Music Perception
Fall 1998, Vol. 16, No. 1, 27–42

Syntactic Processing in Language and Music: Different Cognitive Operations, Similar Neural Resources?

ANIRUDDH D. PATEL
The Neurosciences Institute, San Diego

Does the processing of structural relations in music have anything in common with the syntactic processing of language? Important differences in the form, purpose, and use of syntactic structures in the two domains suggest that the answer should be "no." However, recent event-related brain potential (ERP) data suggest that some aspect of syntactic processing is shared between domains. These considerations lead to a novel hypothesis that linguistic and musical syntactic processing engage different cognitive operations, but rely on a common set of neural resources for processes of structural integration in working memory ("shared structural integration resource" hypothesis). This hypothesis yields a nonintuitive prediction about musical processing in aphasic persons, namely, that high- and low-comprehending agrammatic Broca's aphasics should differ in their musical syntactic processing abilities. This hypothesis suggests how comparison of linguistic and musical syntactic processing can be a useful tool for the study of processing specificity ("modularity") in cognitive neuroscience.

LANGUAGE and music afford two instances of rich syntactic structures processed by the human brain. In both domains, discrete elements are ordered in hierarchical patterns according to certain principles of combination (Raffman, 1993). Experienced listeners in a given linguistic or musical culture show implicit knowledge of syntactic patterns and principles in a number of ways, including judgments of correctness, memory advantages for rule-governed sequences, and production of plausible substitutions when linguistic or musical sequences are recalled less than perfectly (Blacking, 1973; Sloboda, 1985).

The existence of these two distinct syntactic systems in the human mind raises the following question: does the processing of structural relations in music have anything in common with the syntactic processing of language? This question addresses the much-debated issue of specificity or modular-

Address correspondence to Aniruddh D. Patel, The Neurosciences Institute, 10640 John Jay Hopkins Dr., San Diego, CA 92191. (e-mail: apatel@nsi.edu)

ity of cognitive operations in the human brain, specifically whether the neural operations underlying linguistic syntactic processing are "informationally encapsulated" (Elman, 1990; Fodor, 1983).

In the first part of this article, I discuss syntactic processing in language in terms that allow a comparison with music. In the second part of the article, I present event-related potential (ERP) data suggesting a similarity in the brain's handling of syntactic structure in the two domains. In the discussion, I propose that what is shared between linguistic and musical syntactic processing are neural resources involved in structural integration in working memory. Based on this hypothesis, I suggest an experiment to test specifically the issue of shared resources. I conclude that contrasting the similarities and differences of linguistic and musical syntactic processing provides a way of studying questions of processing specificity in both domains.

Syntactic Processing in Language and Music

Language and music have important differences in the form, purpose, and use of their syntactic structures, so it might seem puzzling to ask if these structures have any commonalties in terms of cognitive or neural processing.[1] For example, language uses grammatical categories not found in music (e.g., nouns and verbs), and vice versa (e.g., intervals, chords). In sentences, words are hierarchically structured in systems of obligatory dependencies that may be characterized as *head-dependent* (e.g., Ades & Steedman, 1982; Bresnan, 1982; Chomsky, 1995; Pollard and Sag, 1994). Musical tones and chords are linked by probabilistic dependencies (Meyer, 1956; Narmour, 1990, 1992), and the hierarchical structuring of these events may be described in terms of *head-elaboration* relations (Lerdahl & Jackendoff, 1983). Syntactic conventions in language are used to signal "who-did-what-to-whom" (thematic relations), while in music, syntax participates in the dynamic articulation of mobility vs. closure, or "tension-resolution" (Swain, 1997). Finally, speakers generally abide by the norms of linguistic syntax in order to transmit a clear semantic message. Composers and performers of music, free from this "semantic imperative," readily violate syntactic conventions for structural or aesthetic reasons (Sloboda, 1985, 1991).

These differences, however, should not obscure interesting similarities at the level of the *psychological experience* of linguistic and musical syntax.

1. Because the syntax of music is not homogeneous either within or across cultures, it is necessary to choose a particular type of music for comparison to language (Sloboda, 1985). This article focuses on Western music of the tonal period (approximately 1600–1900). The choice is motivated by the large body of theoretical structural analyses of tonal Western music that may be drawn on for comparisons with linguistic syntax.

For example, in both domains it is possible to experience structural ambiguity and resolution in the perception of an unfolding sequence (Sloboda, 1985). In addition, in both domains items take on context-dependent psychological *functions* that are purely relational: for example, "subject" and "predicate" in language (Chomsky, 1965) and "tonic" and "dominant" in functional harmony (Dahlhaus, 1990).[2] Finally, both language and music perception crucially depend on memory and integration in the perception of structural relations between elements.

In this article, I focus on the issue of structural integration, using a recently proposed psycholinguistic theory to make this notion explicit for language and to allow comparison with music. "Syntactic Prediction Locality Theory" (Gibson, 1998) provides a metric for structural integration and structural maintenance (memory) costs during syntactic processing. One basic intuition behind the theory is that syntactic parsing involves simultaneous processes of prediction and structural integration. Prediction is elicited by knowledge of what word categories are required to complete a grammatical sentence given the input up to the current moment. For example, if a listener hears "The reporter who...," he or she knows that at that point at least two obligatory verbs are required to finish the sentence: a verb for the embedded clause initiated by "who" and a main verb for "reporter." Maintaining these predictions until the required words are encountered is associated with a memory cost.

Integration is the other side of prediction. When each new word is encountered, it must be integrated with the current structure. In Gibson's model, integration cost increases as distance between the current word and prior dependent words increases.[3] For example, consider the relationship between the words "reporter" and "attacked" in sentences that begin "The reporter who attacked the senator..." vs. "The reporter who the senator attacked..." In the first case, when "attacked" is reached, integration with its dependent "reporter" is relatively easy because the co-dependent noun and verb are nearly adjacent in the string. In the second sentence, the integration between "attacked" and "reporter" (now the object of the verb) is more difficult, because it must cross an intervening noun phrase ("the senator").[4] This difference in processing cost is supported by psycholinguistic data such as reading times (see Gibson, 1998, and references therein).

