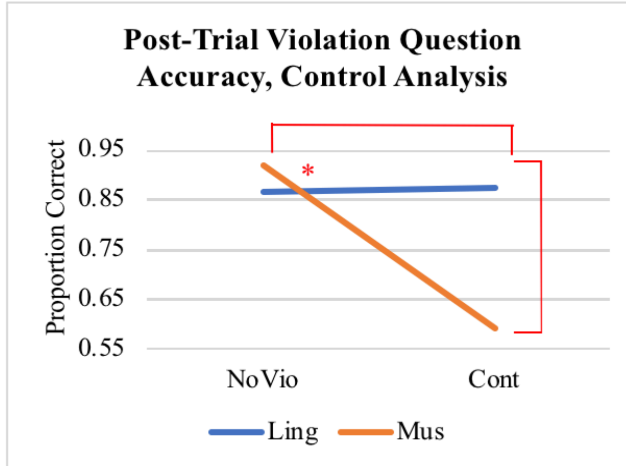


Figure 10. Mean proportions correct for the post-trial violation question for the Linguistic and Musical No Violation and Control conditions. The red brackets denote main effects of domain (Linguistic > Musical) and structure (No Violation > Control) and the red star an interaction between structure and domain.

The Question RT Violation analysis revealed main effects of domain [F (1, 18) = 18.1, p < .001, η^2 = .501 with post-hoc tests revealing a mean difference of .288 (Linguistic > Musical), t = 4.123, p_{bonf} < .001] and structure [F (1, 18) = 13.04, p = .002, η^2 = .42 with post-hoc tests revealing a mean difference of .202 (No Violation > Control), t = 3.128, p_{bonf} = .003] along with a statistically significant interaction between structure and domain [F (1, 18) = 16.52, p < .001, η^2 = .479] (Figure 11). Follow-up t-tests revealed a statistically significant difference between the Linguistic No Violation (m = 1.593, SD = .433) and Violation (m = 1.174, SD = .358) conditions [t (18) = 5.718, p < .001]. The corresponding Control analysis for Question RT, however, only revealed a main effect of domain [F (1, 18) = 20.518, p < .001, η^2 = .533 with post-hoc tests revealing a mean difference of .458 (Linguistic > Musical), t = 5.305, p_{bonf} < .001]. (Figure 12).

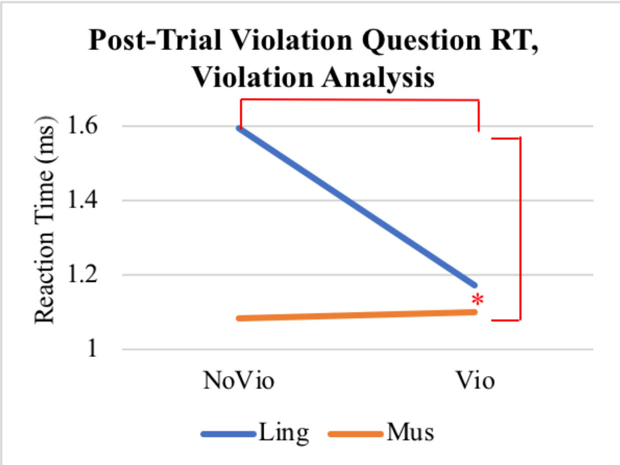


Figure 11. Mean reaction times for the post-trial structural evaluation question for the Linguistic and Musical No Violation and Violation conditions. The red brackets denote main effects of domain (Linguistic > Musical) and structure (No Violation > Violation) and the red star an interaction between structure and domain.

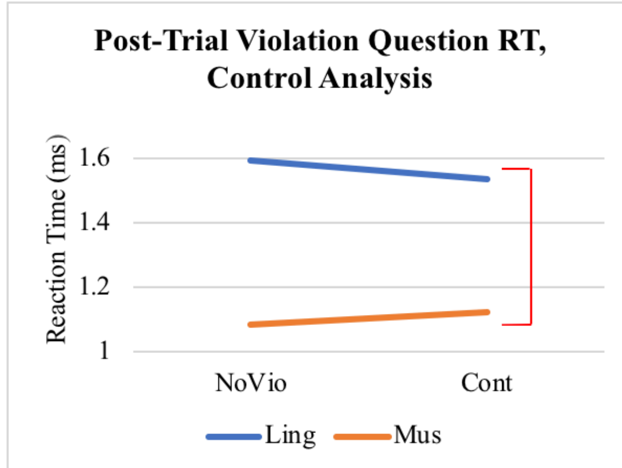


Figure 12. Mean reaction times for the post-trial structural evaluation question for the Linguistic and Musical No Violation and Control conditions. The red bracket denotes a main effect of domain (Linguistic > Musical).

Analysis of latent variables

In many experiments employing a forced two-choice paradigm with direct measurement of accuracy and RT DVs (such as the present study), it is informative to also consider the latent, unobserved variables underlying participant performance. This is because accuracy and RT often tradeoff and therefore examining each in isolation may obscure how the relationship between them can explain a subject's true processing efficiency across independent conditions. One popular method for addressing this issue is the *drift diffusion model* (DDM) (e.g., Ratcliff & McKoon, 2008). In use for decades, DDM is a mathematical-cognitive model which takes as inputs "accuracy, mean response times, and response time distributions", outputs latent parameters encompassing "components of cognitive processing" such as the ability to process information efficiently (the drift rate), an individual's response conservativeness, and nondecision time, therefore allowing for more "detailed explanations of behavior in two-choice discrimination tasks" (Ratcliff & McKoon, 2008)[14].

The DDM was applied to this data set via a simplified implementation known as the "EZ2" method, using software provided by Grasman, Wagenmakers, and van der Maas (2009). A pre-defined model was selected which used several observed parameters for each subject (proportion of errors, the mean RT of correct responses, and the variance in RT of correct responses) and output parameters including the drift rate, boundary separation (preference for one response over another), response conservativeness, and non-response time. Only the drift rate, as the feature most representative of overall processing efficiency, was used for this analysis, and participant performance was only analyzed for 2-back portion of this experiment using this method. The latent parameters output by this model were used to predict the input, observed variables, with all model fits performing well at re-creating the original data. No other models were used for comparison.

One limitation of DDM, including this EZ2 implementation, is that it does not handle perfect performance well. So, in cases where a participant did not make any mistakes (or if they had more than 2 missed button presses) within a condition, an accurate drift rate could not be obtained and instead these missing data points were imputed using the mean of all other participants in that condition. This affected 34/114 data points across all participants and conditions. The 4 participants excluded from the prior analysis were also excluded from this analysis.

[14] Thanks to Micah Johnson and Michael Spezio for their help in thinking through how to implement a DDM-approach for this experiment.

Two additional 2 x 2 [Domain (Linguistic vs. Musical) x Structure (No Violation vs. Violation or Control)] repeated measure ANOVAs were performed, this time using drift rate as the DV, again changing the second level of the structure factor from Violation to Control in the second analysis. Neither ANOVA resulted in any statistically significant main effects, however the Control analysis did yield a marginally significant interaction [$F(1, 18) = 4.278$, $p = .053$, $\eta^2 = .192$]. Figures 13 and 14 show the mean drift rate by domain and structure. Greater positive drift rates indicate increased efficiency in processing and in responding correctly to the 2-back stimuli. Note that for each of these values, the standard deviation is greater than the absolute value of the mean.

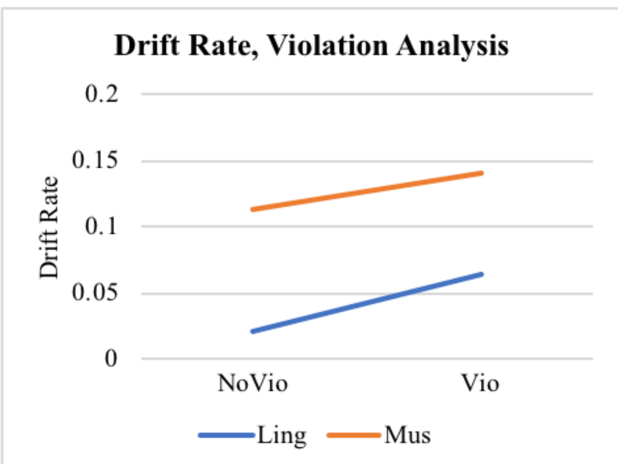


Figure 13. Mean drift rates for the 2-back data for the Linguistic and Musical No Violation and Violation conditions.

