COGNITIVE SCIENCE
A Multidisciplinary Journal

Musical Garden Paths: Evidence for Syntactic Revision Beyond the Linguistic Domain

Gabriele Cecchetti,[a] Steffen A. Herff,[a,b] Martin A. Rohrmeier[a]

[a]*Digital and Cognitive Musicology Lab, École Polytechnique Fédérale de Lausanne*
[b]*The MARCS Institute for Brain, Behaviour and Development, Western Sydney University*

Abstract

While theoretical and empirical insights suggest that the capacity to represent and process complex syntax is crucial in language as well as other domains, it is still unclear whether specific parsing mechanisms are also shared across domains. Focusing on the musical domain, we developed a novel behavioral paradigm to investigate whether a phenomenon of syntactic revision occurs in the processing of tonal melodies under analogous conditions as in language. We present the first proof-of-existence for syntactic revision in a set of tonally ambiguous melodies, supporting the relevance of syntactic representations and parsing with language-like characteristics in a nonlinguistic domain. Furthermore, we find no evidence for a modulatory effect of musical training, suggesting that a general cognitive capacity, rather than explicit knowledge and strategies, may underlie the observed phenomenon in music.

Keywords: Ambiguity; Music and language; Reanalysis; Revision; Syntactic processing; Syntax

1. Introduction

Syntactic parsing accounts for the computational operation of inferring representations of syntactic structure from sequential inputs (Jurafsky & Martin, 2009; Sipser, 2012). For syntactic parsing to be understood as a model of cognition beyond the pure *computational* level of description (Marr, 1982), it is necessary to account for how processing is implemented at the
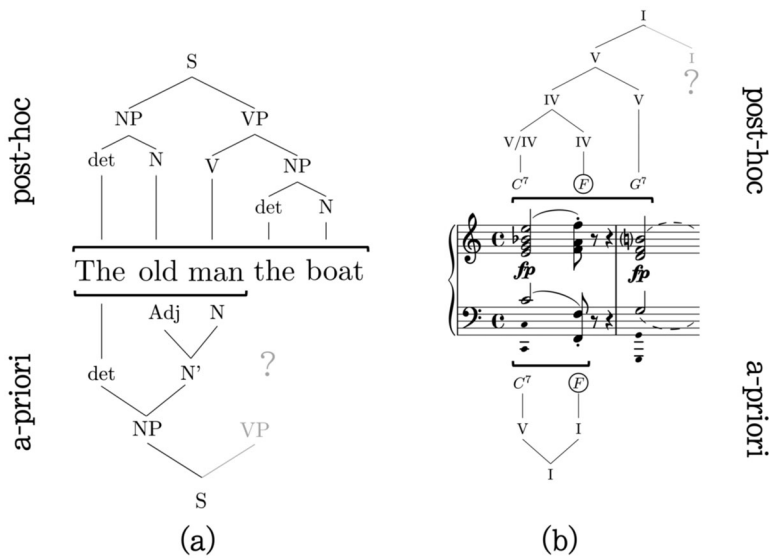
Fig. 1. Syntactic revision in language and music. (a) A-priori (bottom) and post-hoc (top) interpretations of a garden-path sentence. (b) A similar phenomenon is predicted to occur in music (Rohrmeier, 2013), as exemplified here with the changing syntactic interpretation hypothesized to occur in the opening of L. van Beethoven's Symphony op. 21 before (a-priori, bottom) and after (post-hoc, top) the presentation of the third chord. Musical syntactic interpretations are represented as syntactic trees according to Rohrmeier and Neuwirth (2015), and question marks indicate open dependency relations, entailing expectations of future events.

algorithmic level through cognitive parsing strategies that cope with ambiguity, limited memory resources, and with the temporal unfolding of parsing itself (Narayanan & Jurafsky, 1998; Vogelzang et al., 2017). In particular, the revision mechanisms that deal with ambiguity and temporarily misled syntactic interpretations are thoroughly investigated in psycholinguistics (Fodor & Ferreira, 1998; Kaan & Swaab, 2003).

Syntactic organization has also been argued to govern nonlinguistic stimuli, such as music (Fitch & Martins, 2014; Fitch, Hauser, & Chomsky, 2005; Jackendoff, 2007; Lerdahl & Jackendoff, 1983; Patel, 2010), but it has not been empirically investigated whether effects analogous to garden-path effects occur in music and whether the parsing strategies involved in the processing of such structures would resemble those observed in language. In addressing this issue, this paper presents explicit perceptual evidence for a revision effect to occur in the processing of tonal melodies.

1.1. Syntactic revision in language

Structural representations emerge incrementally as a sentence is gradually presented and parsed (Frazier, 1987; Marslen-Wilson, 1973). When processing ambiguous sequences, the representation of structure as perceived may be updated retrospectively upon encountering new information, as prototypically exemplified by the recovery from so-called *garden-path* effects (Fig. 1a; Ferreira & Henderson, 1991; Frazier, 1978). While reading the first part of

the garden-path sentence in Fig. 1a ("The old man…"), the most likely interpretation is to understand "man" as a noun and to expect a Verb Phrase to follow (a-priori interpretation, bottom). After exposure to the second part of the sentence ("… the boat"), the previously most likely interpretation is replaced by a different one, where "man" serves as a verb (post-hoc interpretation, top). Note how parsing "[…] the boat" serves here as a critical event that requires the most likely syntactic role of the word "man" to change retrospectively from noun to verb, although, by this time, "man" lies in the past. This change is retrospective because the interpretation that is most likely after the critical event may differ from the interpretation that is most likely before the critical event not only in terms of how it accounts for the critical event and those that follow, but also in terms of how it accounts for the events that precede the critical event. The occurrence of such a retrospective change of interpretation is associated with cognitively demanding recovery processes that manifest themselves, for example, in slower reading times (Frazier & Rayner, 1982; Meseguer, Carreiras, & Clifton, 2002) and characteristic patterns of brain activity (Meltzer & Braun, 2011) following the critical event itself.