2. No direct analogy between subjects/predicates and tonics/dominants is intended here. The key point is that both domains have syntactic functional categories that have meaning only *in relation to each other,* a kind of "mutually referential" semiosis.

3. "Distance" is measured by the number of new discourse referents between dependent words. For details, see Gibson (1998). The model is only sketched in outline here. The full model provides an explicit metric for both memory and integration costs on a word-by-word basis for any given sentence.

4. Technically the integration is between "attacked" and an empty-category object that is coindexed with the pronoun "who." For the purposes of a simplified exposition, this subtlety is not essential.

Gibson's model also predicts high integration costs for words that cause syntactic reanalysis of the sentence due to a local ambiguity ("garden-path" effects), because these involve integration with alternative structures that are not highly activated when the word is encountered.[5]

What is the significance of this model to music perception? The importance lies not in the details (which are unlikely to be relevant to music perception, because music does not have obligatory dependencies), but in the fact that the model is cast in terms that allow a comparison between syntactic processing in language and music. That is, it seems quite plausible to assume that music perception involves continuous prediction and integration of events into a perceived relational structure in working memory. The word "expectation" may be more suitable than "prediction" for music, because musical syntactic relations are expressed in terms of probabilities rather than obligatory dependencies. "Integration," however, seems to be well-suited to both language and music, because in both cases the mind must integrate incoming elements with remembered previous elements in order to understand their place in the unfolding network of structural relations. These integrations can be easy or difficult. For language, Gibson's model suggests that difficulty is associated with long-distance integrations or integrations that force a revision of the current syntactic structure. For music, difficult integrations are likely to be those that are unexpected because of melodic, harmonic, or rhythmic factors.[6]

Neural Correlates of Syntactic Processing in Language and Music

The data described here were collected as part of a study that used music to examine the language specificity of a known neural correlate of syntactic processing, the P600 brain potential (Patel, Gibson, Ratner, Besson, & Holcomb, in press). The P600 is a positive component of the ERP elicited by words that are difficult to integrate structurally into sentences. For example, Osterhout and Holcomb (1992, 1993) found that in sentences of the type "The broker persuaded to sell the stock was sent to jail," a P600 was elicited by the word "to," relative to the same word in "The broker hoped to sell the stock." "To" is more difficult to integrate in the first

5. Garden-path sentences include "The dog walked to the park was chasing the squirrel," where the word "was" is difficult to integrate because it forces a reinterpretation of "walked" from a main verb to an embedded relative clause.

6. By calling a musical integration "difficult," I do not mean to imply that the listener necessarily experiences a conscious sense of struggle, difficulty, or displeasure, simply that the processing and psychological experience of the event are different than if the same event occurred in an expected context. This difference can be demonstrated by using a number of techniques from psychomusicology, including reaction times (Bigand & Pineau, 1997) and event-related potentials (Besson, 1997).

sentence because it forces a reanalysis of the verb "persuaded" from a simple active verb (as in "The broker persuaded his client to sell...") to a more complex reduced-relative clause (i.e., "The broker [who was] persuaded to sell…").[7] The P600 is thus quite different from the better-known language ERP component, the N400, a negative-going waveform sensitive to semantic relations rather than syntactic ones (Kutas & Hillyard, 1980, 1984). One important feature shared by both waveforms is that they can be observed without frank anomalies (as in the preceding example) and are thus not simply generic brain "error detectors" for syntax and semantics.

We set out to test the language specificity of the P600 by examining the brain's response to syntactic incongruities in both language and music. Western tonal music was chosen because of its syntactic properties of harmony and key, which allow a listener to detect harmonic anomalies (e.g., out-of-key notes or chords) in novel sequences, analogously to the way competent speakers of a particular language can detect a syntactic error in a sentence they have never heard before.[8] If syntactic anomalies in both domains elicited the P600, this would indicate that the component does not uniquely reflect syntactic parsing in language. The impetus to undertake such a comparison came from work by Besson and Faïta (1995) showing that harmonic anomalies elicit a positive-going ERP waveform whose maximum is in the 600-ms range.

METHODS

Fifteen musically trained subjects listened to sentences and musical phrases while ERPs were recorded from 13 sites around the head (for details of EEG recording, see Patel et al., in press).[9] Both linguistic and musical sequences contained a target region that was syntactically congruous, moderately incongruous, or highly incongruous with the preceding structural context. This hierarchy was created with principles of phrase structure for language, and principles of harmony and key-relatedness for music. Two levels of incongruity were used so that comparisons could be made at both levels between linguistic and musical waveforms in terms of amplitude, polarity, and scalp distribution. In such comparisons, highly similar patterns of ERPs across the scalp are thought to reflect a common or largely overlapping set of neural generators (Rugg & Coles, 1995).

---

7. The P600 has also been called the "syntactic positive shift" and has been elicited by both closed-class words (such as "to" in the above example) and open-class words (such as nouns and verbs), for example, in reversals of word-order or number-agreement violations (Hagoort, Brown, & Groothusen, 1993).

8. In this article, the term "harmonic anomaly" is simply shorthand for "unexpected given the prevailing harmonic context," not a judgment of artistic value.

9. All subjects had significant musical experience (mean, 11 years), had studied music theory, and played a musical instrument (mean, 6.2 hours/week). None had perfect pitch. Musically trained subjects were selected because music perception research indicates that they are more likely to be sensitive to harmonic relations than their untrained counterparts (Krumhansl, 1990). For both sentences and musical phrases, subjects were asked to make off-line acceptability judgments approximately 1500 ms after the end of each sequence. These judgments were collected to ensure attention to the stimuli and were not used in analysis of the ERP data.

