Approaches to studying prediction in music and language

A thesis submitted by

Allison Fogel

in partial fulfillment of the requirements for the degree of

Master of Science

in

Psychology

Tufts University

May 2017

Adviser: Aniruddh D. Patel

Abstract

Prediction or expectancy is thought to play an important role in both music and language processing. However, prediction is currently studied independently in the two domains, limiting research on relations between predictive mechanisms in music and language. In Study 1, I developed a melodic cloze probability task (modeled on the standard linguistic cloze probability task), in which listeners are presented with the beginning of a novel tonal melody and asked to sing the note they expect to continue the melody. Participants' responses varied in consistency across melodies with different underlying harmonic structures. In Study 2, a sentence comprehension paradigm was used to explore lexical prediction strength in individuals with and without musical training, as indexed by the amplitude of the ERP component known as the frontal positivity; no relationship was observed between this component and participants' degree of musical training.

Table of Contents

List of Figures

General Introduction

Recent years have seen growing interest in cognitive and neural relations between music and language. Although there are clear differences between the two— for example, language can convey specific semantic concepts and propositions in a way that instrumental music cannot (Slevc & Patel, 2011)— they share several features. For example, both language and music involve the generation and comprehension of complex, hierarchically-structured sequences made from discrete elements combined in principled ways (Koelsch, Rohrmeier, Torrecuso, & Jentschke, 2013; Patel, 2003), and both rely heavily on implicit learning during development (Barbara Tillmann, Bharucha, & Bigand, 2000).

While neuropsychology has provided clear cases of selective deficits in linguistic or musical processing following brain damage (e.g., Peretz, 1993), several neuroimaging studies of healthy individuals suggest overlap in the brain mechanisms involved in processing linguistic and musical structure. One early demonstration of this overlap came from event-related potential (ERP) research, which revealed that a component known as the P600 is observed in response to syntactically challenging or anomalous events in both domains (Patel, Gibson, Ratner, Besson, & Holcomb, 1998). Later research using MEG and fMRI provided further suggestions of neural overlap in structural processing, e.g., by implicating Broca's region in the processing tonal-harmonic structure (e.g., LaCroix, Diaz, & Rogalsky, 2015; Maess, Koelsch, Gunter, & Friederici, 2001; Musso et al., 2015; Tillmann, Janata, & Bharucha, 2003; though see Fedorenko, McDermott, Norman-Haignere, & Kanwisher, 2012). To resolve the apparent contradiction between evidence from neuropsychology and neuroimaging, Patel (2003) proposed the Shared Syntactic

Integration Resource Hypothesis (SSIRH). The SSIRH posits a distinction between domain-specific representations in long-term memory (e.g., stored knowledge of words and their syntactic features, and of chords and their harmonic features), which can be separately damaged, and shared neural resources which act upon these representations as part of structural processing. This "dual-system" model proposes that syntactic integration of incoming elements in language and music involves the interaction (via long-distance neural connections) of shared "resource networks" and domain-specific "representation networks" (see Patel (2012) for a detailed discussion, including relations between the SSIRH and Hagoort (2005)'s "memory, unification, and control" model of language processing).

The SSIRH predicted that simultaneous demands on linguistic and musical structural integration should produce interference. This prediction has been supported by behavioral and neural research (for a review, see Kunert & Slevc, 2015). For example, behavioral studies by Fedorenko, Patel, Casasanto, Winawer, & Gibson (2009) and Slevc, Rosenberg, & Patel (2009) have shown that it is particularly difficult for participants to process complex syntactic structures in both language and music simultaneously (see also Carrus, Pearce, & Bhattacharya, 2012; Hoch, Poulin-Charronnat, & Tillmann, 2011; though cf. Perruchet & Poulin-Charronnat, 2013). Additionally, Koelsch, Gunter, Wittfoth, & Sammler (2005) conducted an ERP study that observed an interaction between structural processing in language and music, as reflected by effects of music processing on the left anterior negativity (LAN, associated with processing syntax in language) and effects of language processing on the early right anterior negativity (ERAN, associated with processing musical syntax).

In addition to structural integration, it has been suggested that prediction may be another process that operates similarly in language and music (Koelsch, 2012a, 2012b; Patel, 2012). Prediction is increasingly thought to be a fundamental aspect of human cognition (Clark, 2013), and is a growing topic of research in psycholinguistics (Van Petten & Luka, 2012; see Kuperberg & Jaeger, 2016 for a recent review). It has become clear that we regularly use context to predict upcoming words when comprehending language (Altmann & Kamide, 1999; DeLong, Urbach, & Kutas, 2005; Tanenhaus et al., 2016; Wicha, Moreno, & Kutas, 2004). This has been demonstrated using event-related potentials (ERP), a brain measure with millisecond-level temporal resolution that allows one to study cognitive processing during language comprehension. Recent evidence from ERP research has suggested that prediction in language processing occurs at multiple distinguishable levels (e.g., syntactic, semantic, phonological) (Kuperberg & Jaeger, 2016; Pickering & Garrod, 2007).

Strong lexical predictions for a specific word occur when multiple types of information within a linguistic context constrain strongly for the semantic features, the syntactic properties, and the phonological form of a specific word. For example, the sentence *"The piano is out of ____"* leads to a strong expectation for the word "tune", so one can refer to this as a high lexical constraint sentence. It is well established that unexpected words following these contexts evoke a larger N400 ERP component (occurring 300-500 ms after the presentation of the final word) than expected words (Kutas & Hillyard, 1984, 1980; Kutas & Federmeier, 2011). Such unexpected words do not necessarily need to be anomalous to produce an N400: predictions can also be violated with words that are perfectly coherent and non-anomalous. For example, if the final word delivered in the above sentence is "place" (i.e., "*The piano is out of place*") this word still violates a lexical prediction for the highly expected word "tune". As in the previous example, the N400 elicited by "place" would be larger than that elicited by "tune," as it is less expected. Moreover, in recent ERP research, violations of specific lexical predictions with other plausible words have also been observed to elicit a late anteriorly distributed positive component. This late frontal positivity has been observed at various time points after the N400, often peaking around 500-900 ms after the presentation of a critical item (Federmeier, Wlotko, De Ochoa-Dewald, & Kutas, 2007; Van Petten & Luka, 2012). Importantly, unlike the N400, the late frontal positivity is not produced by words that follow non-constraining contexts, when comprehenders have no strong prediction for a particular word (e.g. "place" following the context, "After a while, the boy saw the...".

Predictions in language are not always at the level of specific lexical items: they can also be generated at the level of semantic-syntactic statistical contingencies that determine the structure of an event ('who does what to whom') (Kuperberg, 2013). For example, at a certain point in a sentence we might expect a certain syntactic category of word, like a noun-phrase, with certain coarse conceptual features, such as animacy. For example, in the sentence *"Mary went outside to talk to the ____ "* there is no strong indication of which word will come next, but it is clear that it must be an animate noun-phrase (Mary would likely not talk to an inanimate object like a truck). Violations of these semantic-syntactic structural predictions have been observed to elicit a different neural response from the anterior positivity discussed above, namely the P600 (a late posterior positivity, peaking from around 600 ms after onset of the violating word; see Kuperberg, 2007 for a review). This provides evidence that distinct neural signatures may be associated with violations of strong predictions at different representational levels (e.g., a late anterior positivity evoked by violations of strong lexical predictions, Federmeier et al., 2007; a late posterior positivity evoked by violations of strong semantic-syntactic predictions, Kuperberg, 2007; see Kuperberg, 2013 for discussion). The functional significance of these late positivites (both frontal and posterior) evoked by strong prediction violations remains unclear. One possibility, however, is that they reflect the neural consequences of suppressing the predicted (but not presented) information and adapting one's internal representation of context in order to generate more accurate predictions in the future (e.g. see Kuperberg, 2013; Kuperberg & Jaeger, 2016 for discussion).

Turning to music, expectation has long been a major theme of music cognition research. Meyer, 1956 first suggested a strong connection between the thwarting of musical expectations and the arousal of emotion in listeners. In recent years, theories of musical expectation have been brought into a modern cognitive science framework (e.g., Huron & Margulis, 2010; Huron, 2006; Margulis, 2005; Pearce, Ruiz, Kapasi, Wiggins, & Bhattacharya, 2010), and expectation has been studied empirically with both behavioral and neural methods (e.g., Steinbeis, Koelsch, & Sloboda, 2006). It is increasingly recognized that multiple sub-processes are involved in musical expectation (see Huron, 2006, for one theoretical treatment). Empirical research has shown that predictions are generated for multiple aspects of music, such as harmony, rhythm, timbre, and meter (Rohrmeier & Koelsch, 2012). Such expectations are thought to be automatically generated by enculturated listeners (Koelsch, 2012a; Koelsch, Gunter, Friederici, & Schroger, 2000).

Here, we focus on melodic prediction, and specifically on expectations for upcoming notes in monophonic (single-voice) melodies based on implicit knowledge of the melodic and harmonic structures of Western tonal music (Barbara Tillmann et al., 2000). For those interested in relations between predictive mechanisms in music and language, melodic expectancy provides an interesting analog to linguistic expectancy in sentence processing. Like sentences, monophonic melodies consist of a single series of events created by combining perceptually discrete elements in principled ways to create hierarchically structured sequences (Jackendoff & Lerdahl, 2006). Sentences and melodies have regularities at multiple levels, including local relations between neighboring elements and larger-scale patterns, e.g., due to underlying linguistic-grammatical or tonal structure.

Like in language, comprehenders may generate predictions on these discrete levels in music. However, studies of expectation in music have not made this distinction, instead discussing "expectation" as a single overarching concept . Here, we describe a new method of quantifying melodic expectation that can be used to design studies of melodic expectancy that are comparable to the methods that have been used in psycholinguistic research.

Individual differences

Individual differences in prediction tendencies across domains present another way to study the relationship between prediction in language and in music (Patel & Morgan, 2016). Individuals with musical training have been shown to have enhanced abilities in certain aspects of language processing. For example, it has been demonstrated that adults with musical training are better at recognizing speech prosody and emotion (Lima & Castro, 2011; W. F. Thompson, Schellenberg, & Husain, 2004; Zioga, Di Bernardi Luft, & Bhattacharya, 2016) and identifying speech in noise (Slater et al., 2015; Swaminathan et al., 2015). Studies where children are randomly assigned to participate in music or some control activity have also demonstrated enhancements in language processing that can be attributed to musical training (Chobert, Francois, Velay, & Besson, 2014; Moreno et al., 2009). Jentschke & Koelsch (2009) found that children with musical training had larger amplitudes of the ELAN, an ERP component associated with syntactic processing in language, suggesting that musical training may have strengthened this aspect of language processing.

Because prediction is vital for successful music processing, it is possible that musical training may be associated with a greater tendency to predict upcoming sequential information in general. Both music and language are implicitly acquired via statistical learning, the domain-general ability to extract regularities from the statistics of the environment (Saffran, Aslin, & Newport, 1996). Statistical learning is a core mechanism at work in learning music and language and vital to forming predictions in both domains (Dienes & Longuet-Higgins, 2004; Clement Francois & Schon, 2014; Loui, Wessel, & Kam, 2010; Rohrmeier & Rebuschat, 2012).

Adult musicians have been demonstrated to perform better than non-musicians at auditory statistical learning, learning the statistical probabilities of novel sequences faster or more accurately (Shook, Marian, Bartolotti, & Schroeder, 2013; Skoe, Krizman, Spitzer, & Kraus, 2013). Francois, Chobert, Besson, & Schon (2013) randomly assigned children to training in either music or painting and found that only the music group improved on a statistical learning task based on speech segmentation. Similarly, a number of studies have found increased neural indices of learning statistical regularities in adult musicians compared to non-musicians, even in the absence of behavioral differences; neural measures may be more sensitive to early learning effects (Francois, Jaillet, Takerkar, & Schon, 2014; Francois & Schön, 2011; Paraskevopoulos, Kuchenbuch, Herholz, & Pantev, 2012). While these studies have focused on auditory statistical learning, it is possible that this ability may transfer to other domains. Evidence of this effect is currently limited and mixed. Vasuki, Sharma, Demuth, & Arciuli (2016) found a musician advantage in auditory but not visual statistical learning. However, Vassena, Kochman, Latomme, & Verguts (2016) have demonstrated enhanced cross-modal prediction in musicians, who showed increased sensitivity to statistical structure in both auditory and visual modalities.