Theoretical and empirical literature in linguistics presents diverging accounts of which processing mechanisms underlie sentence processing and, specifically, the garden-path effect and recovery from it (Sprouse & Lau, 2013). In particular, in cases of syntactic ambiguity, it is debated whether only one syntactic representation is parsed at any given time (serial parsing) or rather several alternatives among the plausible ones are parsed simultaneously (parallel parsing). From a serial-processing perspective, behavioral and Event-Related-Potential (ERP) evidence is interpreted as suggesting that separate early and late processes are involved in sentence comprehension (Friederici, 1995; Friederici & Mecklinger, 1996): the former are argued to implement a first parsing attempt that rapidly assigns a structural interpretation to the incoming information, while the latter implement any adjustments to the outcome of the early processing (e.g., through reanalysis) if incompatible information (e.g., the critical event in a garden-path sentence) is presented (Meltzer & Braun, 2011). However, alternative interpretations of ERPs (Hagoort, 2003) alongside conflicting behavioral evidence (Hickok, 1993; Nicol & Pickering, 1993; Trueswell, Tanenhaus, & Garnsey, 1994) rather support a ranked parallel perspective, whereby multiple coexisting interpretations are continuously ranked and eventually pruned based, for example, on lexical (MacDonald, Pearlmutter, & Seidenberg, 1994; McRae, Spivey-Knowlton, & Tanenhaus, 1998; Trueswell et al., 1994) and complexity (Gibson, 1991) constraints. Serial and parallel models are not easy to disambiguate, as in many cases, they make broadly compatible predictions (Gibson & Pearlmutter, 2000; Lewis, 2000). In particular, while garden-path effects have a natural explanation within a serial-processing perspective (Frazier & Rayner, 1982; Friederici, 1995), parallel-processing models can also offer alternative explanations for the same effects (Gibson & Pearlmutter, 2000): as a consequence, observing a retrospective change of the most likely syntactic interpretation as perceived before and after the critical event does not rule out either family of accounts, although the proposed underlying mechanism would be different. Specifically, in serial-processing accounts, such retrospective change results from the need to generate a new representation of the stimulus once encountering the critical event (*reanalysis*; Frazier & Rayner, 1982). Differently, in a parallel-processing perspective, it is the likelihood and ranking of

alternative coexisting parses that is updated (*reranking*; Gibson & Pearlmutter, 2000). In either case, it is possible to define a "preferred" interpretation as the only (in a serial account) or top-ranked (in a parallel account) representation that is generated by the processor at any given time. The phenomenological effect exemplified in Fig. 1a, the switch from one preferred interpretation to a different one, is then characterized irrespectively of the underlying serial or parallel algorithmic account by the following qualitative conditions:

(1)　One of the plausible interpretations of an ambiguous stimulus, the a-priori interpretation, is initially preferred;

(2)　A critical event occurs that is unlikely under the (currently preferred) a-priori interpretation;

(3)　If the critical event is consistent with an alternative structural interpretation of the entire stimulus, including the critical event and those preceding it, this new post-hoc interpretation becomes the preferred one.

Although the term *revision* is often employed equivalently to *reanalysis*, for conciseness, we use here the term (*syntactic) revision* to refer to a phenomenon occurring under conditions (1)–(3), which is a plausible phenomenological manifestation of either a putative *reanalysis* process as well as of a putative *reranking* process in terms of a change of preferred interpretation (depending on the assumed underlying serial or parallel model). In this study, we seek to demonstrate the existence of this phenomenological effect of syntactic revision in a nonlinguistic domain, such as music.

1.2. Syntax and syntactic revision in music

Complex hierarchical structure has been theorized in the musical domain (Herff, Harasim, Cecchetti, Finkensiep, & Rohrmeier, 2021; Lerdahl & Jackendoff, 1983; Rohrmeier, 2011, 2020; Schenker, 1935; Steedman, 1984), where syntax captures idiom-specific recursive dependency relationships linking musical events, which, in turn, motivate the corresponding patterns of creation and resolution of expectancy (Cecchetti, Herff, Finkensiep, & Rohrmeier, 2021; Rohrmeier, 2013). Some degree of formal analogy between such syntactic structures in music and those in language has been highlighted repeatedly in the literature (e.g., Baroni, Maguire, & Drabkin, 1983; Bernstein, 1976; Jackendoff, 2009; Katz & Pesetsky, 2011). It has also been proposed that the listeners' experience of abstract musical structure is the result of a parsing process based on generative rules (Berent & Perfetti, 1993; Jackendoff, 1991; Rohrmeier & Pearce, 2018), and that common neural and cognitive resources are involved in linguistic and musical syntactic processing (Koelsch, 2011; Patel, 2010). The properties of such a putative musical syntactic processor have been discussed on theoretical grounds. Jackendoff (1991), arguing in favor of a parallel processing architecture, predicted a musical "retrospective reanalysis" effect based on a "selection function" singling out one preferred parse in the presence of ambiguity:

> The processor is computing multiple analyses in parallel, and [(1)] evidence has accumulated for one of these to be chosen as most plausible by the selection function.

However, [(2)] subsequent events in the musical surface lead to a relative reweighting of the analyses being computed by the processor. The selection function thereby [(3)] "changes horses in midstream" jumping to a different analysis. The phenomenological effect of such an occurrence will be a "retrospective reanalysis" of the passage as it is heard. (p. 223)

As highlighted by the numbering added in brackets, this prediction is fully compatible with conditions (1)–(3). Nevertheless, experimental evidence for the very existence of such a garden-path effect in music has yet to be found.

Musical garden paths are frequently presented as a compositional device in analytical accounts of Western tonal music (Caplin, 1998; Lewin, 1986; Martin & Vande Moortele, 2014; Rohrmeier, 2013; Schmalfeldt, 2017; Temperley, 2001). A common example is displayed in Fig. 1b. While listening to the opening bars of L. van Beethoven's *Symphony* op. 21, the most likely a-priori interpretation for the first two chords (bottom) is replaced by a new post-hoc interpretation when the third chord is encountered (top). Note how the F chord (circled), initially likely heard as a tonic (I in the key of F major), may be reinterpreted as a subdominant (IV in the key of C major) when the third chord intervenes.

However, unlike the effect of updating expectations over future events (Pearce & Wiggins, 2012; Sears, Spitzer, Caplin, & McAdams, 2020), the perceptual and cognitive nature of revision of musical structure has not received much empirical attention. Additionally, it is even unclear whether revision should exist at all in music: while the success of the parsing process in language is subject to the evolutionary pressure of effectively formulating (Friederici, Chomsky, Berwick, Moro, & Bolhuis, 2017) and communicating (Pinker & Jackendoff, 2005) propositional content, the nature of musical communicative interactions may not require arbitrary specificity and deterministic agreement among interactants (Cross, 2009; Fitch, 2006; Jackendoff, 2009). If reaching an unambiguous and definitive parse would not be crucial in music as it is in language, especially in the absence of formal musical training, it is not granted that processing musical structure would rely on spontaneous strategies to repair failed parsing attempts, requiring the formation and maintenance of structural representations that may be subject to a retrospective update. As a consequence, even if music theory predicts the existence of musical syntactic revision, such a phenomenon may occur spontaneously with lesser frequency or even not occur at all during music listening.