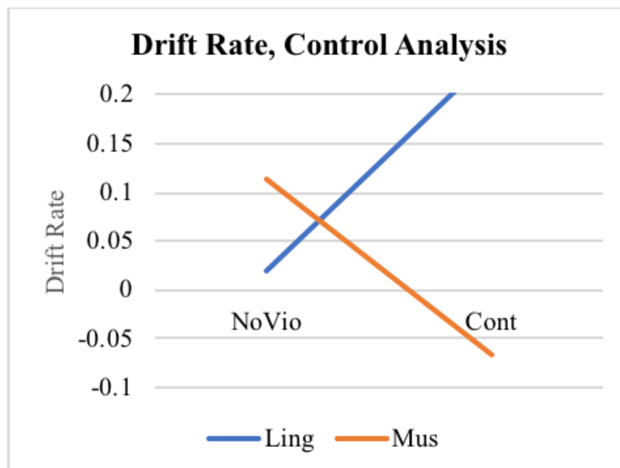


Figure 14. Mean drift rates for the 2-back data for the Linguistic and Musical No Violation and Control conditions.

Serial position effects

A one-way repeated measures ANOVA with four levels (the position of critical 2-back stimulus at either the 4th, 5th, 6th, or 7th serial position of the sequence) was conducted with accuracy of participants' 2-back response as a DV. There was a statistically significant effect of serial position [$F(3, 54) = 3.481$, $p = .022$, $\eta^2 = .162$], with none of the follow-up Bonferroni p-tests reaching a threshold of statistical significance. However, a series of paired sample t-tests indicated there were statistically significant differences in accuracy between serial positions 4 (m = .724, SD = .197) and 7 (m = .57, SD = .193) [t (18) = 2.37, p = .029]., as well as 5 (m = .689, SD = .217) and 7 [t (18) = 2.299, p = .034], with position 6 (m = .627, SD = .223) not differing from any of the others. This overall trend reflects increased difficulty in the 2-back task as the sequences progressed.

Analysis by group

All of the previously reported ANOVAs were repeated adding in participant classification (musician or non-musician) as a between-subject factor. In only one case, the Violation analysis of the Question Accuracy DV, was there a statistically significant between-subject effect [F (1, 17) = 8.961, p = .008, η^2 = .345]. Follow-up t-tests revealed this was not driven by differences in the Musical conditions, however, but by the Linguistic Violation condition [t (17) = 2.482, p = .024] in which non-musicians (m = .844, SD = .111) had more difficulty answering correctly than musicians (m = .955, SD = .084).

Discussion

Recent work has suggested that overlaps in processing between music and language may be due to shared reliance on domain-general neural resources such as working memory or cognitive control (e.g. Rogalsky et al., 2011; Slevc, Reitman, & Okada, 2013). Conversely, other research teams continue to claim that Patel's SSIRH is parsimonious and cross-domain usage of syntax specific resources, not WM capacities, are responsible for such observations (e.g. Kunert, Willem & Hagoort, 2016). Experiment 2 proposed to help resolve this tension by investigating the degree to which the processing of both musical and linguistic phrases, with and without structural violations of listener expectation, affected participants' ability to respond to a simultaneous test of WM. The post-trial evaluations of the linguistic and musical stimuli were also examined.

Effects on WM task performance

The primary analyses of accuracy and reaction time data from the numerical 2-back task repeatedly yielded statistically significant main effects of Musical and Linguistic structure on accuracy in both the Violation and Control analyses. That is, compared to the conditions without any violation or salient event (i.e. a harmonically congruous chord sequence or grammatically correct sentence), those trials which featured either a structural violation (i.e. an out-of-key chord or an ungrammatical word) or another salient event (i.e. a chord in a different timbre or a word spoken in a different voice) led to more errors responding to the 2-back number. However, without an interaction between structure and domain, and given that the same pattern was observed in the structural Violation and in the Control conditions, this can be attributable most likely to these unexpected auditory events distracting participants attention.

The ANOVAs of the RTs for correctly-answered 2-back trials revealed no statistically significant effects. Inspection of Figures 7 and 8 shows a clear trend of increased interference (i.e. increased RT) when comparing musical structural violations or timbre changes to those musical sequences without any unexpected events. It could be that the low number of trials per condition (8) in this experiment made it more difficult to uncover statistically significant patterns here, though, again, given that the directional trend is the same in both the violation and control analyses, there can be yet no conclusions drawn with respect to RT.

An initial analysis of drift rate attempted to take both accuracy and RT data into account simultaneously to explore participants overall processing efficiency during the 2-back task. This approach was similarly unfruitful, and its usefulness may also have been hampered by low numbers of trials per condition.

Overall, there is not sufficient evidence to support the primary hypotheses of this experiment, namely that the presence of a structural violation, but not a salient change in timbre, in the musical sequences would lead to increased processing difficulties for a WM task and that these differences would be comparatively greater than those in the language conditions. Instead, there is only an indication from these data that the intrusion of any unexpected auditory event interferes with our capacity to correctly complete a simultaneous test of WM.

Effects of structural evaluation

Although not the primary interest of this experiment, data was also collected and analyzed for participants' post-trial structural evaluation (i.e. the violation question response) of the auditory material they were presented with. Their 'accuracy' on this task (that is, their ability to respond whether or not there was a theoretical harmonic violation in the music or a grammatical error in the language) showed statistically significant interactions between structure and domain, driven by drastically worse performance (reduction by nearly 1/3) when evaluating musical sequences containing both out-of-key violations (a sharp IV chord) and chords played with a different timbre (orchestral strings as opposed to the usual presentation on piano). This indicates participants tended to be equally confounded both by the presence of structural violations in music (what are often, likely erroneously, referred to as syntactic violations) and sudden changes in timbre. It is worth noting that, by contrast, in all linguistic conditions as well as in the non-violation musical condition participants consistently performed well (above 85% response accuracy).

The RT data for the post-trial question are more difficult to interpret because participants had up to 5 seconds to respond and therefore were under less time pressure. However, Figures 11 and 12 make it clear that responses to the music sequences were faster (< 1.2 seconds for all conditions) whereas participants took more time to evaluate the linguistic material (> 1.5 seconds). The notable exception was the linguistic violation condition, where participants' performance was comparable to the music conditions in deciding quickly that there was a grammatical error, which drove an interaction in this analysis.

Although there were no specific prior hypotheses for this part of the analysis, these patterns of results lead to some interesting considerations. As was the case in Experiment 1, participants were much better at classifying harmonically congruent sequences as being free of error than they were at identifying theoretical harmonic violations as being incongruent within the context of a key. Additionally, responses were divided as to whether the intrusion of a new timbre constituted a violation (59.2% responding that it did not), indicating the instructions for this experiment were ambiguous in this respect.

Conclusions

These results from the post-trial structural evaluations constitute additional evidence, in line with Experiment 1, that the theoretical bases of the HEVP used in much of the literature concerning the relationship between music and language should be more critically questioned. Even choosing a harmonic violation that Experiment 1 demonstrated was generally interpretable as such by musicians and non-musicians alike, the addition of a simultaneous WM task led to poor classification performance in this experiment (69.1% as opposed to 88.33% in Experiment 1). Future work could more systematically investigate the relationship between musical congruency judgments with and without the presence of a secondary task.

The primary analyses of the effect of violations on 2-back task performance did not yield definitive results. It would likely be incorrect, however, to draw from this the conclusion that there is no relationship between musical sequence processing and domain-general WM resources. As mentioned, the current experiment was limited by a low number of trials per conditions (due to timing and budget constraints) as well as missing data points due to the difficulty of the dual task paradigm, even after training. Instead, further similar research should be done so that the field can better appreciate the specific differences and similarities between how music and language engage WM processes.