The musician advantage in language processing could also be attributable to changes in other aspects of cognition that in turn impact predictive tendencies in language. Cognitive control has been suggested as one possible domain-general mechanism involved in processing both music and language (Fedorenko, 2014; Slevc & Okada, 2015). Cognitive control may be involved in prediction in language as part of resolving conflicts between a predicted structure and conflicting input, and enhanced cognitive control has been demonstrated in musicians (Bialystok & Depape, 2009). Musicians have also been shown to have enhanced verbal short term memory (Hansen, Wallentin, & Vuust, 2013; Wallentin, Nielsen, Friis-Olivarius, Vuust, & Vuust, 2010) and verbal working memory (Clayton et al., 2016; Franklin et al., 2008; Zuk, Benjamin, Kenyon, & Gaab, 2014). Working memory may then impact language processing and prediction tendencies (Boudewyn, Long, & Swaab, 2013; Traxler, Williams, Blozis, & Morris, 2005; Van Petten, Weckerly, McIsaac, & Kutas, 1997).

While many links have been suggested between musicianship and language processing abilities, the impact of musical training on predictive tendencies in language has not been well explored. In Study 2, we used a paradigm that has previously been used to index individual differences in prediction in language to examine the relationship between musical training and prediction.

Study 1: Studying musical and linguistic prediction in comparable ways: the melodic cloze probability method[1]

Introduction

In order to study relations between the cognitive mechanisms of prediction in sentences and melodies, it is necessary to measure prediction in these two types of sequences in comparable ways. In sentence processing, one typical method of measuring lexical expectancy is the cloze probability task, in which participants are asked to complete a sentence fragment with the first word that comes to mind (Taylor, 1953). For a given context, the percentage of participants providing a given continuation is taken as the "cloze probability" of that response. The cloze probability of an item is therefore a straightforward measure of how expected or probable it is. In addition to measuring the cloze probability of a particular word in relation to its context, it is also possible to use the cloze task to measure the 'lexical constraint' of a particular context by calculating the proportion of participants who produce a given word (see Federmeier et al., 2007). For example, a sentence such as *"The day was breezy so the boy went outside to fly a …"* would likely elicit the highly-expected continuation *"kite"* from most participants, and thus be a 'strongly lexically constraining' context. In contrast, a sentence such as *"Carol always wished that she'd had a …"* would elicit a more varied set of responses, and thus be a 'weakly lexically constraining' context.

[1] Previously published as Fogel, Rosenberg, Lehman, Kuperberg, & Patel (2015)

While expectancy in music has been measured in various ways over the years, to date there has been nothing comparable to the cloze probability method in language, i.e., a production-based task in which a person is presented with the beginning of a short coherent sequence and then asked to produce the event she thinks comes next.[2] Most behavioral studies of expectancy in music have used perceptual paradigms, such as harmonic priming paradigms or ratings of how well a tone continues an initial melodic fragment. Harmonic priming paradigms consist of a prime context followed by a target event, in which the degree of tonal relatedness between the two is manipulated. Typically, harmonically related targets are processed faster and more accurately than unrelated targets (Tillmann, Poulin-Charronnat, & Bigand, 2013). These studies have shown that chords that are more harmonically related to the preceding context are easier to process, while there is a cost of processing chords that are less related or unrelated to the context (Tillmann, Janata, Birk, & Bharucha, 2003). Another genre of priming studies has shown that timbre identification is improved when a pitch is close in frequency to the preceding pitch and harmonically congruent with the preceding context (Margulis & Levine, 2006). In studies using explicit ratings of expectancy, listeners are asked to rate how well a target note continues a melodic opening, e.g., on a scale of 1 (very bad continuation) to 7 (very good continuation) (e.g., Schellenberg, 1996). More recently, a betting paradigm has been used in which participants place bets on a set of possible continuations for a musical passage, and bets can be distributed across multiple possible outcomes (Huron, 2006). The betting paradigm has the advantage of providing a measure of the *strength* of an

[2] Waters, Townsend, and Underwood (1998) used what they refer to as a "musical 'cloze' task," but theirs was a multiple-choice task where participants selected one of several pre-composed sections of musical notation.

expectation for a specific item. However, like the “continuation rating” task, this task requires post hoc judgments, and is therefore is not an online measure of participants’ real-time expectations. ERPs and measures of neural oscillatory activity can provide online measures of expectation in musical sequences (e.g., Fujioka, Ween, Jamali, Stuss, & Ross, 2012; Pearce et al., 2010), but such studies have focused on perception, not production.

A handful of studies have used production tasks to measure musical expectancy, but they differ in important ways from the standard linguistic cloze probability task. Some studies have used extremely short contexts, in which participants are asked to sing a continuation after hearing only a single two-note interval, or even a single note (Carlsen, 1981; Povel, 1996; Schellenberg, Adachi, Purdy, & McKinnon, 2002; Thompson, Cuddy, & Plaus, 1997; Unyk & Carlsen, 1987). (Lake, 1987) presented two-note intervals after establishing a tonal context consisting of major chords and a musical scale. However, no prior singing-based study of melodic expectation has used coherent melodies as the context (Some studies using piano performance have used very long contexts, in which pianists have been asked to improvise extended continuations for entire piano passages, Schmuckler, 1989, 1990). Also, in all of these studies (and unlike in the linguistic cloze probability task), participants were asked to produce continuations of whatever length they chose in response to brief stimuli. The closest analog to a musical cloze task comes from a study of implicit memory for melody, in which listeners first heard a set of novel tonal melodies and then heard melodic stems of several notes and were asked to “sing the note that they thought would come next musically” (Warker & Halpern, 2005). However, the structure of the melodic stems was not manipulated, and the focus of the study was on implicit memory, not on expectation.

In order to advance the comparative study of prediction in language and music, it is necessary to develop comparable methods for studying prediction in the two domains. To this end, we have developed a melodic cloze probability task. In this task, participants are played short melodic openings drawn from novel coherent tonal melodies, and are asked to sing a single-note continuation. In an attempt to manipulate the predictive constraint of the melodies, the underlying harmonic structure of each opening (henceforth, 'melodic stem') was designed to either lead to a strong expectancy for a particular note, or not (see Methods for details). For each melodic stem, the cloze probability of a given note is calculated as the percentage of participants producing that note. The predictive constraint of a melodic stem is determined by examining the degree of agreement between participants' responses. For example, if all participants sing the same note after a particular stem, the stem has 100% constraint. On the other hand, if the most commonly sung note is produced by 40% of the participants, then the stem has 40% constraint.

The melodic cloze probability method allows the cloze probabilities of notes to be quantitatively measured, and thus provides a novel way to study how different structural factors (e.g., local melodic interval patterns vs. larger-scale harmonic structure) interact in shaping melodic expectation. The method can also be used to test quantitative models of melodic expectation, such as Narmour's (1990) "Implication-Realization" model, using naturalistic musical materials. In the future, the method can facilitate the design of studies comparing predictive mechanisms in language and music, e.g., by systematically manipulating constraint and cloze probabilities across linguistic and musical stimuli in behavioral or ERP studies of expectancy (cf. Tillmann & Bigand, 2015).

Methods

Participants

50 participants (29 female, 21 male, age range 18-25 years, mean age 20.3 years) took part in the experiment and were included in the data analysis (eight further participants were excluded due to difficulties with singing on pitch; see "Data Analysis"). All participants were self-identified musicians with no hearing impairment who had a minimum of 5 years of musical experience within the past 10 years (playing an instrument, singing, or musical training); 22 (44%) reported "voice" as one of their instruments. Participants had received a mean of 9.0 years of formal musical training on Western musical instruments (SD = 4.8) and reported no significant exposure to non-Western music. Participants were compensated for their participation and provided informed consent in accordance with the procedures of the Institutional Review Board of Tufts University.

Materials

The stimuli consisted of 45 pairs of short novel tonal melodies created by Jason Rosenberg, a professional composer. Stimuli were truncated in the middle, creating "melodic stems." The melodies ranged across all 12 major keys and employed variety of meters (3/4, 4/4 and 6/8 time signatures). Each stem was 5-9 notes long (M = 8.38 notes, SD = 0.83), and was played at a tempo of 120 beats per minute (bpm). Note durations varied from eighth notes (250 ms) to half notes (1000 ms). Stems contained no rests, articulation indications, dynamic variability, or non-diatonic pitches. All stimuli were created using Finale software with sampled grand piano sounds. Across all melodies, the highest and lowest pitch were A5 (880.0 Hz) and D3 (146.8 Hz), respectively, and the mean pitch was near E4 (329.6 Hz). On average, stems had a pitch range of 11.4 semitones (distance between the highest and lowest pitch in the stem, *SD* = 3.2 st). Male participants heard the melodic stems transposed down one octave. The average stem duration was 5.02 seconds (*SD* = 1.23).

Each stimulus pair consisted of two stems in the same musical key: one was an “authentic cadence” (AC) version, which was designed to create a strong expectation for a particular note, and the other was a “non cadence” (NC) version, which was designed to *not* generate a strong expectation for a particular note. AC stems ended preceding a strong beat within the meter on the 2nd, 5th, or 7th scale degree and with an implied authentic cadence that would typically be expected to resolve to a tonic function. NC stems ended with an implied IV, iv, or ii harmony, with the last presented note never on the 2nd or 7th scale degree and rarely on the 5th. The two stems in each pair were identical in length, rhythm, and melodic contour; they differed only in the pitch of some of their notes, which influenced their underlying harmonic structure (see **Figure 1** for an example). On average, the two stems of an AC-NC melodic pair differed in 48.3% of their notes (*SD* = 28.5%). When notes of an AC-NC pair differed, they remained close in overall pitch height, on average 1.90 semitones apart (*SD* = 0.38).

Figure 1. (a) Authentic cadence (AC) and (b) Non cadence (NC) versions of one melodic pair (see text for explanation). The figure shows the AC and NC stems in Western music notation. Shown beneath each stem is a possible interpretation of the underlying implied harmonic progression (e.g., D, G, D, A chords in the AC stem), and harmonic functions (I, or tonic chord; IV, or subdominant chord; V, or dominant chord, vi or submediant chord). The stems of a pair are identical in length, key, rhythm, and melodic contour, and each consists of single stream of notes with no accompaniment. In this pair the stems differ only in the identity of the final two notes, which are slightly lower in the second stem. Crucially, this small physical change alters the underlying harmonic progression.

The extent to which the two groups of stems projected a sense of key was compared using the Krumhansl-Schmuckler key-finding algorithm (Krumhansl, 1990). This model is based on "key-profiles" of each potential key, which represent the stability of each pitch in the key, i.e., how well it fits in a tonal context (Krumhansl & Kessler, 1982). The pitch distribution of a given melody, weighted by duration, is compared to the key-profile of each key, and a correlation value is calculated. When correlations with the profiles of each potential key were calculated for each stem, the mean correlation with the correct key for AC stems ($r(22) = .70$) did not differ significantly from the mean correlation with the correct key of NC stems ($r(22) = .73$), $t(44) = 1.24$, $p = 0.22$ (averaging and statistics were performed on Fisher transformed correlation coefficients). The two groups of stems therefore did not differ in the degree to which they projected a sense of key.

Procedure

Stimuli were played to participants over Logitech Z200 computer speakers at a comfortable listening volume within a sound attenuated room. The experiment was presented using PsychoPy (v1.79.01) on a MacBook Pro laptop, and sung responses were recorded as .wav files using the computer's built-in microphone.

Each participant was instructed that s/he would hear the beginnings of some unfamiliar melodies and would need to "sing the note you think comes next." Participants were asked to *continue* the melody—not necessarily complete it—on the syllable "la." Each trial began when the participant pressed a button to hear a melodic stem. Immediately after the end of the last note of each stem, the word "Sing" appeared on the screen and participants were given five seconds to sing the continuation, after which they rated their confidence in their response on a 7-point Likert scale (1 = *low*, 7 = *high*).

Each participant was presented with 24 AC and 24 NC melodic stems (only one version from each AC-NC pair) in one of eight randomized presentation orders. (Three pairs were removed from analysis due to differences in the melodic contours of the two stems, hence data from 45 pairs was analyzed.) At the beginning of the experiment, each participant completed a pitch-matching task in which they heard and were asked to sing back a series of individual tones (F4, A4, B3, G#4, A#3, D4, C#4, and E♭4 [corresponding to 349.2, 440.0, 246.9, 415.3, 233.1, 293.7, 277.2, 311.1 Hz, respectively]; one octave lower for male participants). This was used to evaluate participants' singing accuracy. Before the experimental trials began, participants were familiarized with the experimental procedure with a block of practice items, which ranged from simple scales and familiar melodies to unfamiliar melodies.