For language, sentences varied in syntactic structure before a fixed target phrase (see Figure 1, top). The carrier sentence has three possible syntactic structures depending on the word(s) in the slot. The differences in syntax are intended to vary the difficulty with which the target phrase (shown in the shaded region) can be integrated into the sentence's structure. Condition I is a simple declarative sentence. Condition II is much more complex since "endorsed" functions as an embedded relative clause (i.e., "Some of the senators [who were] endorsed..."). Type II sentences cause processing difficulties because of a strong preference to interpret the "endorsed" verb as a simple active verb when it is first encoun-

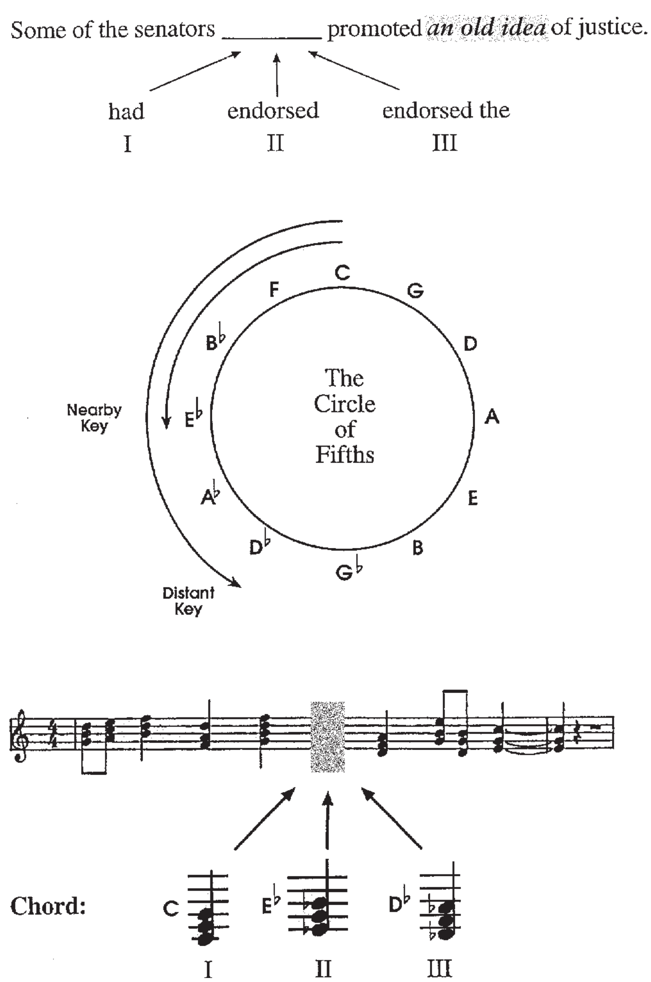

Fig. 1. Examples of varying difficulty of structural integration in language (top) and music (bottom). In both domains, the target region (shaded) varies in how easily it can be structurally integrated with the rest of the sequence, depending on whether Element I, II, or III is inserted in the sequence (see text for details). In the case of the musical phrase, the three target chords are chosen according to their position on the circle of fifths relative to the key of the musical phrase: same-key, nearby-key (three steps away), or distant-key (five steps away).

tered (Frazier & Rayner, 1982; Trueswell, Tanenhaus, & Garnsey, 1994). This preference is accounted for by theories based on structural representations, frequency-based arguments, or memory and integration costs (e.g., Frazier & Rayner, 1982; Gibson, 1998; MacDonald, Pearlmutter, & Seidenberg, 1994). As a result of the parsing preference, when the target region is encountered in Condition II, subjects are engaged in syntactic reanalysis processes that were initiated upon perceiving the previous word (e.g., "promoted"). This causes syntactic integration difficulty that is not present in condition I. In contrast to conditions I and II (which are grammatical), in condition III the occurrence of the target phrase renders the sentence ungrammatical. Thus the target phrase should generate the greatest integration difficulty (and thus the largest P600) in this condition.

A sample musical phrase is shown in the bottom part of Figure 1. The phrase represents an ordinary harmonic progression, with voice-leading patterns and rhythms representative of popular rather than classical styles. The target chord position is indicated by the shaded region. For a musical phrase of a given key, the target chord could come from (I) the same key, (II) a "nearby key", or (III) a "distant key." "Nearby" and "distant" were defined as three or five steps counterclockwise away from the key of the phrase on the circle of fifths. Music perception research using reaction time measures has shown that musically trained listeners perceive harmonic distance between keys in a manner that reflects key distance on the circle of fifths, suggesting that the target chords would vary systematically by how easily they could be structurally integrated into the rest of the sequence (Bharucha & Stoeckig, 1986, 1987). The target chord was always the tonic chord of the key in question and was always a quarter note in duration. By choosing tonic chords of keys three and five steps away, the "nearby"- and "distant"-key chords did not differ in the number of accidentals they introduced into the phrase. This ensured that any difference in positivities elicited by these chords was not due to a difference in the *number* of out-of-key tones, but rather to the harmonic *distance* of the chord as a whole from the native key of the phrase.

Each subject heard 30 sentences of type I, II, and III, plus 60 filler sentences, randomized and divided into five blocks of 30 sentences each. Sentences were presented as connected speech (female voice) at an average rate of 4.4 syllables per second, and had an average overall duration of 4 s. Each subject also heard 36 musical phrases of type I, II, and III, plus 36 filler phrases without harmonic incongruities, randomized and divided into four blocks of 36 phrases each. Language and music blocks were alternated in presentation. Musical phrases varied in length from 7 to 12 chords, and had between 4 and 9 chords before the target. These phrases averaged about 6 s in duration (tempo = 120 beats/minute). For musical phrases sharing the same harmonic progression, such as the three versions of the phrase of Figure 1, some of the pre- and post-target chords were played in different inversions when appearing with different targets (though two beats before and one beat after the target were always held constant). These variations were introduced to avoid identical repetition of any particular context (for similar reasons, sentences sharing the same target phrase were given different first nouns). All 12 major keys were equally represented among the musical phrases. In order to avoid priming of the in-key target, the principal chord of a key was always avoided before the target position.