And there is certainly still ample experimental work to be done. Debates over the relationship between syntax processing, Broca's area, executive resources, and the cognitive domains of music and language will surely continue for decades to come. Although the current results from Experiments 1 and 2 will do little to resolve these questions, the following lessons can be learned from this literature review and these studies: 1) The reality of musical experience in the brain and in the world is far more varied than is often acknowledged in laboratory settings. An individual's culture and background playing music clearly influences how they interpret music and this should have consequences for any investigation, whether it be neuroscientific, linguistic, sociological, or otherwise. 2) The differences in behavioral interpretations between musicians and non-musicians of even stripped-down, Western tonality-based chord sequences, as demonstrated in these experiments, should be taken into account in future experiments investigating the neural substrates of different cognitive domains. Brain functional, anatomical, and activation data should be interpreted in conjunction with behavioral data. 3) These observations are yet more ways in which music belies simple analogy with language (see **Appendix A** for another example). Nonetheless, it is still a fruitful scientific and humanistic enterprise to understand why and to what extent music, language, and other domains of cognition are similar or even co-dependent. The answers to these questions may help us treat neurological disorders, understand the evolution of our species, and teach children in innovative ways. Consequently, we should fully support further efforts to carefully understand how our evolving notions of syntactic processing, facets of natural language, and the full range of human musical experiences interrelate.

***

Appendix A: Language, Music, and Recursion

It may seem strange that in such a long literature review about the neural computations enabling language and music processing that recursion has not been discussed yet, except for mention of Fitch and Martins' idea that a recursive "scannable stack buffer" structure is key to the processing of complex hierarchical information across domains (Fitch & Martins, 2014). This is far from the only hypothesis along these lines, however. Before exploring a variety of perspectives on the potentially recursive nature of music in the mind, though, an obviously key question is: what exactly is recursion?

It is a difficult question, since it is a term used in a wide variety of fields, from mathematics and computer science to linguistics and philosophy. In the context of brain sciences, two helpful general definitions are "(1) [the] embeddedness of phrases within other phrases, which entails keeping track of long-distance dependencies among phrases and (2) the specification of the computed output string itself, including meta-recursion, where recursion is both the recipe for an utterance and the overarching process that creates and executes the recipes" (Coolidge, 2011). Although there is a fair amount of theoretical confusion in the literature due to many authors failing to carefully define their terms, most implicitly use the first definition (e.g. the embedding of a noun phrase within a noun phrase which changes the meaning of the entire phrase).

In Michael C. Corballis' 2011 book *The Recursive Mind* – in which he defines recursion as "a procedure that calls itself, or…a constituent that contains a constituent of the same kind" – he lays out an evolutionary account in which our species developed a "general cognitive capacity...to share our thoughts and emotions" via recursion. Corballis postulates that not only language but also theory of mind and mental time travel are dependent on a form of embedded recursion (e.g. thinking about others' thoughts or remembering past remembrances) and that this "is the primary characteristic that distinguishes the human mind from that of other animals" (Corballis, 2011). Corballis' theory takes for granted that the debate surrounding recursion's place in language comprehension and production is settled, though, and this is far from true.

Perhaps the most influential recent paper on this subject is "The Faculty of Language: Who Has It, What Is It, and How Did It Evolve" from Marc Hauser, Noam Chomsky, and Tecumseh Fitch, published in *Science* in 2002. In this article, the authors distinguish between the faculty of language in a "broad sense" and a "narrow sense" (FLN) wherein FLN is a recursive capability unique to human and responsible for the hierarchical, generative, and limitless nature of human language (Hauser, Chomsky, & Fitch, 2002). FLN is the most specific formulation of Chomsky's Minimalist Program which has had as a tenet for decades that a recursive Merge operation allows for the integration of syntactic objects into a meaningful whole (Chomsky, 1995). The recursion-based nativism of FLN, as an explanation for the supposed "innate dispositions" underlying "universal grammar" (Hauser, Chomsky, & Fitch, 2002), was an explosive and controversial proposition in the field, prompting some researchers to adopt the idea that "recursion is accomplished by the hierarchy" and is separable from a "general learning mechanism" (Perlovsky, 2009). However, FLN has been met with substantial pushback as well.

The most popular and direct response to Hauser, Chomsky, and Fitch came 3 years later from Pinker and Jackendoff who claimed that FLN is too focused on syntax and ignores a variety of other language adaptations, specifically in phonology, that are also "uniquely human, discretely infinite, and not recursive" (Pinker & Jackendoff, 2005). Other researchers have gone further and said that the lack of a rigorous definition of recursion in FLN is problematic and, in a mathematical sense, it would be better to use the term "inductive definition" (Tomalin, 2007).

The trouble with using the term recursion in linguistic analysis stems back to empirical observations that "naturally occurring speech contains only a limited amount of complex recursive structure" (Christiansen & Chater, 1999). So, from a usage and comprehension perspective, even if Chomskyan linguistics can identify recursive structures in certain phrases, this is not terribly informative in helping understand how language is produced and processed in the brain (Michael Diercks, personal communication, 10/31/2017). One recent critique of the adoption of the term recursion by scientists studying language is even more forceful: it has been demonstrated that, computationally, supposedly recursive and hierarchical structures, such as center-embedded sentences, can be parsed using the rules of non-recursive, less complex context-free grammars (Paap & Partridge, 2014). Evidence such as this casts into doubt the notion that any innate recursive mechanism exists for the language domain, let alone for other cognitive capacities.

Nonetheless, that has not stopped cognitive neuroscientists, especially those steeped in the Chomskyan tradition, from applying their understanding of our species' recursive capabilities towards their experimental designs. For example, one paradigm presents subjects with "simple" and "embedded" artificial grammar phrases of alternating A's and B's in an effort to localize hierarchical grammatical processing, for instance "in Broca's area and the adjacent rim of the ventral premotor cortex" (Bahlmann, Schubotz, & Friederici, 2008). Even a recursion optimist such as Corballis has criticized this type of task as being too simplistic and possible to parse via non-recursive strategies (Corballis, 2007). Other studies, however, have taken more care to distinguish between the structure of stimuli and the cognitive processes involved in understanding them. In 2012, Martins and Fitch found, for instance, that participants can implicitly distinguish between recursively and iteratively-created fractal structures in the visuospatial domain and that this ability is correlated positively with verbal working memory skills. There is still a long way to go, though, in determining to what extent this is true for grammatical or musical structures.

At the least, it is difficult to say with any certainty that recursion is a fundamental property of human language. This is the main reason it was not fully discussed as a concept in my main literature review. On top of this issue, though, what is there to make of the role of recursion in musical processing? The answer is, so far, very little. There appears to be only one study looking at the connection between recursion and music[15], from the same team behind the visual fractal experiment described briefly above. Using a similar design, they present evidence that humans can represent recursion in the auditory domain and that this ability is correlated with other recursive abilities, however with the caveat that the capacity "to build recursive algorithms that generate hierarchical structures... [does not mean that] humans actually represent the recursive character of these structures... [or] use these representations productively" (Martins et al., 2017).

So, where does this leave us? While we should not entirely rule out the possibility that a computationally recursive procedure is important to one or more domains of cognition, there is not yet enough evidence to make it the focus of this thesis. Although work and theories from the likes of Corballis, Fitch, and Martins is intriguing, we must keep in mind Martins' assessment that, "given that brain computations are opaque to observers (until behavioral correlates have been found), definitions focused on algorithmic properties (such as 'a recursive function is one that calls itself') may not be entirely relevant for empirical research" (Martins, 2012). It is still

[15] There is, additionally, some unpublished work from Jonah Katz and David Pesetsky from MIT relating what they term "The Identify Thesis for Language and Music". In an attempt to update Jackendoff and Lerdahl's GSM of music with modern linguistics, they claim that music processing too employs a recursive Merge operation.

worthwhile to consider theories such as the “scannable stack” in Broca’s area and to continue investigating the neural correlates of information processing in language, music, and other domains, however it will likely still be a long time before a mathematically rigorous model, based on recursion or otherwise, becomes apparent and useful in this field.