Data Analysis

We extracted the mean fundamental frequency of the sung note using Praat (Boersma, 2001). The pitch of the sung note was determined by rounding the measured mean fundamental frequency to the closest semitone in the Western chromatic scale (e.g., A4 = 440 Hz), with the deviation from the frequency of this chromatic scale tone recorded (in cents, i.e., in hundredths of a semitone). The sung response was also represented in terms of its scale degree within the key of the stem in question. Responses were generalized across octaves for the purpose of this study. Participants' responses to the pitch-matching portion of the experiment were also analyzed; if any participant's pitch-matching responses did not round to the same note that was presented, or if their responses to at least 25% of the experimental trials were more than 40 cents away from the nearest semitone, the participant's responses were excluded from further analysis (8 participants were omitted for these reasons). Additionally, reaction times were measured using a sound onset measurement script in Praat (a sound's onset was detected when the sound reached a level -25 dB below its maximum intensity for a minimum of 50 ms) to determine how quickly the continuation was sung after the offset of the last note of the stem.

Results

Participants found the task intuitive and uncomplicated, suggesting that the melodic cloze probability task provides a naturalistic way to measure melodic expectations. On average, participants sang a continuation note with a reaction time of 899 ms (SD = 604 ms), and their sung notes were an average of 1896 ms long (SD = 808 ms). Given that that the melodies had a tempo of 120 BPM, this corresponds to an average time interval of 1.80 beats after the offset of the stem, and a sung note duration of 3.79 beats.

Constraint

The primary dependent variable in our study was the predictive constraint of a melodic stem, as measured by the percentage of participants that sang the most common note after the stem. Figure 2 illustrates how this was computed, based on the AC-NC melodic pair in Figure 1. Figure 2a and 2b show the distributions of sung notes after the AC and NC stems in Figure 1, respectively. Figure 2a shows that 92% of participants that heard the AC stem produced the most commonly sung note (the tonic, D), while Figure 2b shows that no more than 24% of participants that heard the NC stem produced any one note (in this case, there was a tie between C# and A, but in most cases, one pitch class was most common). Thus the constraint of this melodic pair was 92% (or 0.92) for the AC melody and 24% (or 0.24) for the NC melody. For this pair, the AC melody was indeed far more constraining than the NC melody, as predicted.

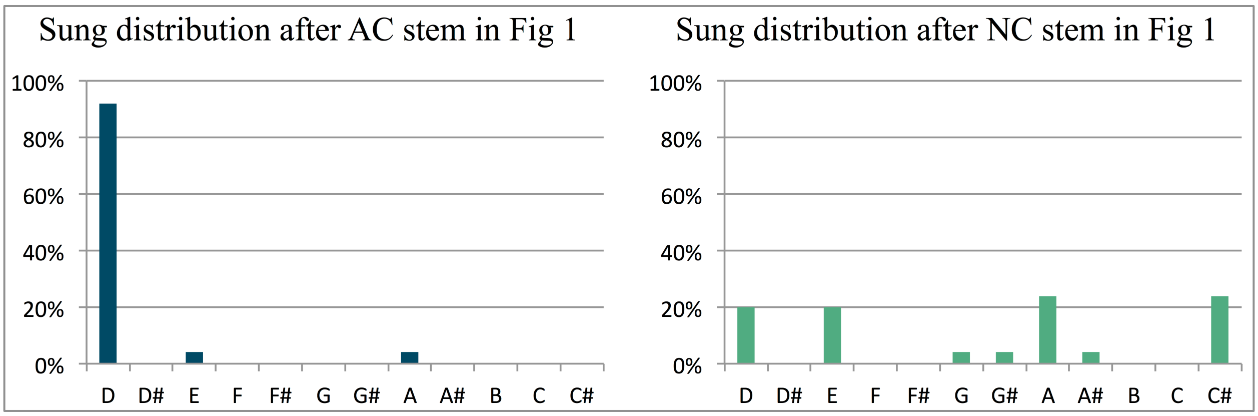


Figure 2. Histograms showing the relative frequency of different notes sung by participants at the end of the AC and NC stems in Figure 1. After the AC stem, most participants (92%) sang the pitch D, which is the 1st scale degree or tonic of the prevailing key of D major. After the NC stem, the note sung varied much more between participants: only 20% sang the pitch D, and no more than 24% of participants sang the same note (a tie between A and C# in this case).

For each AC and NC stem, we computed the constraint as described above. After AC stems, the average constraint was 69% (i.e., on average, 69% of participants sang the same note after hearing an AC stem), while after NC stems, the average constraint was 42% (i.e., on average, only 42% of participants sang the same note after hearing an NC stem). Thus on average, melodic stems in the AC condition did prove to be more constraining than NC stems (AC M = 0.692, SD = 0.171; NC M = 0.415, SD = 0.153), $t(44) = 7.79$, $p < .001$.). This pattern of higher constraint for the AC vs. NC stem was observed in 38 of the 45 item pairs (Figure 3).

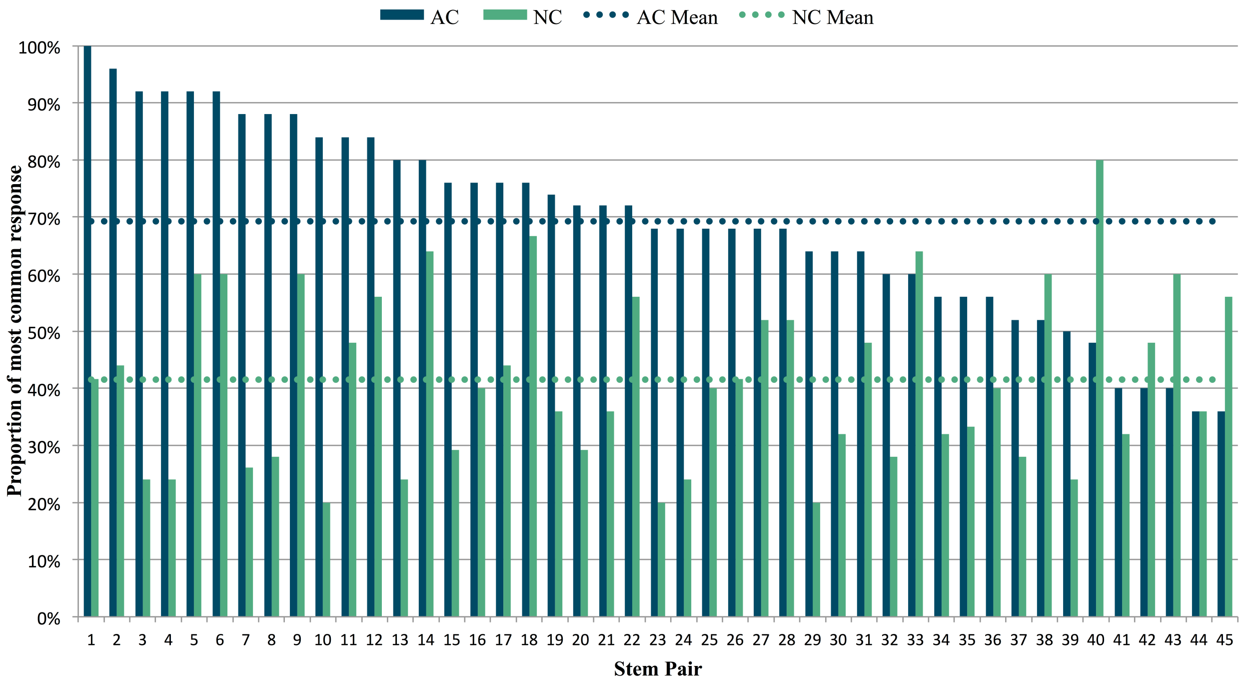


Figure 3. Constraint of AC and NC stems, as calculated by the percentage of participants providing the most common response for each stem. Stem pairs are ranked in order of decreasing constraint for AC stems. The melodic pair shown in Figure 1 corresponds to stem pair 3 in this graph. Dotted horizontal lines show the mean constraint across all AC and NC stems.

On average, participants responded significantly more quickly after AC stems (mean RT = 767 ms, *SD* = 265 ms) than after the NC stems (mean RT = 1033 ms, *SD* = 302 ms), $t(49) = 9.78$, $p < .001$. Additionally, on average participants were significantly more confident in their responses to AC stems ($M = 5.14$, $SD = 0.95$) than to NC stems ($M = 4.36$, $SD = 1.04$), $t(49) = 9.60$, $p < .001$.

Scale Degree

When responses were represented in terms of their scale degree in the key of the stem in question, and compiled across all items in each condition, the distributions for AC and NC items were strikingly different. For 6 of the 7 diatonic scale degrees, the frequency of response differed significantly between AC and NC items based on t-tests of each scale degree with a Bonferroni correction applied (see Figure 4 for p-values). For AC items, responses were heavily weighted around the first note of the scale, or tonic (known as 'do' in solfege). For NC items, responses were more widely distributed; however, they were mainly restricted to in-key diatonic scale degrees.

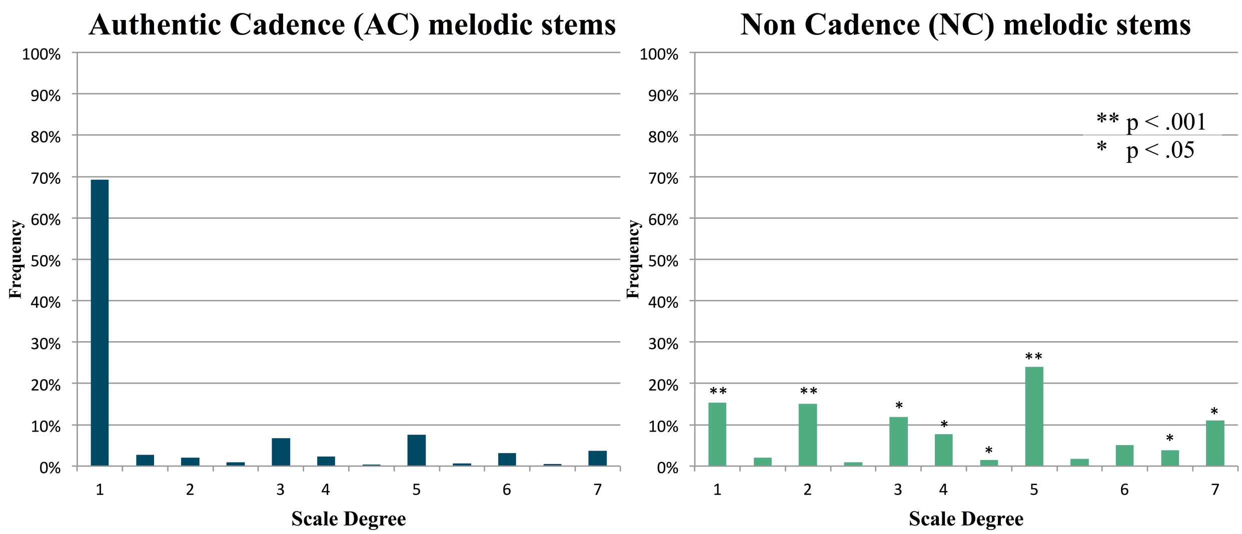


Figure 4. Average of all response distributions to AC and NC stems, shown as scale degrees. Numbers represent diatonic (major) scale degrees (e.g., 1 = tonic, 7 = leading tone, etc.), with asterisks indicating scale degrees with significantly different frequencies between the two conditions.

Figure 5. Examples of two melodic stem pairs with an authentic cadence (AC) stem that was much more constraining than the non-cadence (NC) stem. Stems are shown in black and white, and for each stem the most frequently sung note is shown as a red note head at the end of the stem. The pitch class name of this note and the proportion of listeners who sang the note (i.e., the measured melodic constraint of the stem) are printed next to the red note. These two pairs correspond to stem pairs 4 and 3 in Figure 3 (the stems in panel (b) are the same as in Figure 1).