Empirical approaches to syntactic processing in music have identified musical counterparts of neural markers (Patel, Gibson, Ratner, Besson, & Holcomb, 1998) that are known to be associated with ambiguity (Frisch, Schlesewsky, Saddy, & Alpermann, 2002) and second-pass (re)analysis (Friederici & Mecklinger, 1996; Kaan & Swaab, 2003; Osterhout, Holcomb, & Swinney, 1994) in the linguistic domain. Behavioral interference between the linguistic garden-path effect and generic musical syntactic violations (Slevc, Rosenberg, & Patel, 2009) has also been demonstrated, but the unambiguous nature of the musical stimuli (as opposed to the linguistic ones) does not afford the inference that the competing cognitive processes were analogous at the computational and algorithmic level. Overall, such evidence is consistent with analogous or concurrent processing between musical stimuli and linguistic garden-path sentences, possibly relying on cognitive-control resources shared across domains

(Ogg, Okada, Novick, & Slevc, 2019; Slevc & Okada, 2015). However, cross-domain processing interference alone does not prove the substitution of a previously active representation with a different one: showing the existence of a phenomenon with this feature would be necessary to identify revision. In other words, evidence from cross-domain resource sharing shows that some aspect of the implementation of processing is shared, not necessarily that the same revision processes (as characterized in (1)–(3) above) are performed. Furthermore, despite the abundance of theoretical examples of musical ambiguity (Jackendoff, 1991; Rohrmeier, 2013; Slevc & Okada, 2015), no empirical studies have directly addressed this phenomenon by adopting revisable musical stimuli in a controlled experimental setting, while previous attempts to specifically contrast linguistic and musical syntactic revision with harmonically ambiguous stimuli comparable to garden-path sentences have led to inconclusive results (Ross, 2014).

1.3. Aims and hypotheses

Compared to the linguistic case, establishing a phenomenology of musical processing is hampered by the methodological difficulties of capturing perceptual correlates of syntactic representations in music. In language, the availability of specific syntactic representations may be tested through explicit verbalization or semantic matching (e.g., matching sentences with visual representations of their meaning; Meltzer & Braun, 2011), which is not straightforward to achieve in music. In particular, while most speakers can explicitly report their interpretation of a sentence, it is not to be expected that music listeners, especially untrained listeners, would be able to do the same. To address this issue, the present paradigm was designed to prompt behavioral responses that can be read as proxies of syntactic interpretations, even in the absence of semantic references. By accessing listener's syntactic interpretation of ambiguous tonal melodies, we aim at testing whether such interpretations were revised from an a-priori to a different post-hoc one as a consequence of a disambiguating critical event perceived as unlikely (hence surprising) under the a-priori interpretation. Overall, in analogy to linguistic syntactic revision, we hypothesize that a phenomenon unfolding as outlined in (1)–(3) above and exemplified in Fig. 1 occurs in the processing of ambiguous tonal melodies upon presentation of a disambiguating critical event. We further assess whether such an effect is based on a general cognitive capacity or rather explicit domain-knowledge by considering the impact of formal musical training.

2. Methods

2.1. Participants

Sixty-two participants (median age 25.5, range 18–74) took part in an online study (ethics approval granted by the IRB of the École Polytechnique Fédérale de Lausanne, HREC 037–2020). The sample represents a wide range of musical expertise, as reflected by a median score 30 (range 7–45) in the Musical Training subscale of the Goldsmith MSI (Müllensiefen, Gingras, Musil, & Stewart, 2014). For comparison, Müllensiefen et al. reported a mean score of

26.52 ($SD = 11$) for a large validation sample of Western listeners. All participants reported close familiarity with at least one genre within Western musical practices (e.g., classical, Jazz, and Rock/Pop).

2.2. Stimuli

Fifteen distinct original melodies, collectively ranging from C4 to G5 with 440 Hz tuning and each spanning 2 bars in 4/4 meter at 120 bpm, were synthesized in MuseScore 3.5.0 in the default piano timbre. In their original transposition, melodies were composed with the goal of being interpretable in the key of C-major in the absence of any accidentals, while by flattening pitch B to B flat they can be interpreted in the key of F-major. Specifically, each melody affords to be harmonized with idiomatic chord progressions in either one of the two keys, given the appropriate accidentals. Ambiguous stimuli were then obtained from each melody by mistuning all occurrences of pitch B by a quarter-tone, halfway between B and B flat, as highlighted by the box in Fig. 2. These 15 stimuli, each comprising some mistuned notes, are used in the main experimental task. All stimuli are available as Material S4.

Importantly, modes and keys are not only sets of notes (e.g., scales), but come with specific typical melodic and harmonic motions that determine (functional) relationships between notes (Bostwick, Seror, & Neill, 2018; Large, Kim, Flaig, Bharucha, & Krumhansl, 2016; Lerdahl, 2001). Hearing a melody in a key results in attributing interpretations to each note, specifying their relationships to all other notes (cf. Schenker, 1935). These key-specific relationships may be expected to be updated when a melodic excerpt is suddenly perceived in a different key. As an example, Fig. 2 reports two alternative tree analyses for one of the melodies based on the established generative grammar for tonal harmony proposed by Rohrmeier and Neuwirth (2015). Specifically, the formalism models hierarchical harmonic structure in terms of an Abstract Context Free Grammar (Harasim, Rohrmeier, & O'Donnell, 2018) based on two rule types: preparation, $X \rightarrow YX$, and prolongation, $X \rightarrow XX$, where X and Y stand for chord symbols (expressed, e.g., as Roman numerals). The two tree analyses capture the syntactic constituency structure that a listener may perceive when interpreting the melody in C major or F major, respectively. The tuning of note B as a B natural or B flat disambiguates between the two alternatives: a B natural may be interpreted as the third of a V ("leading tone") in C major, which then prepares a C major chord; a B flat may be interpreted in F major as, for example, the seventh of a V, which then prepares an F major chord, or as the third of a ii, which then prepares a C major chord as the V of F.

In this framework, revision occurs when a listener's preferred parse for the melody, captured by one of the two tree structures, is made implausible by the occurrence of a key-defining chord at the end of the melody, and eventually the listener's preferred parse for the entire stimulus, including the chord, is best represented by the other tree structure, consistently with conditions (1)–(3). Crucially, since the C major scale contains no B flat and the F major scale contains no B natural, a listener hearing the quarter-tone note B half flat as a $\hat{7}$ in the key of C major might find more appropriate to replace the quarter-tone note with an equal-tempered B natural, while a listener hearing the quarter-tone note as a $\hat{4}$ in the key of F major may find more appropriate to replace it with a B flat. In other words, the preferred
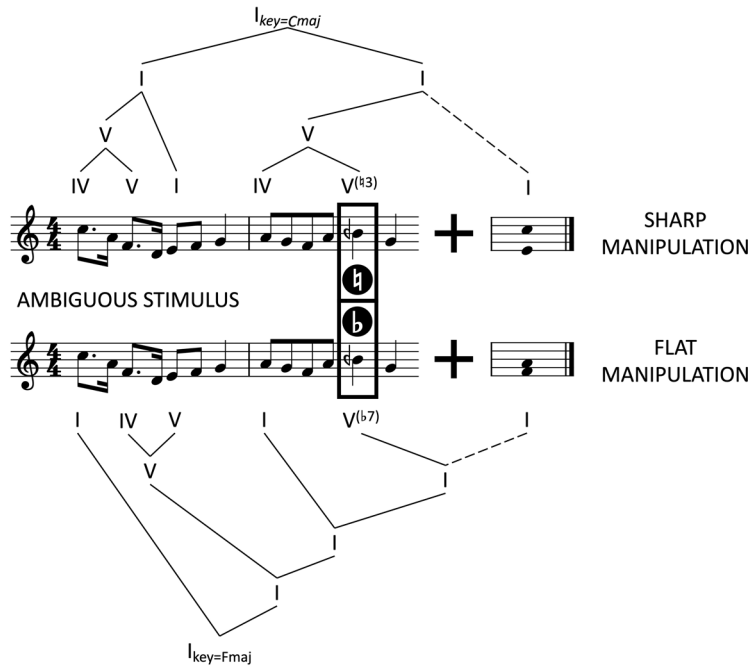
Fig. 2. Example of a stimulus. A tonally ambiguous stimulus, with a key-defining note mistuned by a quarter tone (in the box). The presentation of a sharp or flat manipulation, in the form of a two-voiced chord (root and third of C major or F major, respectively), at the end of the melody may bias the interpretation of the stimulus toward the corresponding key. We also highlight two different tree analyses after Rohrmeier and Neuwirth (2015), exemplifying the two interpretations. The B half-flat may be heard either (top) as the leading-tone in a dominant (V) chord in C, to be tuned upward as a B natural, or rather (bottom) as the seventh of a dominant chord in F, to be tuned downwards as a B flat. Since the note is tuned halfway between B and B flat, both interpretations are plausible until the chord is presented.