RESULTS

Detailed analyses of the results are given in Patel et al. (in press): only the points relevant to the current discussion are reviewed here. Grand average ERPs to target phrases in simple, complex, and ungrammatical sentences are shown in Figure 2. A hierarchy of effects is observed, with target phrases in ungrammatical sentences eliciting the greatest positivity, followed by

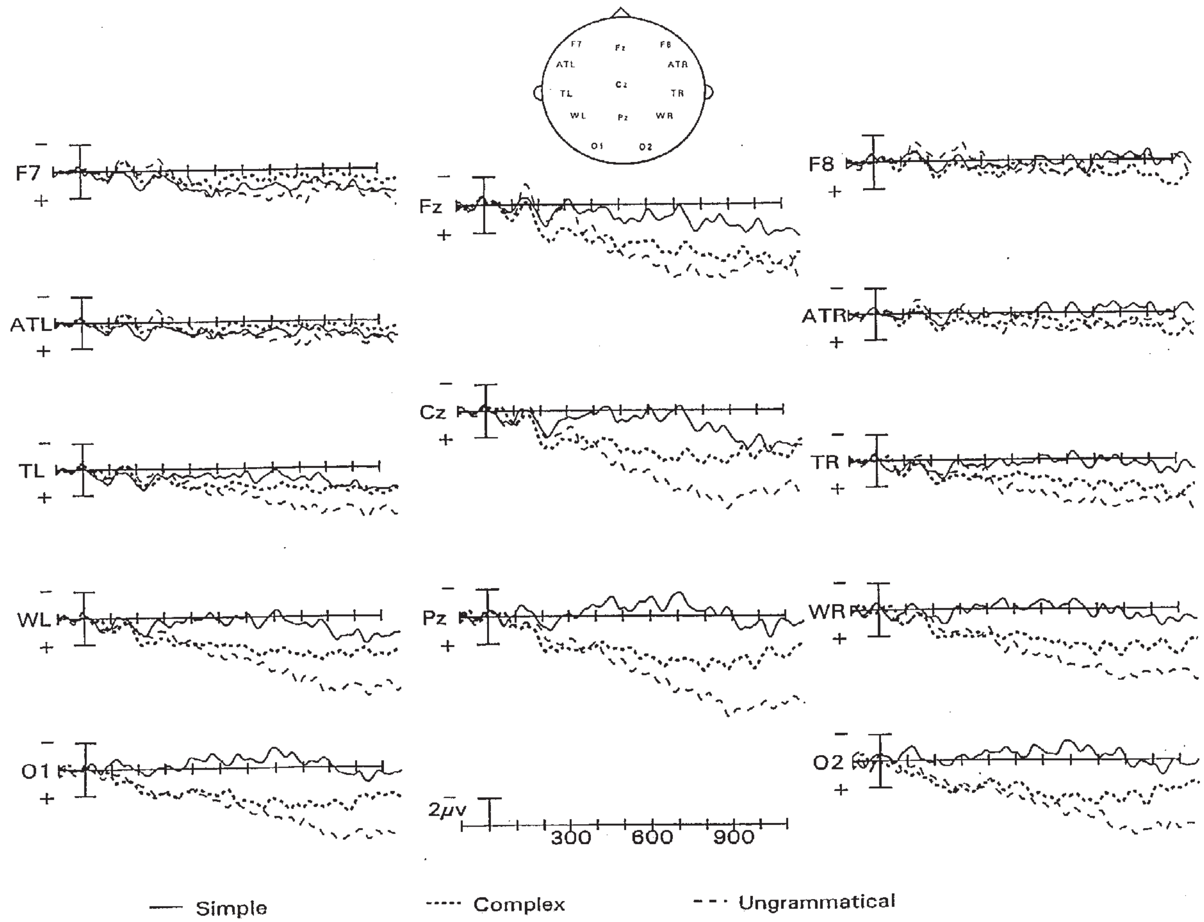

Fig. 2. Grand average ERPs from 13 scalp sites time-locked to target phrases in grammatically simple, complex, and ungrammatical sentences. Each plot represents averages made over approximately 1300 trials. A key to electrode positions is given at the top of the figure. (Fz, Cz, Pz = frontal, central, parietal midline sites; F7, F8 = left and right frontal sites; ATL, ATR = anterior temporal left/right; TL, TR = temporal left/right; WL,WR = Wernicke's left/right homolog; O1,O2 = occipital left/right.)

target phrases in grammatically complex and grammatically simple sentences. A repeated measures analysis of variance (ANOVA) of mean waveform amplitude revealed that both the complex and the ungrammatical sentences were significantly more positive than the simple sentences in the 500–800 ms range and that the difference between the complex and ungrammatical sentence reached significance in the 800–1100 ms range. A hierarchy of effects is also visible in music (Figure 3), with the distant-key chord eliciting the greatest positivity, followed by the nearby-key chord and the in-key chord. Results revealed that the nearby-key and distant-key chords were significantly more positive than the in-key chord in the 500-800 ms range, and that the difference between the distant- and nearby-key targets was also significant in this range.