Variability

While AC stems were on average significantly more constraining than their matched NC stems, there was considerable variability across AC-NC pairs in the degree of difference in constraint between members of a pair (see Figure 3). 38 out of 45 pairs demonstrated the expected pattern, with the AC stem proving more constraining than the NC stem. For instance, the stem pair in Figure 5a has a highly constraining AC stem, with 92% of participants singing the same note, the melody's tonic pitch, C (in Figures 5 and 6, the most commonly sung note is shown as a red note head after the end of each stem). Why might this be? This stem is short, contains only one rhythmic value, and has very clear harmonic implications, beginning with an unambiguously arpeggiated tonic triad (C-E-G) and concluding with a similarly outlined complete dominant triad (G-B-D). This stem also ends on the leading tone of B, i.e., the seventh scale degree of the diatonic major scale, which customarily resolves to the tonic scale degree, particularly near the end of a phrase. Further structural factors that may contribute to the high degree of agreement on the final pitch are 1) the melody's consistent downwards contour, which seems to close in on middle C, and 2) the fact that the tonic note is heard very close to the end of the phrase, which may make it more likely to be replicated. Turning to the NC stem in Figure 5b, it is similar in many respects to the AC stem, yet very different in constraint, with the most commonly sung note (F) being produced by just 24% of participants who heard this stem. What might account for this? The NC stem does not have any resolution-demanding dominant pitches at its conclusion, and as a result lacks a clear sense of harmonic direction. Instead, the melody follows a downwards pattern of melodic thirds (E-G, C-E, A-C) whose continuation is ambiguous. The most commonly chosen completion of F could be explained as the next logical pitch in the chain of descending thirds, after A-C. Thus, when faced with a stem where harmonic direction is underdetermined, subjects may have recruited an alternative strategy of melodic pattern continuation.

Another example of an AC stem that proved to be highly constraining is shown in Figure 5b (same melodic pair as in Figure 1). As with the melody in Figure 5a, the AC stem begins on the tonic note and returns to it as the most expected continuation, with an overall melodic range that emphasizes the octave generated above the first scale degree. The melody's interior arpeggiates two chords, first the tonic (D-F#-A) in measure 1, then the subdominant (G-B-D) in measure 2. The subdominant chord frequently serves a syntactic role of "predominant," a harmonic function that signals the initiation of a cadence. This is indeed how measure 3 is structured, with a heavily implied dominant harmony via scale degrees 2 and 7, and a melodic contour that insures D as a plausible completion due to an implied F#-E-D melodic descent and a unresolved leading tone of C#. The less constraining NC stem in Figure 5b, by contrast, ends on the sixth scale degree (the submediant). Unlike the leading tone, this note lacks a strong tendency to resolve in a particular way. It may plausibly serve as part of a stepwise motion to or away from the dominant, or as part of an arpeggiation of a predominant harmony; in either case, it negates the cadential function of the third measure and points to no obvious melodic completion.

Figure 6. Examples of two melodic stem pairs with an authentic cadence (AC) stem that was less constraining than the non-cadence (NC) stem. Stems are shown in black and white, and for each stem the most frequently sung note is shown as a red note head at the end of the stem. The pitch class name of this note and the proportion of listeners who sang the note (i.e., the measured melodic constraint of the stem) are printed next to the red note. These two pairs correspond to stem pairs 40 and 45 in Figure 3.

Contrasting with these stems, where subjects' responses to stems adhered to the AC/NC designations, there were several items where the constraint of the NC stem unexpectedly *exceeded* that of the AC stem. For example, after the NC stem in Figure 6a, 80% of participants sang the same note (F#, the 5th scale degree). In this particular melody, we believe this reflects the tendency for a large melodic interval to be followed by stepwise motion in the opposite direction. This "gap-fill" pattern (Meyer, 1956; Narmour, 1990) likely strongly influenced the continuation most participants chose, which involved singing a note (F#) one step down from the last note of the stem (G#), following a large leap of a sixth to an already contextually unstable note (scale degree six). Additionally, this stem has a strongly implied compound melody, wherein most of the topmost notes form a rising, stepwise pattern of B-C#-D#-E, which leads to an F# if this pattern is continued. Meanwhile, the unexpectedly low constraint of the AC stem in Figure 6a was perhaps due to the lack of a strong tendency note (like the leading tone) as its last pitch, and the obscuring of the underlying harmonic implications by the relative rhythmic complexity of the melody. That is, the unpredictable and syncopated rhythm may have reduced the strength of the expectancy for the tonic scale degree (Schmuckler & Boltz, 1994). Similarly, in the stem pair in Figure 6b, the most common continuation for the NC stem was a gap-filling motion to fill the exceptionally wide upward leap of an octave from Bb4 to Bb5. Landing on Ab, which 56% of subjects agreed on, helps close that gap with a downwards step and continues the melody on the more stable pitch of scale degree 5. This note also has the advantage of mirroring the first note of the melody, thus promoting melodic symmetry. The AC stem of this melodic pair presented no such clearly determined ending. If subjects opted to fill in the large upwards octave gap to Ab with a downwards step, they would land on the unstable fourth scale degree (Gb). On the other hand, if they were to resolve the melody with a cadence on the tonic note (Db), they would land far from the final note of the stem, going against a general tendency in melodic expectation for pitches that are proximate in frequency to the previous note (see section on modeling below).

Based on the above observations, it is clear that underlying harmonic structure, which was manipulated in the AC vs. NC stems, does not alone determine melodic expectation. Melodic factors that likely contributed to increased constraint in our melodies include (but are not limited to) rhythmic simplicity, gap-fill pattern, compound-line implication, leading-tone resolution, and pattern completion. In this way, stems in which linear, contrapuntal, rhythmic and harmonic parameters were closely coordinated produced reliable agreement on melodic completions, while examples with a conflict or ambiguity between those factors were prone to considerably less consensus.

Musical Experience

Prior research suggests that musical training enhances sensitivity to underlying harmonic structure (Koelsch, Schmidt, & Kansok, 2002). Since implicit harmony was used to guide the listeners' expectation for a tonic note after AC stems, we sought to determine if participants with greater degrees of musical training were more likely to sing the tonic after AC stems. Thus across AC stems, we correlated each participant's total years of formal musical training with their frequency of responding with the tonic. (Thus for example, if a participant sang the tonic after half of the AC stems they heard, their frequency of responding with the tonic to an AC stem would be 0.5) When all AC items were included in the analysis, there was no significant correlation with years of formal musical training, $r(48) = 0.035$, $p = 0.812$. However, when we divided AC stems according to the scale degree of their final note, an interesting pattern emerged. On average, after AC stems that ended on the 7th scale degree, participants sang the tonic 81% of the time, and in these melodies, there was a significant correlation between participants' years of formal training and their frequency of responding with the tonic, $r(48) = .45$, $p = .001$ (see Figure 7). This relationship with musical training was also observed with AC stems that ended on the 5th scale degree, where participants sang the tonic 55% of the time on average, $r(48) = .33$, $p = .02$. (The relationship was not seen for AC stems that ended on the 2nd scale degree, where participants sang the tonic 57% of the time on average.)

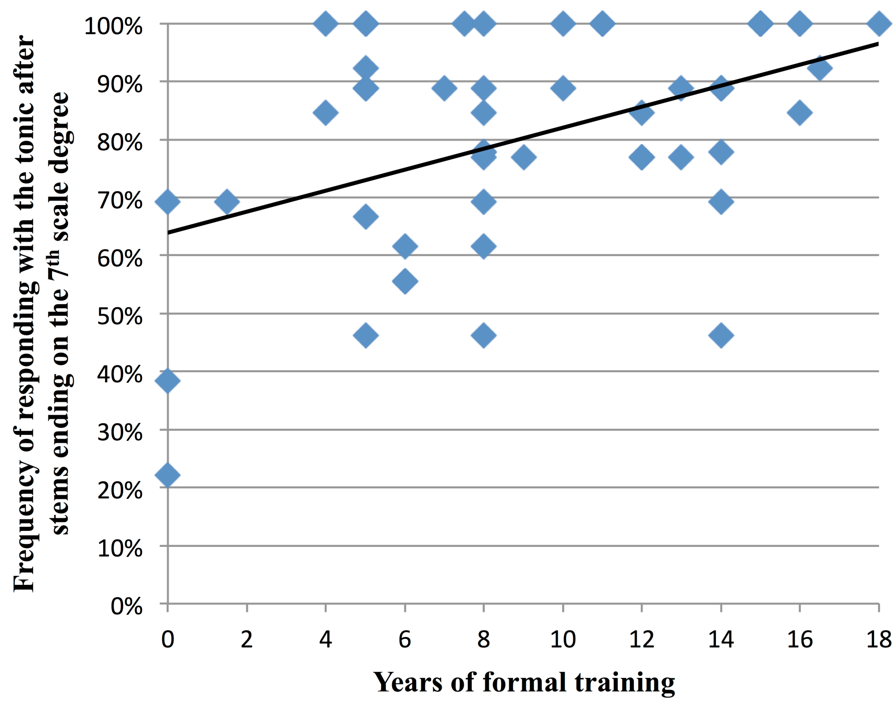


Figure 7. Relationship between participants' years of formal musical training and how often they sung the tonic after melodic stems that ended on the 7th scale degree. On average, participants sang the tonic 81% of the time after these stems (data for all 50 participants are shown: due to some data points lying directly on top of each other, fewer than 50 data points are visible on the graph).

Discussion

We introduce the melodic cloze probability task, in which participants hear the opening of a short, novel tonal melody and sing the note they expect to come next. This task, which is modeled on the well-known cloze probability task in psycholinguistics, has not previously been used to study expectancy in the field of music cognition. Participants found the melodic cloze task easy to do, demonstrating that expectancy can be measured in a comparable way across linguistic and musical domains.

Prior work using singing to study melodic expectancy has focused on responses to two-note intervals (see introduction for references). Of these studies, the closest task to ours is Lake (1987), who had participants sing extended continuations in response to a two-note interval preceded by a tonal context. Unlike the current study, the tonal context was not the opening of a novel coherent melody, but a sequence of notes consisting of a major chord, a scale, and another major chord, which served to establish a strong sense of key before the two-note interval. One might ask how our results compare to those of Lake, since one can conceive of our stimuli as also consisting of a key-inducing context followed by a final two-tone interval (i.e., the final two tones of the melodic stem).

While the last two notes of our stems clearly contribute to our results, our findings cannot be attributed to only hearing this final interval in a generic tonal context. A number of our stems are identical in the scale degrees of their final two notes, yet they elicit very different patterns of results from participants (see Figure 8 for an example). This different pattern of responding to the same final interval reflects differences in the *structure* of the preceding notes. Thus our paradigm and results are not simply a replication of Lake (1987), and show the relevance of using melodically coherent materials as contexts for production-based studies of melodic expectation. Similarly, we note that our results are not simply a replication of the well-known probe-tone results of Krumhansl & Kessler (1982), since the pattern of responding was not just a reflection of the tonal hierarchy, and depended on the structure of the heard melody (e.g. Figures 1 and 2).

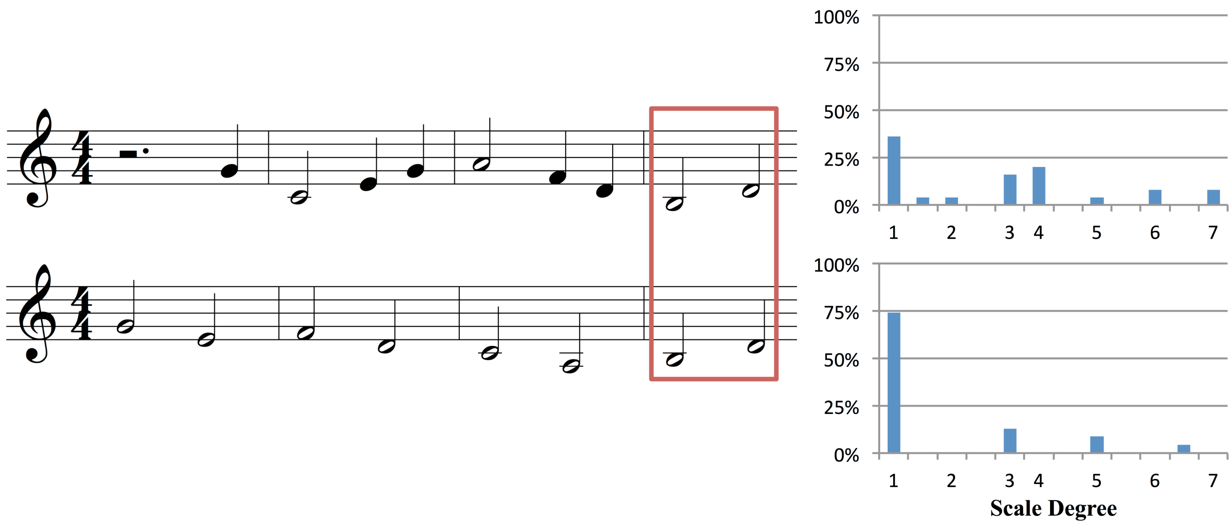


Figure 8. Example of stems that have the same final two notes but elicit different patterns of responses from participants. Stems have been transposed from their original keys to C major in order to facilitate comparison. Both stems end with scale degrees 7 and 2. The distribution of sung responses (expressed as scale degrees) is shown to the right of each stem.