equal-tempered approximate tuning of the quarter-tone note can be used as a proxy to infer the listener's syntactic interpretation.

Note that, in principle, listeners may have perceptual biases that deviate from equal-tempered tuning, thus making quarter-tones only an approximation of the perceptual midpoint between scale tones. In order to mitigate this potential source of variability, we make sure listeners are primed to equal temperament by presenting equal-tempered melodies throughout the experiment, and we further account for systematic individual biases across participants by allowing for the corresponding random effect in our analyses.

2.3. Experimental task

In the presentation phase of each trial, a stimulus was played randomly in one of the 12 chromatic transpositions. Over the 30 trials of the experimental task, each stimulus was presented twice in random order with different manipulations. No proximity constraints were imposed by design, but the effect of proximity between presentations of the same stimulus

was explicitly investigated in the analysis. In each trial, the manipulation consisted in the presentation of a two-voiced chord that removed the key ambiguity: in the original transposition, a C-major chord constituted the *sharp* manipulation (cf. the top staff in Fig. 2, where the manipulation is displayed as the chord in the last bar) and an F-major chord constituted the *flat* manipulation (cf. the bottom staff in Fig. 2). Following the manipulation, two behavioral variables were measured:

a. *Surprise Rating*. Participants were asked to rate how surprising the final chord sounded to them. Ratings were provided on a quasi-continuous Visual Analog Scale (Hayes & Patterson, 1921) ranging from *Expected* to *Surprising*.

b. *Tuning Response*. Participants were then presented with the stimulus again, additionally transposed by an ascending or descending tritone and preceded by 3 s of white noise to minimize proactive interference from the manipulation at the end of the first presentation toward the second presentation. This time, the stimulus was interrupted right before the first occurrence of the detuned pitch. Participants were then instructed to select a note to continue the melody in the way that most closely resembled how they remembered the melody itself from the presentation phase of that trial. Two options were given, corresponding to the *sharp* (in the original transposition, B) and *flat* (B flat) tuning of the mistuned note, respectively. Each option was associated randomly with a key on the participants' keyboard (Q or P), and participants could play either option arbitrarily until they were ready to confirm their response with another key press.

In order to address the hypothesis that a phenomenon occurring under conditions (1)–(3) (cf. Introduction) plays a role in the processing of the ambiguous stimuli, we need to be able to access and compare the a-priori and the post-hoc interpretations of each melody. This is achieved through the two behavioral measures. In each trial, the Surprise Rating carries information concerning the a-priori interpretation of the melody. A low Surprise Rating suggests that a strong a-priori interpretation of the stimulus was formed, which happened to be the same as the one implied by the manipulation (e.g., C-major for the *sharp* manipulation in the original transposition), whereas a high Surprise Rating suggests that the a-priori interpretation was different to the one implied by the manipulation. Specifically, a high Surprise Rating indicates that the manipulation was perceived as having a low likelihood conditional to the a-priori interpretation up to that point in the melody (Pearce & Wiggins, 2012), thus potentially serving as a critical event consistently with condition (2). In turn, the selection of one tuning over the other in the Tuning Response is a proxy of the post-hoc interpretation, as it captures a representation of the melody that participants accessed after the manipulation had been presented.

Since each melody was presented with both manipulations, we can compare the corresponding Surprise Ratings. A small difference between the two Surprise Ratings suggests either that no strong a-priori interpretation was formed in either trial, so that both manipulations resulted in only average surprise, or that the a-priori interpretation had changed across the two trials, resulting in the two different manipulations being perceived as similarly surprising. None of these scenarios matches both conditions (1) and (2) in our working definition

of revision: condition (1) requires an a-priori interpretation to be preferred for the melody prior to, and independently of, the manipulation, so that only one of the two manipulations is perceived as an unlikely critical event inducing the need for revision to occur as per condition (2). On the contrary, a large difference between the two Surprise Ratings indicates that an interpretation is strongly preferred for the melody in both trials, and that in the more surprising trial, this interpretation is at odds with the manipulation, so that in such a trial, both conditions (1) and (2) are satisfied. If revision occurs in this trial, we further expect the posthoc interpretation, as captured by the Tuning Response, to be updated in accordance with the manipulation, so that condition (3) is also satisfied. In other words, if a melody undergoes revision, we expect its two Tuning Responses to be different from each other and in music-theoretical agreement with the manipulation in the respective trials. While it is possible that this latter scenario may also occur alongside a small difference in Surprise Ratings, we conservatively only interpret a systematic co-occurrence of this circumstance with a large difference in Surprise Ratings as evidence for the occurrence of revision as defined at the outset in terms of conditions (1)–(3).

2.4. General procedure

After providing informed consent, participants performed a memory task based on a continuous-recognition paradigm (Herff, Olsen, & Dean, 2018; Herff, Olsen, Dean, & Prince, 2018; Shepard & Teghtsoonian, 1961), where the 15 ambiguous stimuli were presented twice in random order and transposition. Results from the memory task are reported in Cecchetti, Herff, and Rohrmeier (2021). During the memory task, participants had the chance to familiarize themselves with the stimuli, potentially (but not necessarily) settling on some preferred parsing that could eventually be prone to be revised. Whether this happened or not does not impact the interpretation of the results reported in this study. Following the memory task, participants took part in the main revision experiment described above and finally answered the Goldsmith Music Sophistication Index (MSI) questionnaire (Müllensiefen et al., 2014). The entire experimental session lasted 45–60 min in total.

2.5. Analysis

For each melody and participant, a new variable is defined, *Congruency*, with the three categories *Congruent*, *Incongruent*, and *Stable*. A melody falls in the *Congruent* category if the Tuning Responses to both presentations of the melody are consistent with the manipulation adopted in the corresponding trial (e.g., *sharp* Tuning Response in a trial where the *sharp* manipulation was presented, and *vice versa*). A melody falls in the *Incongruent* category if the opposite happens in both trials, while it falls in the *Stable* category if the Tuning Response is the same irrespective of the manipulation. By definition, the occurrence of revision for a given melody as music-theoretically predicted would place that melody into the *Congruent* category. In addressing our main hypothesis that syntactic revision occurs in music, we show then that *Congruent* Tuning Responses are more likely as the difference in Surprise Rating (*DiffSurprise*) between the two presentations of the same melody increases.