Differences in the morphology of the ERP waveforms in language and music were due to surface acoustic differences. For example, the musical targets elicit a negative-positive (N1-P2) complex 100 ms after onset, and once again 500 ms later, because of the onset of the following chord. The salient N1-P2 complexes occur because each chord in the sequence was temporally separated from the next chord by about 20 ms of silence. In connected speech, the lack of pauses between words makes the N1-P2 complex refractory (Holcomb & Neville, 1991). In order to compare linguistic and musical ERP waveforms without contamination by these domain-specific effects, difference waveforms were computed by subtracting the waveform of Condition I from the waveforms of Conditions II and III (Figures 4 and 5). Then, linguistic and musical difference waves were compared in the latency range of the P600: a repeated-measures ANOVA was conducted in the 450–750 ms range for both the II – I and III – I comparisons, with domain (i.e., language vs. music) as a main variable. Of particular interest were interactions of domain and electrode site, as differences in scalp distribution would suggest different source generators for the waveforms. However, results showed no main effect and no interaction. Thus in the latency range of the P600, the positivities to structurally incongruous elements in language and music do not appear to be distinguishable.

A new and unexpected result of this study was the observation of a brief right-hemisphere negativity in response to out-of key target chords (Figures 3–5, right electrode sites, 300–400 ms). This asymmetry showed a reliable Condition x Hemisphere x Electrode Site interaction in this time range, reflecting its right anterotemporal distribution.[10]

10. Space limitations prevent discussion of this component in the current article, and interested readers are referred to Patel et al. (in press) for discussion. Here I merely note that it is clear that the observed right anterotemporal negativity does not resemble the semantic N400 (which differs in symmetry, duration, and scalp distribution).

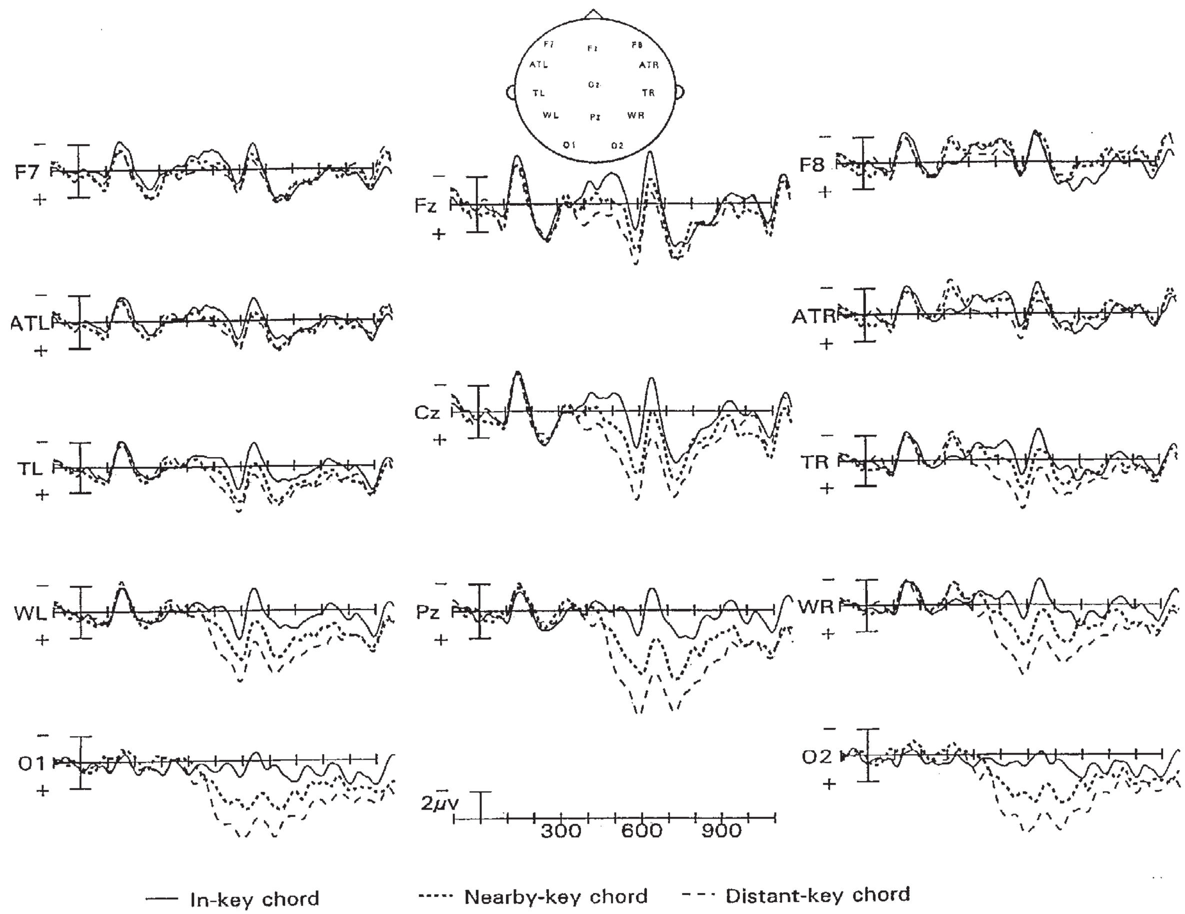

Fig. 3. Grand average ERPs time-locked to the three target chord types in a phrase of a given musical key. Each plot represents averages made over approximately 1600 trials. Note that the temporal epoch of the figure also includes the chord following the target chord (onset of second chord is 500 ms after target onset).