In addition to being the first study to obtain cloze probabilities for musical notes, to our knowledge the current study is the also the first to manipulate the predictive *constraint* of musical sequences as part of research on melodic expectation. By using pairs of monophonic melodic openings (or ‘stems’) matched in length, rhythm, and melodic contour, but differing in implied harmonic structure, we show that underlying harmonic progressions can strongly guide melodic expectations. Specifically, there was significantly more consistency in participants’ responses to melodic stems ending on an implied authentic cadence (AC condition) than in their responses to stems ending non-cadentially (NC condition), as reflected by a higher percentage of participants singing the most common continuation for items in the AC condition. In other words, AC stems were more highly constraining than NC stems on average.

However, our data also clearly indicate that expectations based on larger-scale implied harmony interact with expectations based on melodic structure. That is, despite the fact that the harmonic differences between the AC and NC melodies in each pair were similar, we observed considerable variability in the constraint of melodies. In some pairs, the AC stem was considerably more constraining than the NC stem, but in other pairs the difference in constraint was mild, and in seven pairs the NC stem was actually equal to or more constraining than the AC stem (Figure 3). Analysis of two such 'reversed constraint' pairs (Figure 6) suggested that factors related to rhythmic simplicity, gap-fill pattern, compound line implication, and pattern completion may have been involved in overwhelming harmonic expectations. Further investigation of the factors driving the observed large variation in constraint among melodies is clearly warranted. From our results it is clear that expectancies related to melodic patterns (e.g., gap-fill) may sometimes trump those related to tonality.

Indeed, the variability in constraint observed in our data (Figure 3) suggests that the melodic cloze task is well suited for use in future studies aimed at exploring the relative contributions of melodic and harmonic patterns in shaping melodic expectation. Such studies can help test and improve quantitative models of melodic expectation (e.g., Eerola & Toiviainen, 2004; Krumhansl, Louhivuori, Toiviainen, & Eerola, 1999; Margulis, 2005; Pearce, 2005; Pearce & Wiggins, 2006; Schellenberg, 1996, 1997).

The musical cloze probability task has further uses in the field of music cognition. For example, this paradigm can be used to investigate how different factors influence melodic expectancy. While we manipulated only the harmonic structure of melodies in the present experiment, the influence of any other factor (e.g., melodic contour, rhythm, dynamics, etc.) on musical expectations could be explored in subsequent studies by composing melodies in pairs and manipulating the one factor while keeping other factors constant. Additionally, the task could be varied to have participants sing multiple-note continuations, as has been done in previous studies (Carlsen, 1981; Lake, 1987; Glenn Schellenberg et al., 2002; Thompson et al., 1997; Unyk & Carlsen, 1987). This would allow responses to be examined on longer timescales than just the first sung note. In addition, it would reduce the possibility that participants are responding by *completing* the melodic sequences with the sung note, instead of *continuing* them (as instructed). This is an important issue, as the note sung after the stem may differ depending on whether listeners treat it as a continuation or a completion (Aarden, 2003, cf. Huron, 2006).

Of course, the melodic cloze paradigm does have its limitations. By focusing on what pitch a person sings, it cannot give independent measures of all the different types of expectations which may be at play at a given point in a melody, such as timbral expectations (if listening to complex textures) or rhythmic expectations. To study these sorts of expectations, modifications of the paradigm presented here would be necessary. For example, if studying rhythmic expectations, at the end of each stem one could ask participants to press a bar for as long as they thing the next note will last.

The melodic cloze task can also be used to examine musical expectations in different populations. We observed a significant correlation between formal musical training and a tendency to sing the tonic after AC stems that ended on the 7th or 5th scale degrees. It has been suggested that having more musical experience leads to greater sensitivity to harmonic cues, which is consistent with our finding and with neural research on harmonic processing (Koelsch et al., 2002). Future studies could use the melodic cloze method to investigate how different kinds of musical experience might impact expectancy formation. For example, expectations may differ between musicians who have been educated in music theory vs. those who have experience singing or improvising without reading music. Additionally, the melodic cloze paradigm could be used in studies with children, to investigate how melodic expectations develop (cf. Corrigall & Trainor, 2014).

Obtaining melodic cloze probabilities is crucial for future research comparing predictive processing in music and language, as it allows for the comparison of the effects of violating predictions of comparable strength in the two domains (cf. Tillmann & Bigand, 2015). Previous studies comparing expectancy violations in music and language have typically chosen violations that are intuitively thought to be comparable in the two domains. By using a cloze paradigm to quantify cloze probabilities for possible continuations in both domains, it is possible to compare effects of violations of the same degree, using normed stimuli (cf. Featherstone, Waterman, & Morrison, 2012). For example, this will allow comparison of brain responses to plausible violations of expectations, instead of to frank structural violations (which rarely occur in naturalistic sequences). Also, studies that probe interactions between simultaneously presented music and language expectancy violations can be more precisely calibrated, in order to further elucidate cognitive and neural relations between language and music processing.

Study 2: Comparing linguistic prediction in musicians and non-musicians

Introduction

Given the importance of prediction in music processing, it has been suggested that musical training may be associated with a greater tendency to predict upcoming sequential information in general. In Study 1, we found that musical training was associated with a higher tendency to form predictions for notes based on hierarchical structure, as measured by participants' sung responses. As discussed above, it is possible musical training may be associated with prediction in language as well.

In psycholinguistic research, ERPs have recently been used to quantify individual differences in prediction. Wlotko, Federmeier, & Kutas (2012) examined individual differences in predictive tendencies between older and younger adults as indexed by the amplitude of the late anterior positivity, a component that has been observed after the N400, often peaking around 500-900 ms after the presentation of a critical item (Federmeier et al., 2007; Van Petten & Luka, 2012). This component is elicited by violations of predictions for *specific* lexical items; unexpected but plausible words in constraining contexts elicit an increased late frontal positivity compared to cloze matched unexpected words in nonconstraining contexts. Wlotko et al. (2012) found that younger adults showed a greater tendency to predict; most younger adults showed a late frontal positivity, while only a minority of older adults showed this effect.

Here, we used a similar sentence comprehension paradigm to explore language prediction strength in individuals with and without musical training, using a design in which target nouns fulfilled or violated contextual predictions for specific lexical items or verb-argument event structure.

Methods

Participants

Participants were recruited from online postings. Thirty-three volunteers (19 men, 14 women) participated in the ERP experiment. All were right-handed native English speakers (with no exposure to any other language before age 5) between ages 18 and 32 (M = 21.8, SD = 3.8). Participants had normal or corrected-to-normal vision and reported no history of psychiatric or neurological disorders. They provided informed consent in accordance with the procedures of the Institutional Review Board of Tufts University and were compensated for their time.

Twenty-one of these participants indicated that they had taken part in some amount of formal musical training (such as private lessons on an instrument) in their lifetimes, and 15 had at least five years of musical training. Sixteen participants reported having at least five years of musical experience within the past 10 years. Three participants majored in music, but none were professional musicians.

Materials

Verbs that were not highly predictive of any particular upcoming direct object (as measured by cloze probability) were selected. For each verb, two scenarios were created. One was a high constraint context, designed to create an expectation for a particular upcoming word (for example, the constraining context in Figure 9 constrains for the specific word "*swimmers*"). The other was a low constraint context that did not lead to an expectation for any particular word (the non-constraining context in Figure 9 could be completed by multiple possible words). Each scenario consisted of two context sentences (matched for length across constraining and non-constraining contexts) and a third sentence containing the critical word. The third sentence was identical across constraining and non-constraining scenarios aside of the critical word itself, which was always four words from the end of the sentence (e.g., *"Hence, they cautioned the ____ to be wary."* for the example item in Figure 9). Cloze probability norming was conducted for these scenarios; each participant saw only one of the two contexts for a given verb. For the set of items included in the experiment, high constraint contexts had a mean lexical constraint of 79% (*SD* = 10.1%) and low constraint contexts had a mean lexical constraint of 26% (*SD* = 10.1%).

Critical words were counterbalanced such that the same critical words appeared in both constraining and non-constraining contexts (see examples in Figure 9). In lexically constraining contexts, critical words were: (1) lexically predictable (the "right" word predicted from the preceding context), (2) lexical prediction violations (fully plausible in the scenario, but violating a prediction for the "right" word), or (3) violations of the semantic-thematic constraints of the preceding context, operationalized here as inanimate words in contexts that predict animate continuations or vice versa. In lexically non-constraining contexts, critical words were were: (4) lexically unpredictable (coherent continuations that are not particularly predicted, as non-constraining contexts do not create any particular expectation) or (5) semantic-thematic (i.e., animacy) violations. Critical words in conditions 2-5 all had zero cloze probability and were also matched on semantic relatedness to their preceding contexts by Latent Semantic Analysis; all were equally unpredictable from the prior context.

Each participant saw 100 critical items (20 in each condition, counterbalanced across participants) as well as 60 filler trials (20 constraining, 40 non-constraining).

Lexically constraining contexts (average constraint: 79%)		**Lexically non-constraining contexts** (average constraint: 26%)	
The lifeguards received a report of sharks right near the beach. Their immediate concern was to prevent any incidents in the sea. Hence, they cautioned the…		*Eric and Grant received the news late in the day. They decided it was better to act sooner than later. Hence, they cautioned the…*	
swimmers	(1) lexically predictable		
trainees	(2) lexical prediction violation	*trainees*	(4) lexically unpredictable (non-violation)
drawer	(3) lexical + animacy violation	*drawer*	(5) animacy violation

Figure 9. Sample stimuli for Experiment 2.

Nineteen participants also took part in a separate cognitive testing session. This session included measures of working memory (Reading Span, Listening Span, Subtract 2 Span, and Operation Span; Unsworth, Heitz, Schrock, & Engle, 2005; Daneman & Carpenter, 1980) as well as a Visual Statistical Learning Task (Fiser & Aslin, 2002). The Visual Statistical Learning (VSL) task requires participants to learn and remember implicitly grouped triplets of static images. Participants are first passively exposed to twelve images, displayed in groups of three. During the test phase of the task, each triplet of images is presented with an "impossible" foil triplet that was never presented during the exposure phase, and participants must identify which of the pair of triplets is more familiar (see Figure 10 for an example).

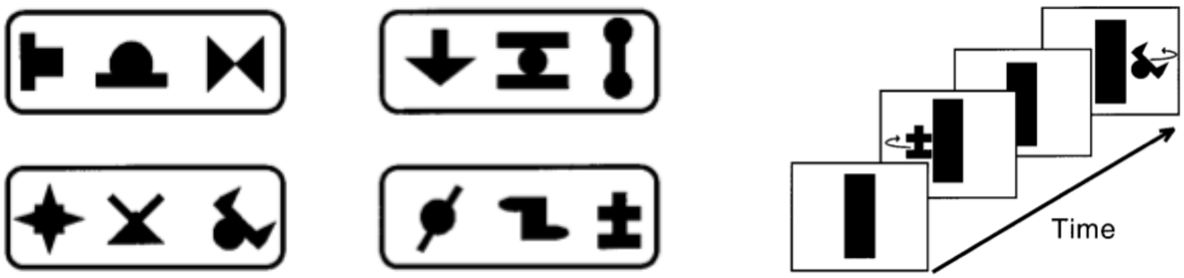


Figure 10. Example of the exposure and test portions of the Visual Statistical Learning Task (Fiser & Aslin, 2002).

Procedure

Trials were self-paced; participants pressed a button to begin each trial. The first sentence of the trial appeared in full, then the participant pressed a button to advance to the second sentence. When the participant pressed to advance to the third (experimental) sentence, each word was displayed individually for 450 ms with an ISI of 100 ms. 32 of the 60 filler trials were followed by yes-or-no comprehension questions to ensure that participants were attending to the scenarios.

ERP acquisition and processing

Data was collected using a BioSemi ActiveTwo EEG system and ActiView v7.05 EEG acquisition software. EEG was recorded from 32 Ag/AgCl electrodes in an elastic cap placed according to the international 10-20 system (see Figure 11 for electrode placement). Electrodes on each mastoid were recorded to serve as the reference, and electrodes below the left eye and beside the right eye were recorded to monitor for blinks and eye movements.

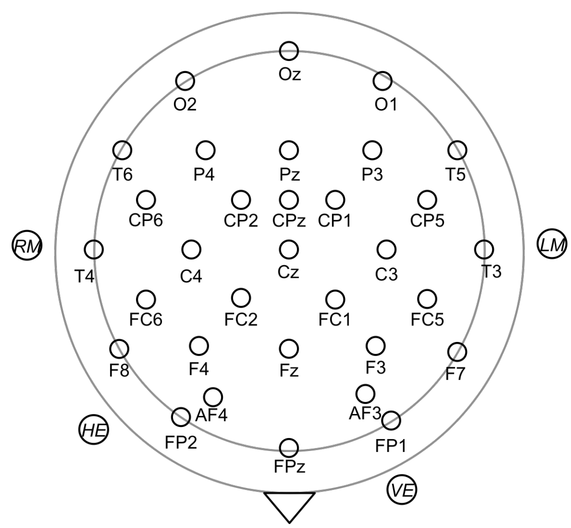


Figure 11. Electrode placement.