Data were analyzed with Bayesian mixed-effects models provided with weakly informative priors ($t(3,0,1)$; Gelman, Jakulin, Pittau, & Su, 2008) and implemented with the R package *brms* (Bürkner, 2018). All noncategorical variables were standardized to null mean and unit standard deviation, and we then fitted the model:

$$Congruency \sim DiffSurprise + MusicalTraining + DiffSurprise \times MusicalTraining$$
$$+ InterveningTrials + (1|Participant) + (1|Stimulus),$$

where $DiffSurprise = MaxSurprise - MinSurprise$ and *MusicalTraining* is quantified by the corresponding subscale of the MSI. We also tested for an effect of the number of intervening trials (*InterveningTrials*) separating the two presentations of the same melody, to account for potential interference effects (Herff & Czernochowski, 2019; Herff et al., 2018). Specifically, it is possible that the two trials of the same stimulus interfere more strongly with one another, the closer they are together. The model also allows for random effects accounting for the individual variability across participants and for the differences across the 15 stimuli. We report coefficient estimates (β), their estimated error (*EE*), as well as evidence ratios (*Odds*) for the individual hypotheses (a given coefficient being larger or smaller than zero), labeled as "significant" (*) at a 0.05 confidence level when exceeding 19 (Milne & Herff, 2020). Data and code are available as Materials S2 and S3.

3. Results

Stable responses (56.67% of the total 930 observations) were most likely in general ($p < .001$ in a one-sided binomial test). In other words, participants tended to form a definite *a-priori* interpretation of each stimulus that was usually retained irrespectively of the manipulation, even if the latter was found in conflict with the interpretation itself. However, in the cases in which changes in the post-hoc interpretation occurred, participants responses were not distributed randomly between *Congruent* and *Incongruent* responses. Instead, we found strong evidence that *DiffSurprise* ($\beta = 0.16$, $EE = 0.07$, $Odds(\beta > 0) = 188.47^*$) carries predictive power toward the *Congruency* of Tuning Responses. Specifically, as shown in Fig. 3 by the upward trend in the blue line, the predicted probability of observing *Congruent* responses for a stimulus increases with the difference between the two surprise ratings, and significantly exceeds the likelihood of observing *Incongruent* responses for $DiffSurprise > 0.91$ ($Odds(0.91 \cdot DiffSurprise > (Intercept_1 + Intercept_2)/2) = 19.01^*$; see Material S1 for details). As discussed in Section 2.3, this suggests that a process satisfying conditions (1)–(3) outlined in the Introduction is observed, and thus supports the occurrence of syntactic revision in music.

We only observed weak to no evidence that this effect was further shaped by *MusicalTraining* ($\beta = 0.11$, EE = 0.08, $Odds(\beta > 0) = 12.28$) or its interaction with *DiffSurprise* ($\beta = -0.04$, EE = 0.06, $Odds(\beta > 0) = 0.36$), or that *InterveningTrials* influenced *Congruency* ($\beta = -0.04$, EE = 0.06, $Odds(\beta < 0) = 2.40$).
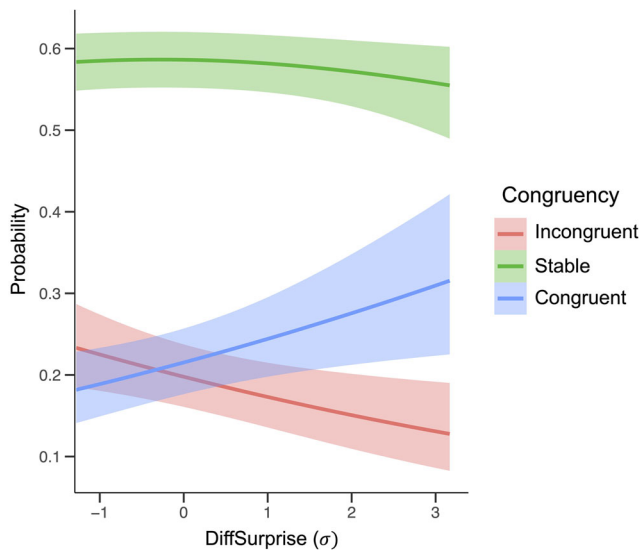
Fig. 3. Marginal effect of DiffSurprise on Congruency. Predicted probability for participants' post-hoc interpretations of a melody to be in accordance with the manipulation in both (Congruent category, blue), neither (Incongruent category, red) or only one (Stable category, green) of the two presentations of the same melody, expressed as a function of the difference in Surprise Rating (DiffSurprise, expressed in standard-deviation units from the mean) between the two presentations of the melody. The prevalence of Congruent over Incongruent responses for increasing values of DiffSurprise supports the hypothesis that syntactic revision occurs in the processing of the musical stimuli. Stable responses are most likely in general.

4. Discussion

In this study, a novel behavioral paradigm was developed to test for the existence of a perceptual correlate of syntactic revision in the domain of music. Data support that the retrograde integration of new information, as provided by the manipulation presented at the end of a stimulus, results in updated syntactic representations, as captured by Surprise and Tuning Responses. Specifically, a bias in favor of revising the preferred tonal interpretation of the ambiguous stimuli in accordance with the manipulation, rather than randomly, was observed selectively in conditions that match psycholinguistic accounts of garden-path-sentence processing (e.g., Frazier, 1978).

We characterized syntactic revision as a phenomenon, whereby (1) listeners form a preferred a-priori interpretation of an ambiguous stimulus, (2) a critical event occurs that is unlikely under the a-priori interpretation, and (3) a change occurs from the a-priori interpretation, which is preferred prior to the critical event, to a different post-hoc interpretation consistent with the critical event. As a consequence, we hypothesized that, if a phenomenon of musical revision exists, the occurrence of conditions (1) and (2) would increase the likelihood of condition (3).

Our stimuli comprised ambiguous melodies allowing participants to form two plausible interpretations. At the end of each melody, a disambiguating final chord was presented and a Surprise Rating was measured. Two different chords were used, corresponding to the two different interpretations. Given a melody, the observation that one chord was perceived as expected and the other as surprising indicates that listeners had formed the same a-priori interpretation of the melody in both cases, and that under this a-priori interpretation, the surprising chord was perceived as unlikely, that is, as a critical event. A large difference in surprise for a given melody indicates then that conditions (1) and (2) were fulfilled in the more surprising trial.

Following the disambiguating chord, we then asked participants to report their memory of the melody as a proxy of the post-hoc interpretation. We observed that, as the difference in surprise increased, participants were not only more likely to report a post-hoc interpretation consistent with the chord in trials where the chord was perceived as highly expected, but crucially also in trials where the chord was perceived as highly surprising. This supports that conditions (1)–(2), captured by the difference in surprise, were predictive toward (3), captured by the reported post-hoc interpretation—which is consistent with the occurrence of musical revision.