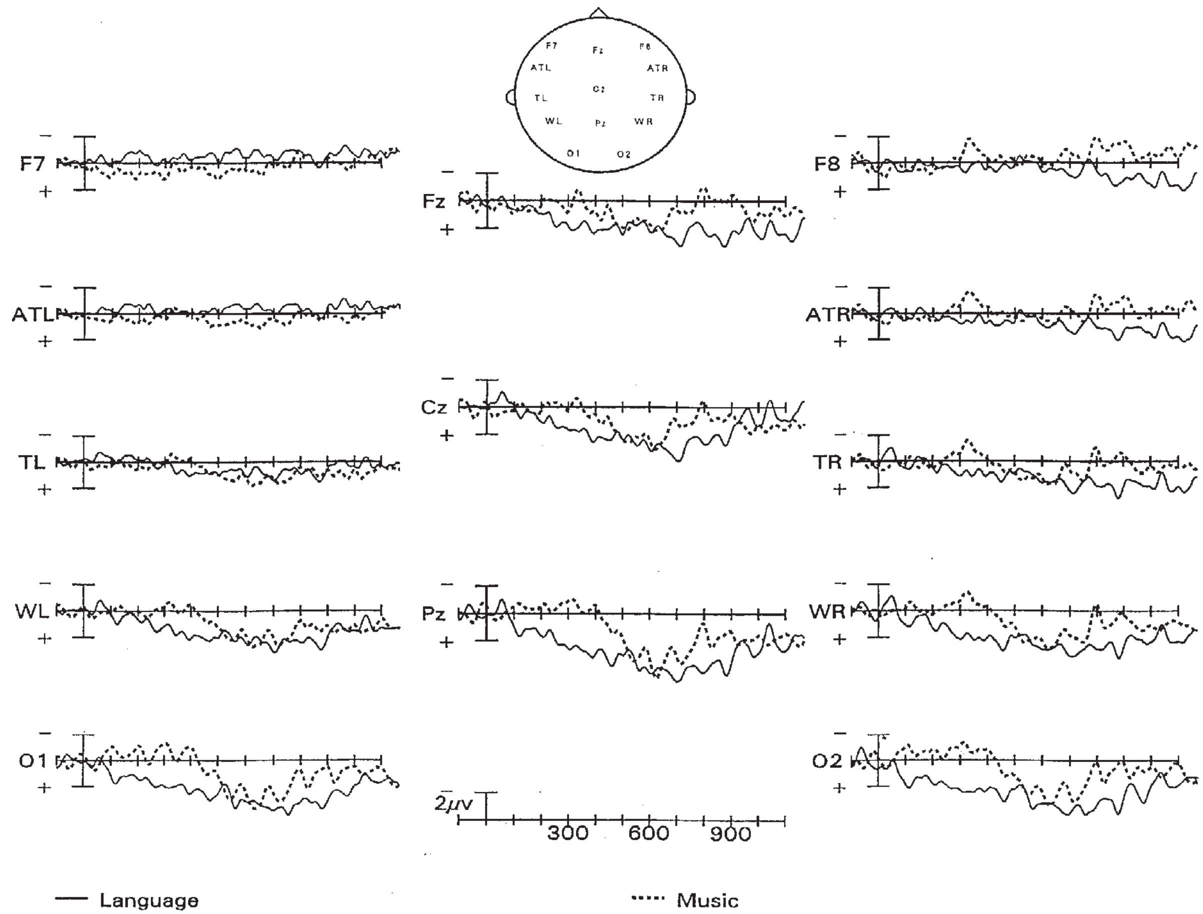

Fig. 4. Difference waves for condition II – I in language and music. The solid line represents the target phrase in a grammatically complex sentence minus the target phrase in a grammatically simple sentence. The dashed line represents the nearby-key target chord minus the in-key target chord.