The EEG signal was amplified and continuously sampled at 512 Hz and digitally filtered online with a 5th order sinc response filter with a half-amplitude cutoff at 102.4 Hz.

EEG and ERP data processing was conducted in EEGLAB v13.5.4 (sccn.ucsd.edu/eeglab; Delorme & Makeig, 2004) and ERPLAB v5.0.0.0 (erpinfo.org/erplab; Lopez-Calderon & Luck, 2014). The EEG was referenced to the left mastoid electrode, the DC offset was removed by subtracting the average voltage of the entire segment, and a high-pass 2nd-order Butterworth infinite impulse response filter with a half-amplitude cut-off of 0.01 Hz was applied. Algorithms in ERPLAB were used to detect blinks, eye movements, and other artifacts. Averages were calculated for each condition from trials free of artifact (14.1% of trials were removed) after subtraction of the 200 ms prestimulus baseline.

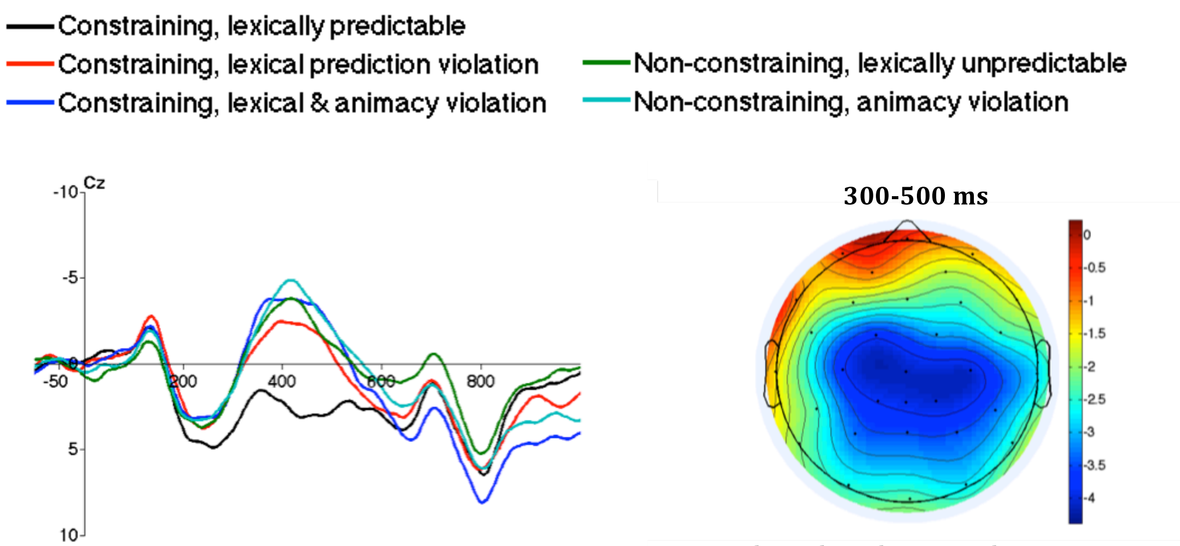
Results

ERPs

In line with prior studies (Federmeier et al., 2007), the N400 was measured as the mean amplitude between 300 and 500 ms at nine central-posterior electrode sites (C3, CP1, P3, Pz, P4, CP2, CPz, Cz, C4). As seen in Figure 12, the N400 was reduced for lexically predictable items compared to all other (cloze-matched) conditions, regardless of lexical constraint; within constraining contexts, the differences in amplitude between lexically predictable items (1) and lexically unexpected items (2) and between lexically predictable items (1) and animacy violations (3) were both significant, $t(32) = 6.20$, $p < .001$; $t(32) = 8.29$, $p < .001$.

The late anterior positivity was measured as the mean amplitude between 600 and 900 ms at ten anterior electrode sites (FP1, FPz, AF3, F7, F3, Fz, F4, F8, AF4, FP2). In this time window, lexical prediction violations (in constraining contexts) elicited a larger positivity than all other items (Figure 13). Lexical prediction violations (2) (in constraining contexts) differed significantly from items that were lexically unpredictable (4) (in non-constraining contexts), $t(32) = 3.21$, $p = .003$.

The P600 was defined as the mean amplitude between 600 and 900 ms at six posterior electrode sites (P3, Pz, O1, Oz, O2, P4). Animacy violations elicited a larger P600 than the remaining conditions, with animacy violations (3) differing significantly from lexically predictable items (1) in constraining contexts, $t(32) = 6.27$, $p < .001$ (see Figure 14).



lexical prediction violations
minus lexically predictable

Figure 12. N400: the N400 is reduced in response to lexically predictable items compared to all other conditions. The central electrode Cz is shown as an example. Voltage map shows the N400 effect as the difference between the constraining, lexical prediction violation and constraining, lexically predictable conditions from 300-500 ms.

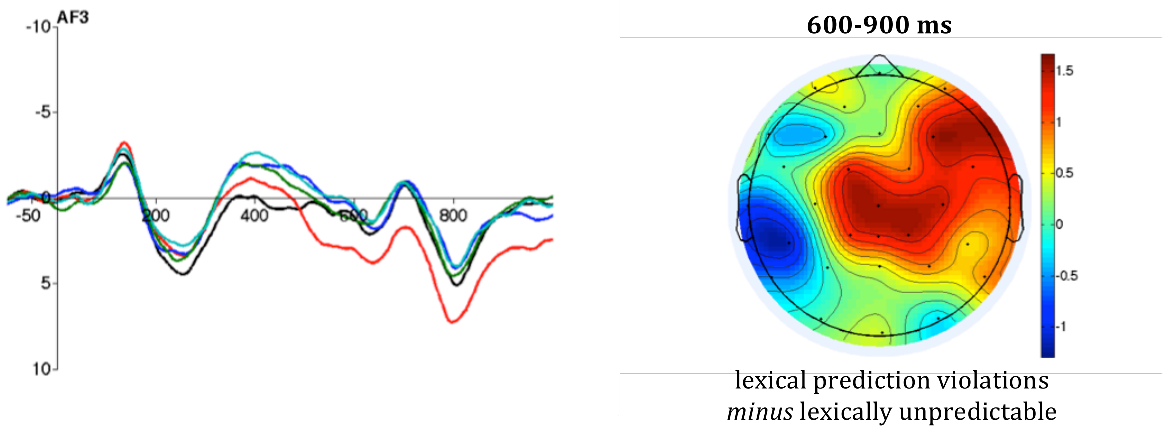


Figure 13. Late anterior positivity. The late anterior positivity is elicited by lexical prediction violations compared to other conditions. The frontal electrode AF3 is shown as an example. Voltage map shows the frontal positivity effect as the difference between the constraining, lexical prediction violation and non-constraining, lexically unpredictable conditions from 600-900 ms.

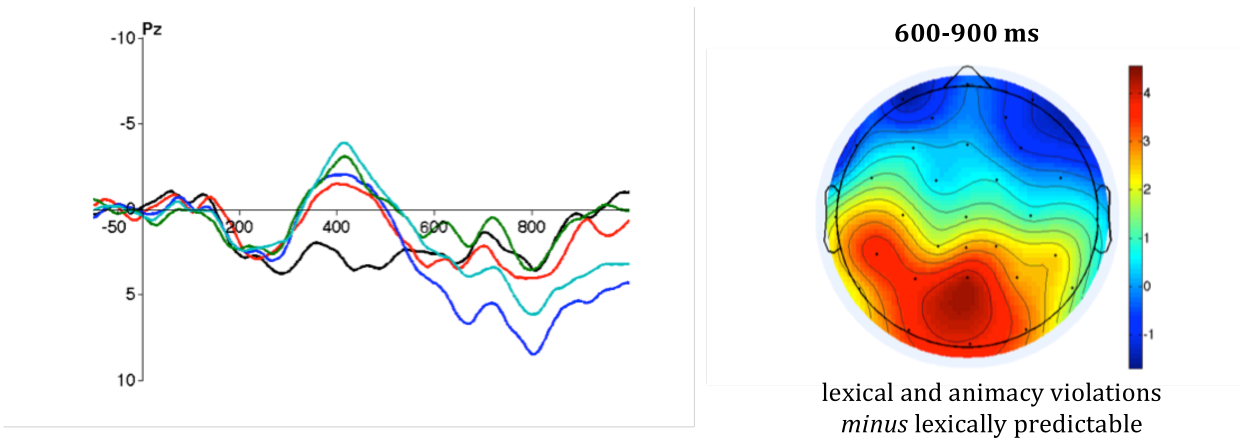


Figure 14. P600. The P600 is elicited by animacy violations as compared to other conditions. The posterior electrode Pz is shown as an example. Voltage map shows the frontal positivity effect as the difference between the constraining, lexical and animacy violation and constraining, lexically predictable conditions from 600-900 ms.

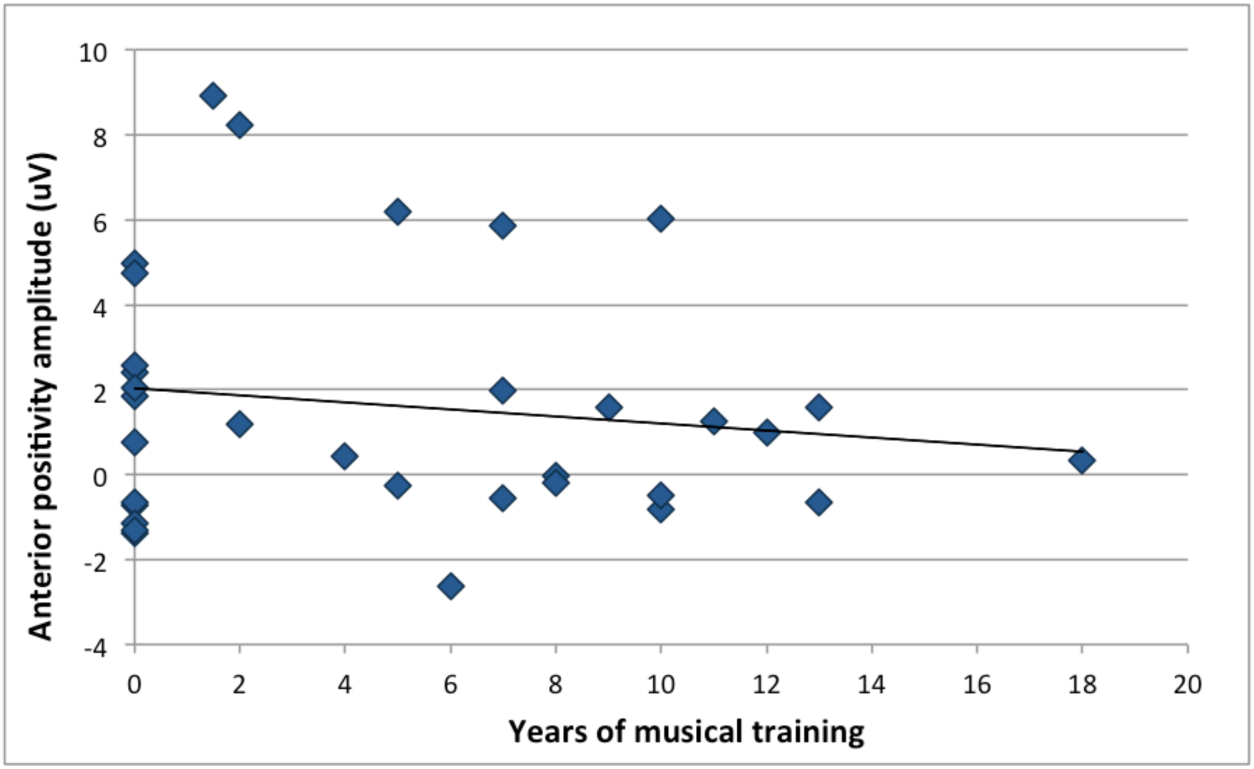


Figure 15. The relationship between the mean amplitude of the frontal positivity and years of musical training. No significant relationship was observed.