Appendix B: Further Implications from Experiment 1

The practice of bridging musicology with more empirical fields is an example of "consilience" or "the synthesis of all branches of human knowledge through the adoption of the scientific method" (Wilson, 1998). Two modern practitioners of such consilience are AD Patel and Fred Lerdahl. Patel is a giant in the study of music-language relations in the brain who has written that, especially in this field, "interactions across traditional boundaries can bear fruit in the form of new ideas and discoveries that neither side can accomplish alone" (Patel, 2009)[16]. Lerdahl, for his part, is best known as one half of the team, along with Ray Jackendoff, behind 1985's *A Generative Theory of Tonal Music* (GTTM), an attempt to bring together Chomskyan linguistics and music theory, especially via the introduction of tree structures to musical analysis (Lerdahl & Jackendoff, 1985). So, continuing in this collaborative tradition, what might Patel's and Lerdahl's work have to say about the findings of Experiment 1?

In his opus *Music, Language, and the Brain*, in a chapter dedicated to his widely influential *shared syntactic integration resource hypothesis* (SSIRH), Patel repeatedly stresses the importance of continuing behavioral work which can directly test the predictions of the SSIRH (particularly regarding "the interaction[s] of musical and linguistic syntactic processing") (Patel, 2009). However, most, if not all, of the existing research either intentionally only recruits non-musicians or does not note the musical backgrounds of its participants. Further, very few modern studies record subjects' interpretations of deviant or incongruent chords, instead usually relying on indirect measures such as neural responses (e.g. Koelsch et al., 2013) or performance

---

[16] Of course, no one researcher can entirely bridge the gaps between fields. A reasonable, recurrent criticism of Patel and other scholars interested in the processing of musical structures is that they are too steeped in Western classical music. In the future, theories should consistently be geared towards understanding the breadth of human musical traditions (see Lawson, 2012 for a good attempt at just this). And there is no shortage of good ethnomusicological work out there just waiting to be integrated into our understanding of how the mind and music interact (e.g. Harvard's "Natural History of Song" project; https://www.naturalhistoryofsong.org/).

on another task (e.g. Fedorenko et al., 2014). This should be considered a gap in the literature. It is difficult to establish a convincing account of 'syntactic' music processing without a robust understanding of how musical sequences are explicitly understood by individuals. Obviously, it is for good reason that scientists often focus on the mechanisms of the brain: there are various aspects of how our mind works that we are not consciously aware of or able to articulate. Nonetheless, linguists, language neuroscientists, and laypeople alike would likely agree that being able to mutually understand sentences (without undergoing any specialized training) is key to both routine communication and knowing how this biological communicative system works – music research should be held to a similar standard.

As mentioned, it would be unfair to claim that there has been no substantial effort to consider the effects of musical background on musical capacities. As will be seen, however, these efforts tend to be older and largely divorced from the current debate about language, music, and syntax. Perhaps as a result, although the papers about to be discussed are oft-cited, there is no in-depth discussion of these types of findings in much of the recent work on music and language processing.

In a 2006 paper, Bigand and Poulin-Charronnat review a host of experiments from the 1980's, 90's, and early 2000's which investigated differences between musically trained and non-trained people on a variety of musical abilities including perceiving tension, learning new musical systems or "idioms", and relating musical variations to a central theme. Importantly for this discussion, they claim the data have converged in one important respect: there are apparently no emergent group differences when it comes to priming harmonic expectations (Bigand & Poulin-Charronnat, 2006). Doesn't this conclusion contradict the findings of Experiment 1? There are at least two reasons this might not be the case. First, the Experiment 1 was not a traditional priming study, instead it explicitly asked for congruency judgments, an aspect missing from previous investigations. Second, modern physiological and neuroscientific techniques may be more sensitive in detecting group differences.

Consider a 2006 paper from Steinbeis, Koelsch, and Sloboda, which investigated musician and non-musician “emotional responses” to chords within modified Bach chorales, which were either theoretically expected or unexpected. Although they found similar heart rate and electrodermal activity responses from both groups across conditions (with slightly faster onset in musicians), they additionally found significantly larger event related potentials (P3 components) in musicians following the unexpected chords, an effect they interpret to be reflective of training (Steinbeis, Koelsch, & Sloboda, 2006). Another study from Koelsch et al. (2005) used functional magnetic resonance imaging to reveal, following “irregular chords”, stronger activation “in the frontal operculum and the anterior portion of the superior portion of the superior temporal gyrus” as a function of musical expertise in both adults and children. These results, while correlational, show that there is already some neuropsychological and neuroscientific evidence supporting the findings of Experiment 1.

Any theory of how the brain works must at least attempt to distinguish between how we are fundamentally wired and what we learn. It is far less interesting and informative, for example, to say that American college-aged people who practiced music as children notice when a chord is harmonically unexpected than the much stronger claims of Patel’s SSIRH. But this is not the only important theory of music and the brain which might be impacted by this recognition.

In examining the relationship between harmonic expectations and musicianship, we must also consider work from Lerdahl and others regarding the role of musical tension. In a 1996 study from Bigand, Parncutt, and Lerdahl, it was reported that the perceived musical tension introduced by a chord varied due to “tonal hierarchies, sensory chordal consonance, horizontal motion, *and musical training*” (italics added). This was largely inspired by Lerdahl’s work on tonal pitch space theory (TPST) which is a component of a theory of tonal tension. In basic terms, in TPST “tonal hierarchies are represented in the form of a multidimensional space in which the distances of chords from the instantiated tonic correspond their relative hierarchical importance” (Bigand, Parncutt, & Lerdahl, 1996)[17]. TPST has more recently been combined with GTTM-style[18] prolongational structure analysis (the tree representations “of the hierarchical event structure in a musical passage”) in an effort to understand the ebbing and flowing of tension throughout music (Lerdahl & Krumhansl, 2007). This work, while interested in what mode of analysis (e.g. sequential vs. hierarchical) listeners use in evaluating tension in music, appears to not be interested in differences among listeners, instead focusing only on global means. But what if, as was indicated in 1996, musical training does matter? What if the ability to track tension and harmony throughout a musical ‘tree’ is dependent on specialized knowledge – would the ‘branches’ bend, or even break, in this case?

Lerdahl and his colleagues, at least, seem to be comfortable with this, pardon the pun, tension. Indeed, he and Krumhansl’s analysis and discussion recognizes that the algorithms they

---

[17] TPST states that standard triadic chords introduce less tension than seventh chords, however V^7 chords particularly are supposed to be low-tension events (e.g. a G7 chord is assigned a distance of 6 within C major, the dominant G is given a 5, and a theoretically incongruent chord such as the Neapolitan sixth (C#) has a distance of 16). This is interesting considering that, in Experiment 1, both musician and non-musician participants were averse to rating sequences ending on V^7 chords as harmonically congruent. Therefore, there should be further work probing this specific prediction of TPST.

[18] Another issue with GTTM is that, for all its popularity in the field, it is still not nearly as well studied or understood as linguistic tree structures are – perhaps for good reason, there have yet to be PhD programs specializing in musical syntax. Even for those scientists and scholars with a solid music background, it is dense and difficult material to work through and requires a fair bit of expertise in linguistics as well. This is admittedly one of the chief issues with consilience, that it is rare for one person to have all the necessary vocabulary and knowledge to work between two or more fields.

use to calculate tonal tension are imperfect, just as the process for determining tree structures is somewhat dependent on preference and context (i.e. the harmonic environment, the rhythmic complexity, etc.). The results from Experiment 1, if true, would indicate, nonetheless, that further work in this vein should pay more explicit attention to the influence of musical expertise. This was more or less the norm in the 20th century research, and it should be re-emphasized going forward.