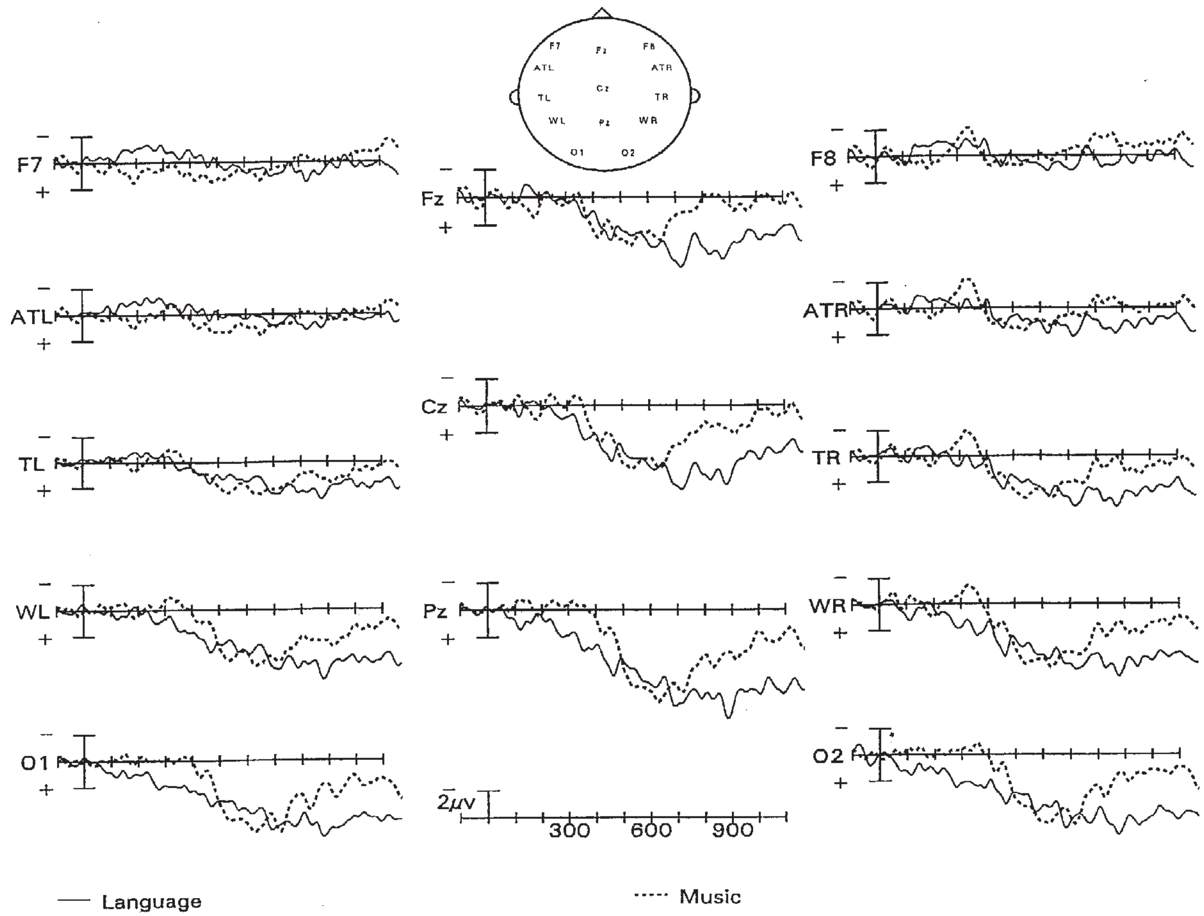

Fig. 5. Difference waves for condition III – I in language and music. The solid line represents the target phrase in an ungrammatical sentence minus the target phrase in a grammatically simple sentence. The dashed line represents the distant-key target chord minus the in-key target chord.

Discussion

The question posed at the beginning of this paper was "does the processing of structural relations in music have anything in common with the syntactic processing of language?" Two lines of reasoning suggest that the answer to this question should be "no." First, the brief review provided in this article of the differences in form, use, and purpose of linguistic and musical syntax suggests that the cognitive operations underlying these two forms of syntax are quite different. Second, it has been demonstrated that the perception of an important aspect of musical syntax (harmonic relations) can be selectively impaired after brain damage, without concomitant linguistic syntactic deficits (Peretz, 1993; Peretz et al., 1994). Yet the data presented here suggest that the story is more complicated. Syntactic incongruities in both domains lead to a positive ERP waveform (P600) similar in polarity, amplitude, and scalp distribution in a specified latency range, suggesting that there is something common to syntactic processing in language and music.

I believe these observations can be reconciled under the following hypothesis: linguistic and musical syntactic processing rely on distinct cognitive operations, but structural integration in both domains relies on a common pool of neural resources. Put another way, the forging of structural links between elements of a sequence in working memory bears a neural cost. Although the principles that forge these links are different in language and music, the brain regions that provide the neural resources that sustain these costs are the same or similar. In making this claim, I wish to be precise in what is meant by structural "integration." For language, I mean the linking of the current input word to past dependents in a string of words, with the assumption that this integration is more costly when dependencies are more distant, when they must reactivate dispreferred structures (as in Gibson, 1998), or when they are simply impossible (in which case a large integration cost is incurred, but to no avail). For music, I mean the linking of the current input element to past probabilistically related elements, with the assumption that integration is more costly when the occurrence of the current element was not implied by past elements. In this view, the P600 is an index of integration cost in both linguistic and musical syntactic processing, and its similarity across domains represents the common resources from which this cost is paid. I call this the "shared structural integration resource" (SSIR) hypothesis. The SSIR hypothesis is consistent with the neuropsychological data of Peretz under the assumption that her subject suffered damage to a domain-specific knowledge base of harmonic relations and not to structural integration processes per se. The deficit results from an inability to access musical harmonic knowledge, rather than a problem with integration itself. Consistent with this idea, selective deficits of harmonic perception have been associated with bilateral damage to temporal association cortices (Peretz, 1993; Peretz et al., 1994), which are likely to be important in the long-term neural representation of harmonic knowledge.

The SSIR hypothesis is of direct relevance to a currently debated issue in psycholinguistics, namely, whether the working memory system used in syntactic processing draws on a dedicated resource pool. Caplan and Waters (in press) review much of the psycholinguistic literature on this topic and argue strongly for a "separate language interpretation resource." Their conclusions are drawn entirely from comparisons between language tasks, but the SSIR hypothesis suggests another way to test the idea of resource modularity in linguistic syntactic processing. Specifically, if a population of persons with reduced available processing resources for language syntax could be found, the SSIR hypothesis predicts that such persons should also show diminished structural integration in musical syntactic processing.

Research in aphasia has suggested that certain aphasics' comprehension deficits are due to a diminution in their processing capacity or efficiency (Frazier & Friederici, 1991; Haarman & Kolk, 1991; Miyake, Carpenter, & Just, 1994). A recent ERP study of syntactic processing in high- vs. low-comprehending agrammatic Broca's aphasics (Wassenaar, Hagoort, & Brown, 1997) revealed that high comprehenders show P600s to linguistic syntactic incongruities whereas low comprehenders do not. If the difference between these populations reflects greatly reduced processing resources in the latter, the SSIR hypothesis predicts that the two groups will show substantial differences in P600s to musical syntactic incongruities. In particular, agrammatic low comprehenders should show an attenuated P600 relative to agrammatic high comprehenders. This experiment has the advantage of comparing one population of brain-damaged persons to another, thus controlling for the presence of brain damage per se. An additional control group of interest would be patients with right-hemisphere damage but without syntactic processing problems. The SSIR hypothesis predicts that such subjects should show musical P600s comparable to the high-comprehending Broca's aphasics, a counterintuitive claim given the typical association between the right hemisphere and musical processing.[11]

In conclusion, the syntactic structures of language and music provide cognitive science with two kinds of valuable information. They provide different windows on the syntactic capabilities of the human mind, and the

11. Right anterotemporal circuits have been implicated in working memory for pitch (Zatorre, Evans, & Meyer, 1994; Patel, Peretz, Tramo, & Labrecque, 1998), so it is possible that diminished musical P600s in patients with right hemisphere damage could be due to a memory deficit. This suggests that persons with right hemisphere damage should be screened for pitch memory before being tested for musical P600s.

similarities and differences between their processing provide new ways of exploring issues of processing specificity in cognitive neuroscience.[12]

References

Ades, A., & Steedman, M. (1982). On the order of words. *Linguistics and Philosophy, 4,* 517–558.