Individual differences

Contrary to our predictions, no relationship was observed between the amplitude of the late anterior positivity and participants' years of formal musical training, r = -.15, p = .41 (Figure 15). Similarly, there was no relationship between years of formal training and the amplitude of the N400 (as measured as the difference between lexically expected and unexpected items in constraining contexts) or P600, r = .19, p = .29; r = -.04, p = .83.

There were no significant relationships observed between participants' years of formal musical training and any of the working memory measures (all *ps* > .30). However, there was a significant relationship between participants' years of musical training and visual statistical learning performance, r = .500, p = .035 (Figure 16).

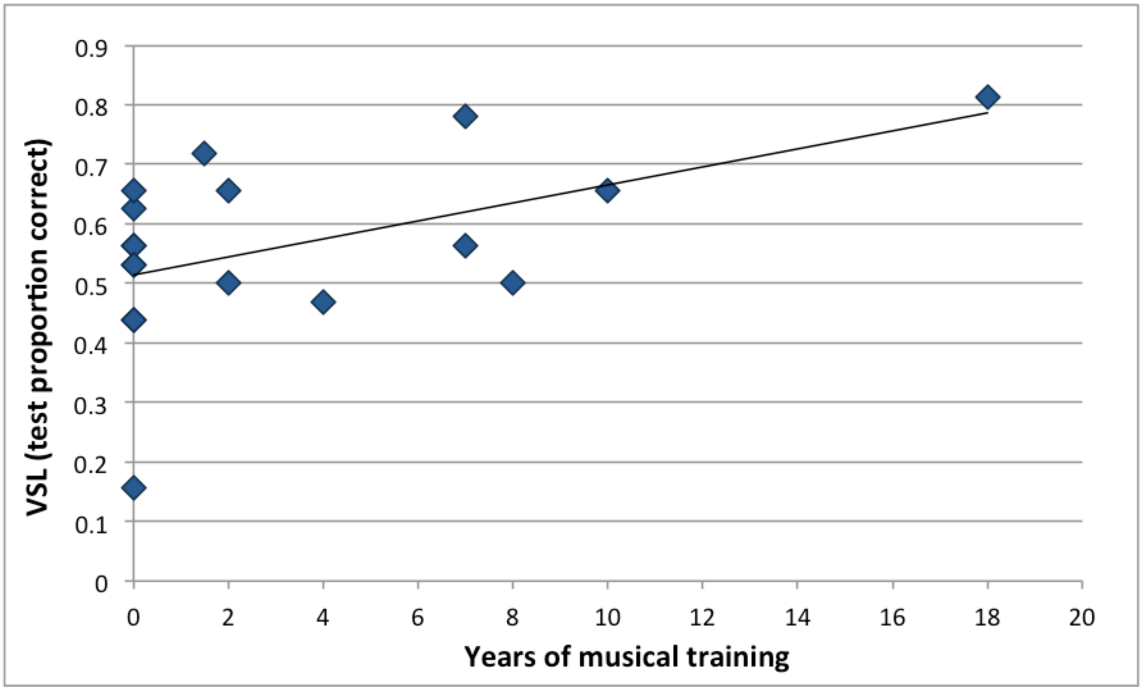


Figure 16. The relationship between performance on the test portion of the Visual Statistical Learning Task and years of formal musical training. A significant relationship was observed, $r = .50$.

Discussion

All expected ERP effects of language stimuli were observed in this study; however, we did not find a relationship between musical training and the amplitude of any ERP component. Contrary to our predictions, musical training does not seem to be associated with the strength of an individual's lexical predictions in language as indexed by the amplitude of the late frontal positivity.

We previously found that musical training was associated with a higher tendency to predict based on *hierarchical* structure (Figure 7). In contrast, the late frontal positivity is elicited by disconfirmed *lexical* predictions. If prediction occurs on multiple discrete levels in music, as it does in language, it is entirely possible that musicians' greater experience with one level of prediction might not be related to another. As music does not contain specific "lexical" content to the same degree as language, this particular type of prediction might not be enhanced in musicians. Instead, predictions related to hierarchical structure (such as grammatical categories) may be strengthened.

However, we did observe a relationship between musical training and Visual Statistical Learning task performance; while causality cannot be determined from this association, it is possible that musical training may enhance an individual's ability to learn statistical regularities and form predictions from the environment. As most previous studies have only demonstrated enhanced *auditory* statistical learning performance in musicians, it is notable that we found cross-modal effects in the visual domain (Shook et al., 2013; Skoe et al., 2013; Francois et al., 2014).

We did not find any relationships between musical training and any test of working memory. This null result is potentially attributable to our population of participants; while we used a correlational approach, most studies of the effects of musical training have compared discrete groups of musicians and nonmusicians. For example, Franklin et al. (2008) recruited trained musicians who had at least nine years of continuous training, played at least 15 hours a week, and were enrolled in an undergraduate or graduate music program; their nonmusician control participants had no history of playing an instrument before age 10 and had never played an instrument for longer than one year. In contrast, our participants were all amateurs with a wide range of experiences. Even most "nonmusicians" reported at least a little musical training at some point in their lives, and only three "musicians" majored in music. While this design has the advantage of measuring the impact of musical training incrementally and therefore potentially reducing confounds that might emerge by recruiting entirely dissimilar groups, we could have failed to recruit participants with the level of musical training or experience necessary to show effects on working memory or language processing.

Of course, it is also possible that musical training does not in fact impact prediction in language. Some studies that have observed differences between musicians and nonmusicians on language-related or cognitive tasks have found that these “effects” are in fact accounted for by some other correlated factor (such as IQ or socio-economic status) that differs between the groups. For example, Boebinger et al. (2015) found that non-verbal IQ (*not* music experience) predicted the ability to recognize speech in noise, and Okada & Slevc (2016) similarly observed that the relationship they observed between musicianship and a number of cognitive tasks did not persist after controlling for general intelligence, SES, and handedness.

General Discussion

In Study 1, we introduced the melodic cloze probability task, in which participants hear the opening of a short, novel tonal melody and sing the note they expect to come next. This task has not previously been used to study expectancy in the field of music cognition; previous methods of measuring melodic expectancy have not used participant-generated continuations. We also successfully manipulated the predictive constraint of short melodic sequences, demonstrating that this aspect of expectancy can be manipulated in a comparable way across linguistic and musical domains.

This paradigm is crucial for future research comparing predictive processing in music and language, as it allows for the comparison of the effects of violating predictions of comparable nature and strength in the two modalities. In contrast to previous studies comparing expectancy violations in music and language, which have typically chosen violations that are simply intuitively thought to be comparable in the two domains, the melodic cloze paradigm can be used to develop studies of the effects of plausible violations of musical expectations (analogous to coherent words that violate lexical predictions) instead of frank violations of musical structure. Hsu, Bars, & Ha (2015) have recently found ERP evidence for differential processing of mispredicted and unpredicted tones in simple sequences; this paradigm could be extended to naturalistic melodies using stimuli developed with the cloze paradigm.

The melodic cloze probability task can also be used to evaluate computational models of melodic expectancy. We have recently used this task to compare participants' expectations to the simplified version of the implication-realization (IR) model of melodic expectancy (Narmour, 1990; Schellenberg, 1997). This model uses melodic factors based on local note-to-note relationships to predict the probability of each possible continuation. The IR model was able to capture some of the variance in participants' expectations, but left much unexplained (Fogel, Morgan, & Patel, 2016). In future work, melodic cloze results can be used to compare additional quantitative models such as simple probabilistic models (Temperley, 2008), the IDyOM model (Pearce, 2005), and hierarchical models (Koelsch et al., 2013; Margulis, 2005). By taking into account more (and different) sources of context, these models may better be able to account for participants' expectations.

In Study 2, we failed to observe an effect of musical training on the late anterior positivity, an ERP response associated with lexical prediction in language. However, we did observe a relationship between musical training and visual statistical learning performance, demonstrating a cross-domain transfer effect of musical training on general predictive tendencies. As our participants were amateur musicians or nonmusicians with varying levels of experience, it is possible that an effect of musical training on the amplitude of ERP components associated with prediction or on working memory performance might be seen in more musically sophisticated participants (such as professional musicians).

If the processes of prediction in music and language are intrinsically linked beyond domain-general prediction as a whole, any relationship could be limited to specific types of prediction that are shared by both domains. As discussed above, lexical prediction in language has no clear analogue in music and may therefore be a poor candidate to observe an association between musical training and language processing; there may be effects of musical training on other specific types of predictions. For example, it could be fruitful to investigate syntactic prediction in musicians and non musicians; Jentschke & Koelsch (2009) have previously found that children with musical training had larger amplitudes of the ELAN, an ERP component associated with syntactic processing in language, and it has frequently been suggested that syntax processing resources may be shared between music and language (Kunert, Willems, & Hagoort, 2016; Patel, 2012). Future work using these approaches has the potential to provide insight into the cognitive and neural relations between prediction in music and language.

References

Aarden, B. J. (2003). *Dynamic Melodic Expectancy*. https://doi.org/10.1024/0301-1526.32.1.54

Altmann, G. T. M., & Kamide, Y. (1999). Incremental interpretation at verbs: Restricting the domain of subsequent reference. *Cognition*, *73*(3), 247–264. https://doi.org/10.1016/S0010-0277(99)00059-1

Bialystok, E., & Depape, A.-M. (2009). Musical expertise, bilingualism, and executive functioning. *Journal of Experimental Psychology: Human Perception and Performance*, *35*(2), 565–574. https://doi.org/10.1037/a0012735

Boebinger, D., Evans, S., Rosen, S., Lima, C. F., Manly, T., & Scott, S. K. (2015). Musicians and non-musicians are equally adept at perceiving masked speech. *The Journal of the Acoustical Society of America*, *137*(1), 378–87. https://doi.org/10.1121/1.4904537

Boersma, P. (2001). PRAAT, a system for doing phonetics by computer. *Glot International*, *5*(9/10), 341–345. https://doi.org/10.1097/AUD.0b013e31821473f7

Boudewyn, M. A., Long, D. L., & Swaab, T. Y. (2013). Effects of working memory span on processing of lexical associations and congruence in spoken discourse. *Frontiers in Psychology*, *4*(February), 60. https://doi.org/10.3389/fpsyg.2013.00060

Carlsen, J. (1981). Some factors which influence melodic expectancy. *Psychomusicology: A Journal of Research in Music …*. Retrieved from http://psycnet.apa.org/journals/pmu/1/1/12/

Carrus, E., Pearce, M. T., & Bhattacharya, J. (2012). Melodic pitch expectation interacts with neural responses to syntactic but not semantic violations. *Cortex*, 1–15. https://doi.org/10.1016/j.cortex.2012.08.024

Chobert, J., Francois, C., Velay, J. L., & Besson, M. (2014). Twelve months of active musical training in 8-to 10-year-old children enhances the preattentive processing of syllabic duration and voice onset time. *Cerebral Cortex*, *24*(4), 956–967. https://doi.org/10.1093/cercor/bhs377

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, *36*(3), 181–204. https://doi.org/10.1017/S0140525X12000477

Clayton, K. K., Swaminathan, J., Yazdanbakhsh, A., Zuk, J., Patel, A. D., & Kidd, G. (2016). Executive function, visual attention and the cocktail party problem in musicians and non-musicians. *PLoS ONE*, *11*(7), 1–17. https://doi.org/10.1371/journal.pone.0157638

Corrigall, K. A., & Trainor, L. J. (2014). Enculturation to musical pitch structure in young children: Evidence from behavioral and electrophysiological methods. *Developmental Science*, *17*(1), 142–158. https://doi.org/10.1111/desc.12100

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning & Verbal Behavior. Vol*, *19*(4), 450–466.