To wrap up, it is worth addressing two strands of criticism. The first claims that our analysis of music and language similarities need not be dependent on any theory of music or language neural processing or even behavioral responses. It may be, then, that comparisons between music and language as just a useful structural "analogy" which can act as a starting point for interesting analyses of these communicative systems (Swain, 1997). The second scoffs at the entire enterprise of GTTM and TPST – didn't traditional tonal harmony die in the 20th century with the advent of 12-tone composition, free form jazz, etc.? And what of the fact that composers as far back as, and surely before, Beethoven, employed theoretical harmonic tonal 'violations' in their music? How does this scientific theorizing help us understand how music really works? Though these viewpoints are divergent, both merit a similar response. It is worthwhile to explore both the promise and the potential pitfalls in Patel's and Lerdahl's theories. This is because we know from a plethora of scientific experiments, some presented above, that we are dealing with real cognitive phenomena. What is left to find out is to what extent the parallels between such disparate, similar systems as music and language are informative when it comes to understanding the mind and behavior. Experiment 1 is a good starting point for delving anew into these considerations. Though, admittedly, it is just that: a starting point.

Appendix C: Stimuli from Experiment 2

Critical Position	Language with Violation (Critical Trials)	Number pairings
7	The client's personal publicist will be speak.	[4, 0, 3, 6, 5, 6, 5]
6	The child will have been plays alone.	[0, 7, 5, 2, 7, 2, 3]
5	The tourist will have seeing lots yesterday.	[1, 7, 3, 2, 3, 7, 0]
4	The parole officer obey the judge's rule.	[6, 8, 0, 8, 9, 6, 8]
7	The adventure will surely have been starts.	[7, 1, 9, 8, 2, 7, 2]
6	The orange power cord had tightening quickly.	[1, 9, 0, 8, 6, 8, 4]
5	The adventurous explorer is travels a lot.	[3, 6, 5, 8, 5, 0, 3]
4	Yesterday Bill was cleans in the yard.	[9, 3, 8, 3, 4, 1, 5]
	Language without Violation (Critical Trials)	
7	The guilty crook still fought the verdict.	[8, 3, 3, 1, 0, 8, 0]
6	The traveler has packed for her journey.	[6, 0, 9, 8, 3, 8, 5]
5	The specialist will be designing our furniture.	[9, 4, 5, 2, 5, 1, 9]
4	The old machine buzzed and beeped loudly.	[0, 6, 7, 6, 4, 7, 5]
7	The cloud was darkening over my head.	[7, 4, 9, 3, 2, 2, 2]
6	The blue bird will sing all morning.	[8, 9, 7, 5, 9, 5, 4]
5	The idle waitress should have been working.	[7, 5, 0, 3, 0, 1, 8]
4	The happy amateur works next to me.	[8, 4, 5, 4, 1, 9, 4]
	Language with Violation (Catch Trials)	
N/A	The pastor will works for you still.	[2, 5, 7, 2, 9, 8, 4]
N/A	The voyage was be ending by tomorrow.	[6, 2, 1, 4, 2, 6, 7]
N/A	The building will be collapse soon enough.	[0, 1, 8, 4, 8, 4, 7]
N/A	The glass will have been smashing completely.	[1, 9, 4, 6, 0, 3, 1]
	Language without Violation (Catch Trials)	
N/A	The critic had not reviewed many plays.	[9, 1, 8, 1, 8, 6, 2]
N/A	The brightness will have faded by night.	[1, 8, 4, 6, 8, 0, 2]
N/A	The plant will have wilted by then.	[8, 8, 4, 1, 6, 7, 0]
N/A	The baby slept but could be woken.	[7, 1, 1, 8, 3, 2, 7]
	Language Control Trials (bolded words in female voice)	
7	I heard the paper that was **ripping**.	[5, 3, 7, 2, 1, 4, 1]
6	Be safe or the car **will** crash.	[5, 6, 3, 2, 7, 2, 8]
5	The boy goes to **bed** on time.	[3, 2, 1, 6, 1, 8, 9]
4	One day I **will** leave this place.	[4, 3, 6, 3, 8, 7, 2]
7	The friends went to the park **today**.	[5, 6, 8, 2, 3, 8, 3]
6	Dad called and said hello **to** me.	[4, 7, 3, 8, 2, 8, 1]
5	The eager student will **have** been learning.	[4, 3, 6, 5, 6, 8, 4]
4	The evil dictator **is** still scheming constantly.	[5, 7, 1, 7, 4, 8, 2]

Critical Position	Music with Violation (Critical Trials)	Number pairings
7	D# A# Cm G# D# A# A	[6, 0, 4, 7, 7, 4, 7]
6	G D Em C G C# G	[7, 4, 4, 8, 8, 8, 3]
5	G# D# Fm C# D D# G#	[8, 8, 9, 7, 9, 6, 2]
4	A E F#m D# A E A	[4, 4, 3, 4, 1, 3, 0]
7	B F# G#m E F# B F	[2, 4, 8, 1, 6, 5, 6]
6	C G Am F G F# C	[0, 4, 8, 9, 4, 9, 3]
5	F C Dm A# B C F	[2, 6, 3, 0, 3, 4, 1]
4	F# C# D#m C B C# F#	[9, 4, 8, 4, 9, 5, 7]
	Music without Violation (Critical Trials)	
7	D A Bm G D A D	[1, 8, 9, 6, 2, 7, 2]
6	D# A# Cm G# D# A# D#	[1, 2, 6, 4, 2, 4, 7]
5	E B C#m A E B E	[8, 3, 3, 0, 3, 4, 2]
4	F C Dm A# F C F	[6, 7, 1, 7, 8, 2, 0]
7	A# F Gm D# A# F A#	[8, 4, 3, 7, 5, 8, 5]
6	D A Bm G D A D	[0, 5, 7, 1, 3, 1, 6]
5	B F# G#m E B F# B	[5, 7, 0, 6, 0, 5, 5]
4	C G Am F C G C	[8, 5, 2, 5, 6, 2, 3]
	Music with Violation (Catch Trials)	
N/A	G# D# Fm D G# D# G#	[7, 8, 1, 5, 8, 6, 3]
N/A	A# F E D# A# F A#	[9, 1, 4, 8, 5, 4, 7]
N/A	B F# G#m E B F# F	[4, 4, 4, 1, 5, 9, 5]
N/A	C# G# G F# C# G# C#	[8, 1, 0, 9, 4, 6, 8]
	Music without Violation (Catch Trials)	
N/A	G D Em C G D G	[1, 9, 2, 9, 5, 4, 5]
N/A	C G Am F C G C	[0, 3, 9, 6, 1, 3, 2]
N/A	C# G# A#m F# C# G# C#	[2, 5, 9, 6, 6, 2, 4]
N/A	E B C#m A E B E	[9, 8, 5, 4, 6, 6, 8]
	Music Control Trials (bolded chords in string timbre)	
7	F# C# D#m B F# C# **F#**	[5, 2, 3, 1, 4, 9, 4]
6	G D Em C G **C** G	[5, 6, 3, 4, 8, 4, 5]
5	G# D# Fm C# **G#** D# G#	[3, 4, 1, 5, 1, 6, 9]
4	A E F#m **E** A E A	[1, 9, 3, 9, 5, 6, 3]
7	C# G# A#m F# C# G# **C#**	[5, 2, 7, 1, 0, 9, 0]
6	D# A# Cm G# D# **A#** D#	[6, 9, 1, 1, 5, 1, 0]
5	F C Dm A# **F** C F	[0, 6, 7, 4, 7, 2, 8]
4	F# C# D#m **F#** B C# F#	[8, 5, 3, 5, 1, 2, 3]

References

Amunts, K., & Zilles, K. (2012). Architecture and organizational principles of Broca's region. *Trends in Cognitive Sciences*, *16* (8), 418–426. doi:10.1016/j.tics.2012.06.005

Amunts, K., Schleicher, A., Bürgel, U., Mohlberg, H., Uylings, H., & Zilles, K. (1999). Broca's region revisited: Cytoarchitecture and intersubject variability. *Journal of Comparative Neurology*, *412* (2), 319–341.