Besson, M. & Faïta, F. (1995). An event-related potential (ERP) study of musical expectancy: Comparison of musicians with nonmusicians. *Journal of Experimental Psychology: Human Perception & Performance, 21,* 1278–1296.

Besson, M. (1997). Electrophysiological studies of music processing. In I. Deliège & J. Sloboda (Eds.), *Perception and cognition of music* (pp. 217–250). Hove, UK: Psychology Press.

Bharucha, J., & Stoeckig, K. (1986). Reaction time and musical expectancy: Priming of chords. *Journal of Experimental Psychology: Human Perception & Performance, 12,* 403–410.

Bharucha, J., & Stoeckig, K. (1987). Priming of chords: Spreading activation or overlapping frequency spectra? *Perception & Psychophysics, 41,* 519–524.

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

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

Bresnan, J. (1982). *The mental representation of grammatical relations.* Cambridge, MA: MIT Press.

Caplan, D., & Waters, G. (in press). Verbal working memory and sentence comprehension. *Behavioral and Brain Sciences.*

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

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

Dahlhaus, C. (1990). *Studies on the origin of harmonic tonality* (R. O. Gjerdingen, Trans.). Princeton, NJ: Princeton University Press.

Elman, J. L. (1990). Representation and structure in connectionist models. In G. T. M. Altmann (Ed.), *Cognitive models of speech processing.* Cambridge, MA: MIT Press.

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

Frazier, L., & Friederici, A. (1991). On deriving the properties of agrammatic comprehension. *Brain and Language, 40,* 51–66.

Frazier, L., & Rayner, K. (1982). Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambiguous sentences. *Cognitive Psychology, 14,* 178–210.

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

Haarman, H., & Kolk, H. (1991). A computer model of the temporal course of agrammatic sentence understanding: The effects of variation in severity and sentence complexity. *Cognitive Science, 15,* 49–87.

Hagoort, P., Brown, C., & Groothusen, J. (1993). The syntactic positive shift (SPS) as an ERP measure of syntactic processing. *Language and Cognitive Processes, 8*(4), 439–483.

Holcomb, P. J., & Neville, H. J. (1991). Natural speech processing: An analysis using event-related brain potentials. *Psychobiology, 19*(4), 286–300.

---

12. I thank Ted Gibson for numerous discussions relating to syntactic structure and parsing and Jennifer Burton for skillful editing. I also thank Patrick Russell, Kevin Long, and two anonymous reviewers for their comments. This work was supported by Neurosciences Research Foundation.

Krumhansl, C. L. (1990). *Cognitive foundations of musical pitch.* Oxford: Oxford University Press.

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

Kutas, M., & Hillyard, S. (1984). Brain potentials during reading reflect word expectancy and semantic association. *Nature, 307,* 161–163.

Lerdahl, F., & Jackendoff, R. (1983). *A generative theory of tonal music.* Cambridge, MA: MIT Press.

MacDonald, M. C., Pearlmutter, N., & Seidenberg, M. (1994). The lexical nature of syntactic ambiguity resolution. *Psychological Review, 101,* 676–703.

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

Miyake, A., Carpenter, P., & Just, M. (1994). A capacity approach to syntactic comprehension disorders: Making normal adults perform like aphasic patients. *Cognitive Neuropsychology, 11,* 671–717.

Narmour, E. (1990). *The analysis and cognition of basic melodic structures.* Chicago: University of Chicago Press.

Narmour, E. (1992). *The analysis and cognition of melodic complexity.* Chicago: University of Chicago Press.

Osterhout, L., & Holcomb, P. J. (1992). Event-related potentials elicited by syntactic anomaly. *Journal of Memory and Language, 31,* 785–806.

Osterhout, L., & Holcomb, P. J. (1993). Event-related potential and syntactic anomaly: Evidence of anomaly detection during the perception of continuous speech. *Language and Cognitive Processes, 8*(4), 413–437.

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

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

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

Peretz, I., Kolinsky, R., Tramo, M., Labrecque, R., Hublet, C., Demeurisse, G., & Belleville, S. (1994). Functional dissociations following bilateral lesions of auditory cortex. *Brain, 117,* 1283–1301.

Pollard, C., & Sag, I. (1994). *Head-driven phrase structure grammar.* Chicago: University of Chicago Press.

Raffman, D. (1993). *Language, music, and mind.* Cambridge, MA: MIT Press.

Rugg, M., & Coles, M. G. H. (1995). The ERP and cognitive psychology: Conceptual issues. In M. Rugg & M. G. H. Coles (Eds.), *Electrophysiology of mind* (pp. 27–39). Oxford: Oxford University Press.

Sloboda, J. (1985). *The musical mind: The cognitive psychology of music.* Oxford: Clarendon Press.

Sloboda, J. (1991). Musical structure and emotional response: Some empirical findings. *Psychology of Music, 19,* 110–120.

Swain, J. (1997). *Musical languages.* New York: Norton.

Trueswell, J. C., Tanenhaus, M. K., & Garnsey, S. M. (1994). Semantic influences on parsing: Use of thematic role information in syntactic disambiguation. *Journal of Memory and Language, 33,* 285–318.

Wassenaar, M., Hagoort, P., & Brown, C. (1997). Syntactic ERP effects in Broca's aphasics with agrammatic comprehension. *Brain and Language, 60,* 61–64.

Zatorre, R. J., Evans A. C., & Meyer, E. (1994). Neural mechanisms underlying melodic perception and memory for pitch. *Journal of Neuroscience, 14*(4), 1908–1919.