We also observed that participants tended to form a strong and stable *a-priori* interpretation of the melodies that was maintained regardless of the number of intervening trials separating the two presentations, so that a surprising manipulation did not deterministically result in participants reporting a revised post-hoc interpretation. Note that, in the linguistic domain, comprehenders also often fail to recover from garden-path effects, rather sticking to some form of good-enough parsing (Ferreira, Christianson, & Hollingworth, 2001). In the musical domain, this behavior may be even more typical, consistently with the understanding that musical syntax serves a different purpose than linguistic syntax (Lerdahl, 2001; Lerdahl & Jackendoff, 1983; Rohrmeier, 2020). In particular, if ungrammaticality is penalized to a minor extent by communicative pressure (cf. Temperley, 2004) in the musical domain, parsing strategies that deal with recovery from temporary failure may not be systematically adopted. Listeners may then ignore conflicting information, or fail to integrate it into a unique coherent parsing together with previous events.

Since a surprising manipulation did only probabilistically, rather than deterministically, lead to revision, one may wonder how often revision occurs when processing stimuli that may allow for it, as those adopted in this study. In this respect, our results should not be taken as a characterization of the frequency of occurrence of revision: as we conservatively focus on a sufficient condition for conditions (1)–(3) to be met, such frequency may be underrepresented by our analysis. However, this exceeds the scope of this proof-of-concept study, which has the goal to establish whether revision occurs at all in musical stimuli.

Characterizing what other conditions influence the occurrence of revision in music is another relevant issue open for further research. Musical training is a natural candidate in this respect. However, while all participants were generally familiar with Western music, no significant effect of explicit musical training on the *Congruency* of *Tuning Responses* was found, nor did musical training modulate the strength of the observed effect of *Diff-Surprise*. Based on these results, the occurrence of musical syntactic revision does not

seem to be explained as a byproduct of explicit domain-specific knowledge, as is acquired through formal training, but likely as the manifestation of a fundamental cognitive operation relying on an implicit syntactic competence, in broad formal analogy to the linguistic one.

It should also be noted that, in each trial of our paradigm, the Tuning Response was determined by parsing a second presentation of the melody. This second parse, in principle, may have been independent of memory of the first (Jackendoff, 1991). The observed systematic dependency of Surprise Ratings and Tuning Responses, however, supports that the two parses were not entirely independent of one another. In other words, the paradigm could be conceived of as inducing a priming effect of the first presentation on the parsing of the second one. Note that the final chord of the first presentation (i.e., the manipulation) is unlikely to have primed the Tuning Response on its own, since the second presentation was transposed. As a consequence, it is a representation integrating both the ambiguous melody and the manipulation that may have primed the Tuning Response in the second presentation, and we showed evidence that this representation may have been revised retrospectively in some trials at least. In principle, the observed bias on the Tuning Response in highly surprising trials may have originated at any point before the participant's Tuning Response, not necessarily while still actively parsing the first presentation of the stimulus. As a consequence, how closely the time course of the phenomenon identified in this study matches the time course of syntactic reanalysis processes in language (Friederici & Mecklinger, 1996; Meltzer & Braun, 2011) remains to be investigated.

More generally, empirical research is still far from characterizing the processing of musical structure in comparable detail as linguistic processing. In language, the existence of a garden-path effect and related phenomena in readers/listeners constrains the characteristics of any plausible model of linguistic processing (cf., e.g., Lewis, 2000), and based on these phenomenological observations, it is then possible to debate, for example, whether a model of parsing should separate syntactic from semantic processing and which one has functional priority (Frazier, 1978; Hagoort, 2003; Sprouse & Lau, 2013), whether processing is serial or parallel (Lewis, 2000), or what factors determine the preferred choices of the human parser (Gibson & Pearlmutter, 1998). A precondition for addressing this type of questions in music is to observe a solid base of phenomena that constrain the properties of a putative musical syntactic processor. In particular, the very existence of an effect of musical syntactic revision has been a crucial yet unsupported assumption of theoretical accounts of musical processing (Jackendoff, 1991; Rohrmeier, 2013; Slevc & Okada, 2015), which now finds preliminary empirical grounding.

In principle, the phenomenon reported here is strictly musical in nature, and it exceeds the scope of the present study to specify the analogy with its linguistic counterpart beyond the broad characterization in terms of conditions (1)–(3) our approach assumed. In particular, our results cannot disambiguate between serial or parallel accounts of syntactic processing, as our assumptions are broad enough to be consistent with both. Further research will also need to clarify whether existing evidence for shared neural and cognitive mechanisms (Fedorenko, Patel, Casasanto, Winawer, & Gibson, 2009; Koelsch, 2013; Patel, 2003) accounts for the phenomenological analogy concerning musical and linguistic revision established in this study–

in particular, whether such shared substrate supports the whole of the processing pipeline (from parsing proper over error detection to revision), and to what extent domain-specific and domain-generic processing modules are involved (Peretz & Coltheart, 2003; Peretz & Zatorre, 2005). In this respect, our study offers a novel empirical paradigm that succeeds in singling out the occurrence of musical revision and that may complement future studies that wish to selectively target this phenomenon.

Overall, by providing the first explicit evidence for syntactic revision of tonally structured music by Western-enculturated listeners, our results show that syntactic revision as a cognitive mechanism is spontaneously deployed in the processing of nonlinguistic stimuli. The present findings suggest that syntactic representations of music develop incrementally during listening, and representations of syntactic relationships linking events in the past appear to persist in memory. Specifically, such representations are prone to retrograde interference due to an ensuing syntactically related critical event and yet robust to merely sensory interference from intervening unrelated trials. Although we found a prevalence of Stable responses, suggesting that revision may not occur frequently in general, the very existence of this phenomenon even in a smaller number of cases challenges the sufficiency of models that only afford the update of expectations toward future events, such as simple Markov chains or n-gram models predicting surface events, as cognitive models of musical structure (cf. Rohrmeier, 2013; Rohrmeier & Pearce, 2018). In particular, modeling a revision effect requires the existence of latent structure (in terms of, e.g., hidden states or nonterminals) encoding alternative, abstract interpretations of a given surface, such as the harmonic function of a given pitch collection and the syntactic relatedness of different harmonic functions. From this perspective, a given musical surface is ambiguous insofar as it can be generated by a multiplicity of latent-structure encodings. Our findings further suggest the necessity of a processing architecture that allows for the retrospective change of the latent structure that is interpreted as generating a portion of the musical surface that belongs in the past: updating transitional probabilities toward states generating events in the future (as in Markov or n-gram models) is not sufficient to account for the observed effect. In summary, these results support the hypothesis that syntactic models that account for abstract musical dependencies based on latent structure are required in describing the cognitive underpinnings of the musical experience, and that qualitatively analogous parsing strategies as those observed in language are likely deployed in music perception.