Anvari, S., Trainor, L., Woodside, J., & Levy, B. (2002). Relations among musical skills, phonological processing, and early reading ability in preschool children. *Journal of Experimental Child Psychology*, *83* (2), 111–130. doi:10.1016/S0022-0965(02)00124-8

Baddeley, A. (2000). The episodic buffer: a new component of working memory? *Trends in Cognitive Sciences*, *4* (11), 417–423. doi:10.1016/S1364-6613(00)01538-2

Baddeley, A.D., & Hitch, G. (1974). Working memory. *Psychology of learning and motivation, 8*, 47-89. doi:10.1016/S0079-7421(08)60452-1

Badre, D., & Frank, M. (2012). Mechanisms of Hierarchical Reinforcement Learning in Cortico–Striatal Circuits 2: Evidence from fMRI. *Cerebral Cortex*, *22* (3), 527–536. doi:10.1093/cercor/bhr117

Bahlmann, J., Schubotz, R., & Friederici, A.D. (2008). Hierarchical artificial grammar processing engages Broca's area. *NeuroImage*, *42* (2), 525–534. doi:10.1016/j.neuroimage.2008.04.249

Besson, M., & Schön, D. (2001). Comparison between Language and Music. *Annals of the New York Academy of Sciences*, *930* (1), 232–258. doi:10.1111/j.1749-6632.2001.tb05736.x

Bigand, E., & Pineau, M. (1997). Global context effects on musical expectancy. *Perception, & Psychophysics, 59* (7), 1098-1107.

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*(1), 100-130. doi:10.1016/j.cognition.2005.11.007

Bigand, E., Tillmann, B., & Poulin-Charronnat, B. (2006). A module for syntactic processing in music?*Trends in Cognitive Sciences*, *10* (5), 195–196. doi:10.1016/j.tics.2006.03.008

Cameron-Faulkner, T., Lieven, E., & Tomasello, M. (2003). A construction based analysis of child directed speech. *Cognitive Science*, *27* (6), 843–873. doi:10.1016/j.cogsci.2003.06.001

Cheung, V. K., Meyer, L., Friederici, A. D., & Koelsch, S. (2018). The right inferior frontal gyrus processes nested non-local dependencies in music. *Scientific Reports,8*(1). doi:10.1038/s41598-018-22144-9

Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge, MA: The MIT Press.

Chomsky, N. (1995). *The minimalist program*. Cambridge, MA: The MIT Press.

Christiansen, M., & Chater, N. (1999). Toward a connectionist model of recursion in human linguistic performance. *Cognitive Science*, *23* (2), 157–205. doi:10.1016/S0364-0213(99)00003-8

Clerget, E., Badets, A., Duqué, J., & Olivier, E. (2011). Role of Broca's area in motor sequence programming: a cTBS study. *NeuroReport*, *22* (18), 965. doi:10.1097/WNR.0b013e32834d87cd

Coolidge, F., Overmann, K., & Wynn, T. (2011). Recursion: what is it, who has it, and how did it evolve?*Wiley Interdisciplinary Reviews: Cognitive Science*, *2* (5), 547–554. doi:10.1002/wcs.131

Corballis, M. (2007). On phrase structure and brain responses: a comment on Bahlmann, Gunter, and Friederici (2006). *Journal of Cognitive Neuroscience, 19* (10), 1581-1583.

Corballis, M. C. (2011). *The recursive mind: the origins of human language, thought, and civilization*. Princeton, NJ: Princeton University Press.

Deutsch, D. (1980). The processing of structured and unstructured tonal sequences. *Perception & Psychophysics,28*(5), 381-389. doi:10.3758/bf03204881

Dominey, P., Hoen, M., Blanc, J.-M., & Lelekov-Boissard, T. (2003). Neurological basis of language and sequential cognition: Evidence from simulation, aphasia, and ERP studies. *Brain and Language*, *86* (2), 207–225. doi:10.1016/S0093-934X(02)00529-1

Fadiga, L., Craighero, L., & D'Ausilio, A. (2009). Broca's Area in Language, Action, and Music. *Annals of the New York Academy of Sciences*, *1169* (1), 448–458. doi:10.1111/j.1749-6632.2009.04582.x

Fedorenko, E. (2014). The role of domain-general cognitive control in language comprehension. *Frontiers in Psychology*, *5*, 335. doi:10.3389/fpsyg.2014.00335

Fedorenko, E., & Varley, R. (2016). Language and thought are not the same thing: evidence from neuroimaging and neurological patients. *Annals of the New York Academy of Sciences*, *1369* (1), 132–153. doi:10.1111/nyas.13046

Fedorenko, E., Behr, M., & Kanwisher, N. (2011). Functional specificity for high-level linguistic processing in the human brain. *Proceedings of the National Academy of Sciences*, *108* (39), 16428–16433. doi:10.1073/pnas.1112937108

Ferreira, F., & Clifton, C. (1986). The independence of syntactic processing. *Journal of Memory and Language*, *25* (3), 348–368. doi:10.1016/0749-596X(86)90006-9

Fiebach, C. J., Schlesewsky, M., Lohmann, G., Cramon, D. Y., & Friederici, A. D. (2005). Revisiting the role of Broca's area in sentence processing: Syntactic integration versus syntactic working memory. *Human Brain Mapping*, *24* (2), 79–91. doi:10.1002/hbm.20070

Fitch, T., & Martins, M. (2014). Hierarchical processing in music, language, and action: Lashley revisited. *Annals of the New York Academy of Sciences*, *1316* (1), 87–104. doi:10.1111/nyas.12406

Forgeard, M., Winner, E., Norton, A., & Schlaug, G. (2008). Practicing a Musical Instrument in Childhood is Associated with Enhanced Verbal Ability and Nonverbal Reasoning. *PLoS ONE*, *3* (10), e3566. doi:10.1371/journal.pone.0003566

Friederici, A.D. (2004). Processing local transitions versus long-distance syntactic hierarchies. *Trends in Cognitive Science, 8* (6), 245-247.

Friederici, A.D. (2011). The brain basis of language processing: from structure to function. *Physiological review*, *91* (4), 1357–1392.

Friederici, A.D., & Rueschemeyer, S.A. (2003). The role of left inferior frontal and superior temporal cortex in sentence comprehension: localizing syntactic and semantic processes. *Cerebral Cortex, 13* (2), 170-177.

Friederici, A.D., Kotz, S., Scott, S., & Obleser, J. (2010). Disentangling syntax and intelligibility in auditory language comprehension. *Human Brain Mapping*, *31* (3), 448–457. doi:10.1002/hbm.20878

Friedrich, R., & Friederici, A.D. (2009). Mathematical Logic in the Human Brain: Syntax. *PLoS ONE*, *4* (5), e5599. doi:10.1371/journal.pone.0005599

Fuster, J. (2001). The Prefrontal Cortex—An Update Time Is of the Essence. *Neuron*, *30* (2), 319–333. doi:10.1016/S0896-6273(01)00285-9

Goucha, T., & Friederici, A.D. (2015). The language skeleton after dissecting meaning: A functional segregation within Broca's Area. *NeuroImage*, *114*, 294–302. doi:10.1016/j.neuroimage.2015.04.011

Gough, P., Nobre, A., & Devlin, J. (2005). Dissociating Linguistic Processes in the Left Inferior Frontal Cortex with Transcranial Magnetic Stimulation. *The Journal of Neuroscience*, *25* (35), 8010–8016. doi:10.1523/JNEUROSCI.2307-05.2005

Grasman, R. P., Wagenmakers, E., & Maas, H. L. (2009). On the mean and variance of response times under the diffusion model with an application to parameter estimation. *Journal of Mathematical Psychology,53*(2), 55-68. doi:10.1016/j.jmp.2009.01.006

Grewe, T., Bornkessel, I., Zysset, S., Wiese, R., Cramon, D., & Schlesewsky, M. (2005). The emergence of the unmarked: A new perspective on the language-specific function of Broca's area. *Human Brain Mapping*, *26* (3), 178–190. doi:10.1002/hbm.20154

Grodzinsky, Y., & Santi, A. (2008). The battle for Broca's region. *Trends in Cognitive Sciences*, *12*(12), 474–480. doi:10.1016/j.tics.2008.09.001

Hagoort, P., & Levelt, W. (2009). The speaking brain. *Science, 326* (5951), 372-373.