Acknowledgments

the draft. MR contributed to the conceptualization, provided supervision, reviewed the draft and acquired funding.

Open access funding provided by Ecole Polytechnique Federale de Lausanne.

Funding

Conflicts of interest

The authors declare no conflicts of interest.

References

Baroni, M., Maguire, S., & Drabkin, W. (1983). The concept of musical grammar. *Music Analysis*, *2*(2), 175–208.

Berent, I., & Perfetti, C. A. (1993). An on-line method in studying music parsing. *Cognition*, *46*(3), 203–222.

Bernstein, L. (1976). *The unanswered question: Six talks at Harvard*. Harvard University Press.

Bostwick, J., Seror, G. A., & Neill, W. T. (2018). Tonality without structure using drones to induce modes and convey moods. *Music Perception*, *36*(2), 243–249.

Bürkner, P.-C. (2018). Advanced Bayesian multilevel modeling with the R package brms. *R Journal*, *10*(1), 395–411.

Caplin, W. E. (1998). *Classical form: A theory of formal functions for the instrumental music of Haydn, Mozart, and Beethoven*. Oxford University Press.

Cecchetti, G., Herff, S. A., Finkensiep, C., & Rohrmeier, M. A. (2021). The experience of musical structure as computation: What can we learn? *Rivista Di Analisi e Teoria Musicale*, *26*(2), 91–127.

Cecchetti, G., Herff, S. A., & Rohrmeier, M. A. (2021). Musical syntactic structure improves memory for melody: Evidence from the processing of ambiguous melodies. Proceedings of the Annual Meeting of the Cognitive Science Society.

Cross, I. (2009). The evolutionary nature of musical meaning. *Musicae Scientiae*, *13*(2_suppl), 179–200.

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

Ferreira, F., Christianson, K., & Hollingworth, A. (2001). Misinterpretations of garden-path sentences: Implications for models of sentence processing and reanalysis. *Journal of Psycholinguistic Research*, *30*(1), 3–20.

Ferreira, F., & Henderson, J. M. (1991). Recovery from misanalyses of garden-path sentences. *Journal of Memory and Language*, *30*(6), 725–745.

Fitch, W. T. (2006). The biology and evolution of music: A comparative perspective. *Cognition*, *100*(1), 173–215.

Fitch, W. T., Hauser, M. D., & Chomsky, N. (2005). The evolution of the language faculty: Clarifications and implications. *Cognition*, *97*(2), 179–210.

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

Fodor, J., & Ferreira, F. (1998). *Reanalysis in sentence processing*. Springer Science & Business Media.

Frazier, L. (1978). *On comprehending sentences: Syntactic parsing strategies* [PhD Thesis]. University of Massachusetts.

Frazier, L. (1987). Sentence processing: A tutorial review. In M. Coltheart (Ed.), *Attention and performance 12: The psychology of reading* (pp. 559–586). Lawrence Erlbaum Associates, Inc.

Frazier, L., & Rayner, K. (1982). Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambiguous sentences. *Cognitive Psychology*, *14*(2), 178–210.

Friederici, A. D. (1995). The time course of syntactic activation during language processing: A model based on neuropsychological and neurophysiological data. *Brain and Language*, *50*(3), 259–281.

Friederici, A. D., Chomsky, N., Berwick, R. C., Moro, A., & Bolhuis, J. J. (2017). Language, mind and brain. *Nature Human Behaviour*, *1*(10), 713–722.

Friederici, A. D., & Mecklinger, A. (1996). Syntactic parsing as revealed by brain responses: First-pass and second-pass parsing processes. *Journal of Psycholinguistic Research*, *25*(1), 157–176.

Frisch, S., Schlesewsky, M., Saddy, D., & Alpermann, A. (2002). The P600 as an indicator of syntactic ambiguity. *Cognition*, *85*(3), B83–B92.

Gelman, A., Jakulin, A., Pittau, M. G., & Su, Y.-S. (2008). A weakly informative default prior distribution for logistic and other regression models. *Annals of Applied Statistics*, *2*(4), 1360–1383.

Gibson, E. (1991). *A computational theory of human linguistic processing: Memory limitations and processing breakdown* [PhD Thesis]. Carnegie Mellon University.

Gibson, E., & Pearlmutter, N. J. (1998). Constraints on sentence comprehension. *Trends in Cognitive Sciences*, *2*(7), 262–268.

Gibson, E., & Pearlmutter, N. J. (2000). Distinguishing serial and parallel parsing. *Journal of Psycholinguistic Research*, *29*(2), 231–240.

Hagoort, P. (2003). How the brain solves the binding problem for language: A neurocomputational model of syntactic processing. *Neuroimage*, *20*, S18–S29.

Harasim, D., Rohrmeier, M., & O'Donnell, T. J. (2018). A generalized parsing framework for generative models of harmonic syntax. Proceedings of the 19th International Society for Music Information Retrieval Conference.

Hayes, M. H. S., & Patterson, D. G. (1921). Experimental development of the graphic scale. *Psychology Bullettin*, *18*, 98–99.

Herff, S. A., & Czernochowski, D. (2019). The role of divided attention and expertise in melody recognition. *Musicae Scientiae*, *23*(1), 69–86.

Herff, S. A., Harasim, D., Cecchetti, G., Finkensiep, C., & Rohrmeier, M. A. (2021). Hierarchical syntactic structure predicts listeners' sequence completion in music. Proceedings of the Annual Meeting of the Cognitive Science Society.

Herff, S. A., Olsen, K. N., & Dean, R. T. (2018). Resilient memory for melodies: The number of intervening melodies does not influence novel melody recognition. *Quarterly Journal of Experimental Psychology*, *71*(5), 1150–1171.

Herff, S. A., Olsen, K. N., Dean, R. T., & Prince, J. (2018). Memory for melodies in unfamiliar tuning systems: Investigating effects of recency and number of intervening items. *Quarterly Journal of Experimental Psychology*, *71*(6), 1367–1381.

Hickok, G. (1993). Parallel parsing: Evidence from reactivation in garden-path sentences. *Journal of Psycholinguistic Research*, *22*(2), 239–250.

Jackendoff, R. (1991). Musical parsing and musical affect. *Music Perception*, *9*(2), 199–229.

Jackendoff, R. (2007). *Language, consciousness, culture: Essays on mental structure*. MIT Press.

Jackendoff, R. (2009). Parallels and nonparallels between language and music. *Music Perception*, *26*(3), 10.

Jurafsky, D., & Martin, J. (2009). *Speech and language processing* (2nd edition). Prentice Hall.

Kaan, E., & Swaab, T. Y. (2003). Repair, revision, and complexity in syntactic analysis: An electrophysiological differentiation. *Journal of Cognitive Neuroscience*, *15*(1), 98–110.

Katz, J., & Pesetsky, D. (2011). The identity thesis for language and music. Retrieved from https://Ling.Auf.Net/Lingbuzz/000959

Koelsch, S. (2011). Toward a neural basis of music perception – A review and updated model. *Frontiers in Psychology*, *2*.