Hanna-Pladdy, B., & Mackay, A. (2011). The relation between instrumental musical activity and cognitive aging. *Neuropsychology,25*(3), 378-386. doi:10.1037/a0021895

Hauser, M.D., Chomsky, N., & Fitch, T. (2002). The faculty of language: what is it, who has it, and how did it evolve? S*cience.* doi:10.1126/science.298.5598.1569

Heffner, C.C., & Slevc, R. (2015). Prosodic structure as a parallel to musical structure. *Frontiers in psychology, 6.* doi:10.3389/fpsyg.2015.01962

Heim, S., Opitz, & Friederici, A.D. (2003). Distributed cortical networks for syntax processing: Broca's area as the common denominator. *Brain and Language*, *85* (3), 402–408. doi:10.1016/S0093-934X(03)00068-3

Hoch, L., Poulin-Charronnat, B., & Tillmann, B. (2011). The Influence of Task-Irrelevant Music on Language Processing: Syntactic and Semantic Structures. *Frontiers in Psychology*, *2*, 112. doi:10.3389/fpsyg.2011.00112

Holdgraf, C. R., Rieger, J. W., Micheli, C., Martin, S., Knight, R. T., & Theunissen, F. E. (2017). Encoding and Decoding Models in Cognitive Electrophysiology. *Frontiers in Systems Neuroscience,11*. doi:10.3389/fnsys.2017.00061

Honing, H. (2009). *Musical cognition a science of listening*. New Brunswick, NJ: Transaction.
Jackendoff, R. (2009). Parallels and Nonparallels between Language and Music. *Music Perception: An Interdisciplinary Journal*, *26* (3), 195–204. doi:10.1525/mp.2009.26.3.195

Jackendoff, R., & Lerdahl, F. (2006). The capacity for music: What is it, and what's special about it? *Cognition*, *100* (1), 33–72. doi:10.1016/j.cognition.2005.11.005

Jaeggi, S., Buschkuehl, M., Perrig, W., & Meier, B. (2010). The concurrent validity of the N-back task as a working memory measure. *Memory*, *18* (4), 394-412. doi:10.1080/09658211003702171

James, C., Oechslin, M., Ville, D., Hauert, C.-A., Descloux, C., & Lazeyras, F. (2014). Musical training intensity yields opposite effects on grey matter density in cognitive versus sensorimotor networks. *Brain Structure and Function*, *219* (1), 353–366. doi:10.1007/s00429-013-0504-z

Jentschke, S., Friederici, A.D., & Koelsch, S. (2014). Neural correlates of music-syntactic processing in two-year old children. *Developmental Cognitive Neuroscience*, *9*, 200–208. doi:10.1016/j.dcn.2014.04.005

Jentschke, S., Koelsch, S., Sallat, S., & Friederici, A.D. (2008). Children with Specific Language Impairment Also Show Impairment of Music-syntactic Processing.*Journal of Cognitive Neuroscience*, *20*(11), 1940–1951. doi:10.1162/jocn.2008.20135

Justus, T., & Hustler, J. (2005). Fundamental issues in the evolutionary psychology of music: Assessing innateness and domain specificity. *Music Perception: An Interdisciplinary Journal, 23* (1), 1-27.

Koechlin, E., & Jubault, T. (2006). Broca's Area and the Hierarchical Organization of Human Behavior. *Neuron*, *50* (6), 963–974. doi:10.1016/j.neuron.2006.05.017

Koelsch, S., & Siebel, W. (2005). Towards a neural basis of music perception. *Trends in Cognitive Sciences*, *9* (12), 578–584. doi:10.1016/j.tics.2005.10.001

Koelsch, S., Gunter, T. C., Wittfoth, M., & Sammler, D. (2005). Interaction between Syntax Processing in Language and in Music: An ERP Study. *Journal of Cognitive Neuroscience,17*(10), 1565-1577. doi:10.1162/089892905774597290

Koelsch, S., Gunter, T., Cramon, D. Y., Zysset, S., Lohmann, G., & Friederici, A.D. (2002). Bach Speaks: A Cortical "Language-Network" Serves the Processing of Music. *NeuroImage*, *17* (2), 956–966. doi:10.1006/nimg.2002.1154

Koelsch, S., Rohrmeier, M., Torrecuso, R., & Jentschke, S. (2013). Processing of hierarchical syntactic structure in music. *Proceedings of the National Academy of Sciences*, *110* (38), 15443–15448. doi:10.1073/pnas.1300272110

Krause, F., & Lindemann, O. (2014). Expyriment: A Python library for cognitive and neuroscientific experiments. *Behavior Research Methods,46* (2), 416-428. doi:10.3758/s13428-013-0390-6

Kunert, R., Willems, R. M., & Hagoort, P. (2016). Language influences music harmony perception: Effects of shared syntactic integration resources beyond attention. *Royal Society Open Science,3*(2), 150685. doi:10.1098/rsos.150685

Kunert, R., Willems, R. M., Casasanto, D., Patel, A. D., & Hagoort, P. (2015). Music and Language Syntax Interact in Broca's Area: An fMRI Study. *Plos One,10*(11). doi:10.1371/journal.pone.0141069

LaRocque, J., Lewis-Peacock, J., & Postle, B. (2014). Multiple neural states of representation in short-term memory? It's a matter of attention. *Frontiers in Human Neuroscience*, *8* (5). doi:10.3389/fnhum.2014.00005

Lawson, S. (2012). Consilience Revisited: Musical and Scientific Approaches to Chinese Performance. *Ethnomusicology,56*(1), 86. doi:10.5406/ethnomusicology.56.1.0086

Lemaire, J.-J., Golby, A., Wells, W., Pujol, S., Tie, Y., Rigolo, L., … Kikinis, R. (2013). Extended Broca's Area in the Functional Connectome of Language in Adults: Combined Cortical and

Subcortical Single-Subject Analysis Using fMRI and DTI Tractography. *Brain Topography*, *26* (3), 428–441. doi:10.1007/s10548-012-0257-7

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

Lerdahl, F., & Krumhansl, C. L. (2007). Modeling Tonal Tension. *Music Perception: An Interdisciplinary Journal,24*(4), 329-366. doi:10.1525/mp.2007.24.4.329

Limb, C., & Braun, A. (2008). Neural Substrates of Spontaneous Musical Performance: An fMRI Study of Jazz Improvisation. *PLoS ONE*, *3* (2), e1679. doi:10.1371/journal.pone.0001679

Maess, B., Koelsch, S., Gunter, T., & Friederici, A.D. (2001). Musical syntax is processed in Broca's area: an MEG study. *Nature Neuroscience*, *4* (5), 540–545. doi:10.1038/87502

Makuuchi, M., Bahlmann, J., Anwander, A., & Friederici, A. D. (2009). Segregating the core computational faculty of human language from working memory. *Proceedings of the National Academy of Sciences*, *106* (20), 8362–8367.

Martins, M. (2012). Distinctive signatures of recursion. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *367* (1598), 2055–2064.

Martins, M., & Fitch, T. (2012). Empirical approaches to recursion. *The Evolution Of Language*. doi:10.1142/9789814401500_0029

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

Miendlarzewska, E., & Trost, W. (2014). How musical training affects cognitive development: rhythm, reward and other modulating variables. *Frontiers in Neuroscience*, *7*, 279. doi:10.3389/fnins.2013.00279

Miyake, A., Friedman, N., Emerson, M., Witzki, A., Howerter, A., & Wager, T. (2000). The Unity and Diversity of Executive Functions and Their Contributions to Complex "Frontal Lobe" Tasks: A Latent Variable Analysis. *Cognitive Psychology*, *41* (1), 49–100. doi:10.1006/cogp.1999.0734

Monti, M., & Osherson, D. (2012). Logic, language and the brain. *Brain Research*, *1428*, 33–42. doi:10.1016/j.brainres.2011.05.061

Musso, M., Weiller, C., Horn, A., Glauche, V., Umarova, R., Hennig, J., … Rijntjes, M. (2015). A single dual-stream framework for syntactic computations in music and language. *NeuroImage*, *117*, 267–283. doi:10.1016/j.neuroimage.2015.05.020