Koelsch, S. (2013). *Brain and music*. John Wiley & Sons.

Large, E. W., Kim, J. C., Flaig, N. K., Bharucha, J. J., & Krumhansl, C. L. (2016). A neurodynamic account of musical tonality. *Music Perception*, *33*(3), 319–331.

Lerdahl, F. (2001). *Tonal pitch space*. Oxford University Press.

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

Lewin, D. (1986). Music theory, phenomenology, and modes of perception. *Music Perception*, *3*(4), 327–392.

Lewis, R. L. (2000). Falsifying serial and parallel parsing models: Empirical conundrums and an overlooked paradigm. *Journal of Psycholinguistic Research*, *29*(2), 241–248.

MacDonald, M. C., Pearlmutter, N. J., & Seidenberg, M. S. (1994). Lexical nature of syntactic ambiguity resolution. *Psychological Review*, *101*(4), 676–703.

Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information*. MIT Press.

Marslen-Wilson, W. (1973). Linguistic structure and speech shadowing at very short latencies. *Nature*, *244*(5417), 522–523.

Martin, N. J., & Vande Moortele, S. (2014). Formal functions and retrospective reinterpretation in the first movement of Schubert's String Quintet: Formal functions and retrospective reinterpretation. *Music Analysis*, *33*(2), 130–155.

McRae, K., Spivey-Knowlton, M. J., & Tanenhaus, M. K. (1998). Modeling the influence of thematic fit (and other constraints) in on-line sentence comprehension. *Journal of Memory and Language*, *38*(3), 283–312.

Meltzer, J. A., & Braun, A. R. (2011). An EEG–MEG dissociation between online syntactic comprehension and post hoc reanalysis. *Frontiers in Human Neuroscience*, *5*.

Meseguer, E., Carreiras, M., & Clifton, C. (2002). Overt reanalysis strategies and eye movements during the reading of mild garden path sentences. *Memory & Cognition*, *30*(4), 551–561.

Milne, A. J., & Herff, S. A. (2020). The perceptual relevance of balance, evenness, and entropy in musical rhythms. *Cognition*, *203*.

Müllensiefen, D., Gingras, B., Musil, J., & Stewart, L. (2014). The musicality of non-musicians: An index for assessing musical sophistication in the general population. *PLoS One*, *9*(2).

Narayanan, S., & Jurafsky, D. (1998). A Bayesian model of human sentence processing. Proceedings of the Annual Meeting of the Cognitive Science Society.

Nicol, J. L., & Pickering, M. J. (1993). Processing syntactically ambiguous sentences: Evidence from semantic priming. *Journal of Psycholinguistic Research*, *22*(2), 207–237.

Ogg, M., Okada, B. M., Novick, J. M., & Slevc, L. R. (2019). Updating musical tonal structure in working memory engages cognitive control. *Auditory Perception & Cognition*, *2*(1–2), 21–46.

Osterhout, L., Holcomb, P. J., & Swinney, D. A. (1994). Brain potentials elicited by garden-path sentences: Evidence of the application of verb information during parsing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*(4), 786–803.

Patel, A. D. (2003). Language, music, syntax and the brain. *Nature Neuroscience*, *6*(7), 674–681.

Patel, A. D. (2010). *Music, language, and the brain* (1st edition). Oxford University Press.

Patel, A. D., Gibson, E., Ratner, J., Besson, M., & Holcomb, P. J. (1998). Processing syntactic relations in language and music: An event-related potential study. *Journal of Cognitive Neuroscience*, *10*(6), 717–733.

Pearce, M. T., & Wiggins, G. A. (2012). Auditory expectation: The information dynamics of music perception and cognition. *Topics in Cognitive Science*, *4*(4), 625–652.

Peretz, I., & Coltheart, M. (2003). Modularity of music processing. *Nature Neuroscience*, *6*(7), 688–691.

Peretz, I., & Zatorre, R. J. (2005). Brain organization for music processing. *Annual Review of Psychology*, *56*(1), 89–114.

Pinker, S., & Jackendoff, R. (2005). The faculty of language: What's special about it? *Cognition*, 95(2), 201-236.

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

Rohrmeier, M. (2013). Musical expectancy: Bridging music theory, cognitive and computational approaches. *Zeitschrift Der Gesellschaft Für Musiktheorie [Journal of the German-Speaking Society of Music Theory]*, *10*(2), 343–371.

Rohrmeier, M. (2020). The syntax of Jazz harmony: Diatonic tonality, phrase structure, and form. *Music Theory and Analysis*, *7*(1), 1–63.

Rohrmeier, M., & Neuwirth, M. (2015). Towards a syntax of the classical cadence. In P. Bergé & M. Neuwirth (Eds.), *What is a cadence? Theoretical and analytical perspectives on cadences in the classical repertoire*, (pp. 287–335). Leuven University Press.

Rohrmeier, M., & Pearce, M. (2018). Musical syntax I: Theoretical perspectives. In R. Bader (Ed.), *Springer handbook of systematic musicology* (pp. 473–486). Berlin Heidelberg: Springer.

Ross, B. (2014). *Music, language and syntactic integration* [PhD Thesis]. University of Cambridge.

Schenker, H. (1935). *Der Freie Satz*. Universal Edition.

Schmalfeldt, J. (2017). *In the process of becoming: Analytic and philosophical perspectives on form in early nineteenth-century music*. Oxford University Press.

Sears, D. R. W., Spitzer, J., Caplin, W. E., & McAdams, S. (2020). Expecting the end: Continuous expectancy ratings for tonal cadences. *Psychology of Music*, *48*(3), 358–375.

Shepard, R. N., & Teghtsoonian, M. (1961). Retention of information under conditions approaching a steady state. *Journal of Experimental Psychology*, *62*(3), 302–309.

Sipser, M. (2012). *Introduction to the theory of computation* (3rd edition). Cengage Learning.

Slevc, L. R., & Okada, B. M. (2015). Processing structure in language and music: A case for shared reliance on cognitive control. *Psychonomic Bulletin & Review*, *22*(3), 637–652.

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

Sprouse, J., & Lau, E. F. (2013). Syntax and the brain. In M. den Dikken (Ed.), *Cambridge handbook of generative syntax* (pp. 971–1005). Cambridge University Press.

Steedman, M. (1984). A generative grammar for Jazz chord sequences. *Music Perception*, *2*(1), 52–77.

Temperley, D. (2001). *The cognition of basic musical structures*. MIT Press.

Temperley, D. (2004). Communicative pressure and the evolution of musical styles. *Music Perception*, *21*(3), 313–337.

Trueswell, J. C., Tanenhaus, M. K., & Garnsey, S. M. (1994). Semantic influences on parsing: Use of thematic role information in syntactic ambiguity resolution. *Journal of Memory and Language*, *33*(3), 285–318.

Vogelzang, M., Mills, A. C., Reitter, D., Van Rij, J., Hendriks, P., & Van Rijn, H. (2017). Toward cognitively constrained models of language processing: A review. *Frontiers in Communication*, *2*.