Novick, J., Trueswell, J., & Thompson-Schill, S. (2005). Cognitive control and parsing: Reexamining the role of Broca's area in sentence comprehension. *Cognitive, Affective, & Behavioral Neuroscience*, *5* (3), 263–281. doi:10.3758/CABN.5.3.263

Oechslin, M., Ville, D., Lazeyras, F., Hauert, C.-A., & James, C. (2013). Degree of Musical Expertise Modulates Higher Order Brain Functioning. *Cerebral Cortex*, *23* (9), 2213–2224. doi:10.1093/cercor/bhs206

Opitz, B., & Friederici, A.D. (2004). Brain Correlates of Language Learning: The Neuronal Dissociation of Rule-Based versus Similarity-Based Learning. *The Journal of Neuroscience*, *24* (39), 8436–8440. doi:10.1523/JNEUROSCI.2220-04.2004

Paap, K., & Partridge, D. (2014). Recursion Isn't Necessary for Human Language Processing: NEAR (Non-iterative Explicit Alternatives Rule) Grammars are Superior. *Minds and Machines*, *24* (4), 389–414. doi:10.1007/s11023-014-9341-y

Patel, A. (2003). Language, music, syntax and the brain. *Nature Neuroscience*, *6* (7), 674–681. doi:10.1038/nn1082

Patel, A. D. (2009). *Music, language, and the brain*. New York, NY: Oxford University Press.

Patel, A. D. (2014). Can nonlinguistic musical training change the way the brain processes speech? The expanded OPERA hypothesis. *Hearing Research,308*, 98-108. doi:10.1016/j.heares.2013.08.011

Patel, A.D. (2005). The relationship of music to the melody of speech and to syntactic processing disorders in aphasia. *Annals of the New York Academy of Sciences*. doi:10.1196/annals.1360.005

Patel, A.D., Iversen, J.R., Wassenaar, M., & Hagoort, P. (2008). Musical syntactic processing in agrammatic Broca's aphasia. *Aphasiology, 7* (8), 776-789.

Peirce JW (2009) Generating stimuli for neuroscience using PsychoPy. *Front. Neuroinform.,2 (*10). doi:10.3389/neuro.11.010.2008

Peretz, I. (2006). The nature of music from a biological perspective. *Cognition*, *100* (1), 1–32. doi:10.1016/j.cognition.2005.11.004

Peretz, I., & Coltheart, M. (2003). Modularity of music processing. *Nature neuroscience*. doi:10.1038/nn1083

Peretz, I., & Hyde, K. (2003). What is specific to music processing? Insights from congenital amusia. *Trends in Cognitive Sciences*, *7* (8), 362–367. doi:10.1016/S1364-6613(03)00150-5

Perlovsky, L. (2009). Language and cognition. *Neural Networks*, *22* (3), 247–257. doi:10.1016/j.neunet.2009.03.007

Perruchet, P., & Poulin-Charronnat, B. (2012). Challenging prior evidence for a shared syntactic processor for language and music. *Psychonomic Bulletin & Review,20*(2), 310-317. doi:10.3758/s13423-012-0344-5

Pinker, S., & Jackendoff, R. (2005). The faculty of language: what's special about it? *Cognition*, *95* (2), 201–236. doi:10.1016/j.cognition.2004.08.004

Ratcliff, R., & Mckoon, G. (2008). The Diffusion Decision Model: Theory and Data for Two-Choice Decision Tasks. *Neural Computation,20*(4), 873-922. doi:10.1162/neco.2008.12-06-420

Rijntjes, M., Weiller, C., Bormann, T., & Musso, M. (2012). The dual loop model: its relation to language and other modalities. *Frontiers in Evolutionary Neuroscience*, *4*, 9. doi:10.3389/fnevo.2012.00009

Rogalsky, C., Matchin, W., & Hickok, G. (2008). Broca's area, sentence comprehension, and working memory: an fMRI study. *Frontiers in Human Neuroscience*, *2*, 14. doi:10.3389/neuro.09.014.2008

Rogalsky, C., Rong, F., Saberi, K., & Hickok, G. (2011). Functional Anatomy of Language and Music Perception: Temporal and Structural Factors Investigated Using Functional Magnetic Resonance Imaging. *The Journal of Neuroscience*, *31* (10), 3843–3852. doi:10.1523/JNEUROSCI.4515-10.2011

Rohrmeier, M. (2011). Towards a generative syntax of tonal harmony. *Journal of Mathematics and Music*, *5* (1), 35–53. doi:10.1080/17459737.2011.573676

Sala, G., & Gobet, F. (2017). When the music's over. Does music skill transfer to children's and young adolescents' cognitive and academic skills? A meta-analysis. *Educational Research Review*, *20*, 55–67. doi:10.1016/j.edurev.2016.11.005

Sammler, D., Koelsch, S., & Friederici, A.D. (2011). Are left fronto-temporal brain areas a prerequisite for normal music-syntactic processing? *Cortex*, *47* (6), 659–673. doi:10.1016/j.cortex.2010.04.007

Sammler, D., Koelsch, S., Ball, T., Brandt, A., Elger, C., Friederici, A.D., … Schulze-Bonhage, A. (2009). Overlap of Musical and Linguistic Syntax Processing: Intracranial ERP Evidence. *Annals of the New York Academy of Sciences*, *1169* (1), 494–498. doi:10.1111/j.1749-6632.2009.04792.x

Schneider, B. A., Avivi-Reich, M., & Mozuraitis, M. (2015). A cautionary note on the use of the Analysis of Covariance (ANCOVA) in classification designs with and without within-subject factors. *Frontiers in Psychology,6*. doi:10.3389/fpsyg.2015.00474

Slevc, L., & Miyake, A. (2006). Individual Differences in Second-Language Proficiency. *Psychological Science*, *17* (8), 675–681. doi:10.1111/j.1467-9280.2006.01765.x

Slevc, R., & Okada, B. (2015). Processing structure in language and music: a case for shared reliance on cognitive control. *Psychonomic bulletin & review, 22* (3), 637–652.

Slevc, R., Reitman, J., & Okada, B. (2013). Syntax in music and language: The role of cognitive control. *CogSci,* 3414-3419.

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

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*(8), 1380-1393. doi:10.1162/jocn.2006.18.8.1380

Swain, J. (1997). The concept of musical syntax. *The Musical Quarterly, 79* (2), 281-308. doi:742247

Szpunar, K.K., Schellenberg E., & Pliner, P. (2004). Liking and memory for musical stimuli as a function of exposure. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30* (2), 370-381.

Tettamanti, M., & Weniger, D. (2006). Broca's area: a supramodal hierarchical processor? *Cortex, 42* (4), 491-494.

Tillmann, B., & Bigand, E. (2004). The Relative Importance of Local and Global Structures in Music Perception. *The Journal of Aesthetics and Art Criticism*, *62* (2), 211–222. doi:10.1111/j.1540-594X.2004.00153.x

Tomalin, M. (2007). Reconsidering recursion in syntactic theory. *Lingua*, *117* (10), 1784–1800. doi:10.1016/j.lingua.2006.11.001

Trehub, S., & Hannon, E. (2006). Infant music perception: Domain-general or domain-specific mechanisms? *Cognition, 101* (1), 73-99.

Tremblay, P., & Dick, A. (2016). Broca and Wernicke are dead, or moving past the classic model of language neurobiology. *Brain and Language*, *162*, 60–71. doi:10.1016/j.bandl.2016.08.004

Vigneau, Beaucousin, Hervé, P.-Y., Jobard, G., Petit, L., Crivello, F., … Tzourio-Mazoyer. (2011). What is right-hemisphere contribution to phonological, lexico-semantic, and sentence processing? Insights from a meta-analysis. *NeuroImage*, *54* (1), 577–593. doi:10.1016/j.neuroimage.2010.07.036

Wilson, E. O. (1998). *Consilience: The Unity of Knowledge*. New York: Alfred A. Knopf.

Zuk, J., Benjamin, C., Kenyon, A., & Gaab, N. (2014). Behavioral and Neural Correlates of Executive Functioning in Musicians and Non-Musicians. *PLoS ONE*, *9* (6), e99868. doi:10.1371/journal.pone.0099868