DeLong, K. a, Urbach, T. P., & Kutas, M. (2005). Probabilistic word pre-activation during language comprehension inferred from electrical brain activity. *Nature Neuroscience*, *8*(8), 1117–21. https://doi.org/10.1038/nn1504

Delorme, A., & Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, *134*(1), 9–21. https://doi.org/10.1016/j.jneumeth.2003.10.009

Dienes, Z., & Longuet-Higgins, C. (2004). Can musical transformations be implicitly learned? *Cognitive Science*, *28*(4), 531–558. https://doi.org/10.1016/j.cogsci.2004.03.003

Eerola, T., & Toiviainen, P. (2004). MIDI Toolbox: MATLAB Tools for Music Research. *MATLAB Tools for Music Research. Jyväskylä, Finland: …*, 96. Retrieved from http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:MIDI+Toolbox#6

Featherstone, C. R., Waterman, M. G., & Morrison, C. M. (2012). Norming the odd: creation, norming, and validation of a stimulus set for the study of incongruities across music and language. *Behavior Research Methods*, *44*(1), 81–94. https://doi.org/10.3758/s13428-011-0137-1

Federmeier, K. D., Wlotko, E. W., De Ochoa-Dewald, E., & Kutas, M. (2007). Multiple effects of sentential constraint on word processing. *Brain Research*, *1146*(1), 75–84. https://doi.org/10.1016/j.brainres.2006.06.101

Fedorenko, E. (2014). The role of domain-general cognitive control in language comprehension. *Frontiers in Psychology*, *5*(APR), 1–17. https://doi.org/10.3389/fpsyg.2014.00335

Fedorenko, E., McDermott, J. H., Norman-Haignere, S., & Kanwisher, N. (2012). Sensitivity to musical structure in the human brain. *Journal of Neurophysiology*, *108*(12), 3289–300. https://doi.org/10.1152/jn.00209.2012

Fedorenko, E., Patel, A. D., Casasanto, D., Winawer, J., & Gibson, E. (2009). Structural integration in language and music: evidence for a shared system. *Memory & Cognition*, *37*(1), 1–9. https://doi.org/10.3758/MC.37.1.1

Fiser, J., & Aslin, R. N. (2002). Statistical learning of higher-order temporal structure from visual shape sequences. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *28*(3), 458–467. https://doi.org/10.1037/0278-7393.28.3.458

Fogel, A. R., Morgan, E., & Patel, A. D. (2016). Measuring and Modeling Melodic Expectation: A New Approach. In *Proceedings of the 14th International Conference on Music Perception and Cognition*. San Francisco.

Fogel, A. R., Rosenberg, J. C., Lehman, F. M., Kuperberg, G. R., & Patel, A. D. (2015). Studying musical and linguistic prediction in comparable ways: The melodic cloze probability method. *Frontiers in Psychology*, *6*(NOV), 1–14. https://doi.org/10.3389/fpsyg.2015.01718

Francois, C., Chobert, J., Besson, M., & Schon, D. (2013). Music training for the development of speech segmentation. *Cerebral Cortex*, *23*(9), 2038–2043. https://doi.org/10.1093/cercor/bhs180

Francois, C., Jaillet, F., Takerkar, S., & Schon, D. (2014). Faster sound stream segmentation in musicians than in nonmusicians. *PLoS ONE*, *9*(7), 26–29. https://doi.org/10.1371/journal.pone.0101340

Francois, C., & Schon, D. (2014). Neural sensitivity to statistical regularities as a fundamental biological process that underlies auditory learning: The role of musical practice. *Hearing Research*, *308*, 122–128. https://doi.org/10.1016/j.heares.2013.08.018

Francois, C., & Schön, D. (2011). Musical expertise boosts implicit learning of both musical and linguistic structures. *Cerebral Cortex*, *21*(10), 2357–2365. https://doi.org/10.1093/cercor/bhr022

Franklin, M. S., Moore, K. S., Yip, C.-Y., Jonides, J., Rattray, K., & Moher, J. (2008). The effects of musical training on verbal memory. *Psychology of Music*, *36*(3), 353–365. https://doi.org/10.1177/0305735607086044

Fujioka, T., Ween, J. E., Jamali, S., Stuss, D. T., & Ross, B. (2012). Changes in neuromagnetic beta-band oscillation after music-supported stroke rehabilitation. *Annals of the New York Academy of Sciences*, *1252*, 294–304. https://doi.org/10.1111/j.1749-6632.2011.06436.x

Hagoort, P. (2005). On Broca, brain, and binding: A new framework. *Trends in Cognitive Sciences*, *9*(9), 416–423. https://doi.org/10.1016/j.tics.2005.07.004

Hansen, M., Wallentin, M., & Vuust, P. (2013). Working memory and musical competence of musicians and non-musicians. *Psychology of Music*, *41*(6), 779–793. https://doi.org/10.1177/0305735612452186

Hoch, L., Poulin-Charronnat, B., & Tillmann, B. (2011). The influence of task-irrelevant music on language processing: syntactic and semantic structures. *Frontiers in Psychology*, *2*(May), 112. https://doi.org/10.3389/fpsyg.2011.00112

Hsu, Y., Bars, S. Le, & Ha, J. a. (2015). Distinctive Representation of Mispredicted and Unpredicted Prediction Errors in Human Electroencephalography. *The Journal of Neuroscience*, *35*(43), 14653–14660. https://doi.org/10.1523/JNEUROSCI.2204-15.2015

Huron, D. (2006). *Sweet Anticipation: Music and the Psychology of Expectation*. *Music Perception* (Vol. 24). Cambridge, MA: MIT Press. https://doi.org/10.1525/mp.2007.24.5.511

Huron, D., & Margulis, E. H. (2010). Musical Expectancy and Thrills. In P. N. Juslin & J. A. Sloboda (Eds.), *Handbook of Music and Emotion: Theory, Research, Applications* (pp. 575–604). New York, NY: Oxford University Press.

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

Jentschke, S., & Koelsch, S. (2009). Musical training modulates the development of syntax processing in children. *NeuroImage*, *47*(2), 735–44. https://doi.org/10.1016/j.neuroimage.2009.04.090

Koelsch, S. (2012a). *Brain and Music*. *Brain and Music*. John Wiley & Sons. https://doi.org/10.1017/CBO9781107415324.004

Koelsch, S. (2012b). Response to target article "Language, music, and the brain: A resource-sharing framework." *Language and Music as Cognitive Systems*. https://doi.org/10.1093/acprof:oso/9780199553426.003.0023

Koelsch, S., Gunter, T. C., Wittfoth, M., & Sammler, D. (2005). Interaction between syntax processing in language and in music: an ERP Study. *Journal of Cognitive Neuroscience*, *17*(10), 1565–77. https://doi.org/10.1162/089892905774597290

Koelsch, S., Gunter, T., Friederici, A. D., & Schroger, E. (2000). Brain Indices of Music Processing: "Nonmusicians'' are Musical. *Journal of Cognitive Neuroscience*, *12*(3), 520–541.

Koelsch, S., Rohrmeier, M., Torrecuso, R., & Jentschke, S. (2013). Processing of hierarchical syntactic structure in music. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(38), 15443–8. https://doi.org/10.1073/pnas.1300272110

Koelsch, S., Schmidt, B.-H., & Kansok, J. (2002). Effects of musical expertise on the early right anterior negativity: an event-related brain potential study. *Psychophysiology*, *39*, 657–663. https://doi.org/10.1017.S0048577202010508

Krumhansl, C. L. (1990). *Cognitive Foundations of Musical Pitch*. *Cognitive Foundations of Musical Pitch*. New York, NY: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195148367.001.0001

Krumhansl, C. L., & Kessler, E. J. (1982). Tracing the Dynamic Changes in Perceived Tonal Organization in a Spatial Representation of Musical Keys. *Psychological Review*, *89*(4), 334–368.

Krumhansl, C. L., Louhivuori, J., Toiviainen, P., & Eerola, T. (1999). Melodic Expectation in Finnish Spiritual Folk Hymes: Convergence of Statistical, Behavioral, Computational Approaches. *Music Perception: An Interdisciplinary Journal*, *17*(2), 151–195. https://doi.org/10.2307/40285890

Kunert, R., & Slevc, L. R. (2015). A Commentary on: "Neural overlap in processing music and speech." *Frontiers in Human Neuroscience*, *9*(June), 1–3. https://doi.org/10.3389/fnhum.2015.00330

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

Kuperberg, G. R. (2007). Neural mechanisms of language comprehension: challenges to syntax. *Brain Research*, *1146*, 23–49. https://doi.org/10.1016/j.brainres.2006.12.063

Kuperberg, G. R. (2013). The Proactive Comprehender: What Event-Related Potentials tell us about the dynamics of reading comprehension. In: Unraveling the Behavioral, Neurobiological, and Genetic Components of Reading Comprehension. In *Unraveling the Behavioral, Neurobiological, and Genetic Components of Reading Comprehension* (pp. 1–23).

Kuperberg, G. R., & Jaeger, T. F. (2016). What do we mean by prediction in language comprehension? *Language Cognition & Neuroscience*, *31*(1), 32–59. https://doi.org/10.1080/23273798.2015.1102299

Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: finding meaning in the N400 component of the event-related brain potential (ERP). *Annual Review of Psychology*, *62*, 621–47. https://doi.org/10.1146/annurev.psych.093008.131123

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

Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: brain potentials reflect semantic incongruity. *Science (New York, N.Y.)*, *207*(4427), 203–205. https://doi.org/10.1126/science.7350657

LaCroix, A. N., Diaz, A. F., & Rogalsky, C. (2015). The relationship between the neural computations for speech and music perception is context-dependent: an activation likelihood estimate study. *Frontiers in Psychology*, *6*(August), 1–19. https://doi.org/10.3389/fpsyg.2015.01138

Lake, W. (1987). *Melodic Perception and Cognition: The Influence of Tonality*.

Lima, C. F., & Castro, S. L. (2011). Speaking to the trained ear: musical expertise enhances the recognition of emotions in speech prosody. *Emotion (Washington, D.C.)*, *11*(5), 1021–31. https://doi.org/10.1037/a0024521

Lopez-Calderon, J., & Luck, S. J. (2014). ERPLAB: an open-source toolbox for the analysis of event-related potentials. *Frontiers in Human Neuroscience*, *8*(April), 213. https://doi.org/10.3389/fnhum.2014.00213

Loui, P., Wessel, D., & Kam, C. (2010). Humans Rapidly Learn Grammatical Structure in a New Musical Scale. *Music Perception*, 377–388. Retrieved from http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2927013/

Maess, B., Koelsch, S., Gunter, T. C., & Friederici, A. D. (2001). Musical syntax is processed in Broca's area: an MEG study. *Nature Neuroscience*, *4*(5), 540–545. Retrieved from http://www.nature.com/?file=/neuro/journal/v4/n5/full/nn0501_540.html

Mandikal Vasuki, P. R., Sharma, M., Demuth, K., & Arciuli, J. (2016). Musicians' edge: A comparison of auditory processing, cognitive abilities and statistical learning. *Hearing Research*, *342*, 112–123. https://doi.org/10.1016/j.heares.2016.10.008

Margulis, E. H. (2005). A model of melodic expectation. *Music Perception*, *22*(4), 663–714. https://doi.org/10.1525/mp.2005.22.4.663

Margulis, E. H., & Levine, W. H. (2006). Timbre priming effects and expectation in melody. *Journal of New Music Research*, *35*(2), 175–182. https://doi.org/10.1080/09298210600835042

Meyer, L. B. (1956). *Emotion and Meaning in Music*. Chicago, IL: University of Chicago Press. https://doi.org/10.2307/843099

Moreno, S., Marques, C., Santos, A., Santos, M., Castro, S. L., & Besson, M. (2009). Musical training influences linguistic abilities in 8-year-old children: More evidence for brain plasticity. *Cerebral Cortex*, *19*(3), 712–723. https://doi.org/10.1093/cercor/bhn120

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

Narmour, E. (1990). *The Analysis and Cognition of Basic Melodic Structures*. Chicago, IL: University of Chicago Press.

Okada, B. M., & Slevc, L. R. (2016). Musical Experience and Executive Function. In *Proceedings of the 14th International Conference on Music Perception and Cognition*. San Francisco.

Paraskevopoulos, E., Kuchenbuch, A., Herholz, S. C., & Pantev, C. (2012). Statistical learning effects in musicians and non-musicians: An MEG study. *Neuropsychologia*, *50*(2), 341–349. https://doi.org/10.1016/j.neuropsychologia.2011.12.007

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

Patel, A. D. (2012). Language, music, and the brain: a resource-sharing framework. In *Language and music as cognitive systems* (Vol. 31, pp. 1–7). P. Rebuschat, M. Rohrmeier, J. Hawkins, & I. Cross. Retrieved from http://vesicle.nsi.edu/users/patel/Patel_LMCS_response.pdf

Patel, A. D., Gibson, E., Ratner, J., Besson, M., & Holcomb, P. (1998). Processing syntactic relations in language and music: An event-related potential study. *Journal of Cognitive Neuroscience*, *10*(6), 717–733. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9831740

Patel, A. D., & Morgan, E. (2016). Exploring Cognitive Relations Between Prediction in Language and Music. *Cognitive Science*, 1–18. https://doi.org/10.1111/cogs.12411

Pearce, M. (2005). *The Construction and Evaluation of Statistical Models of Melodic Structure in Music Perception and Composition*.

Pearce, M. T., Ruiz, M. H., Kapasi, S., Wiggins, G. a, & Bhattacharya, J. (2010). Unsupervised statistical learning underpins computational, behavioural, and neural manifestations of musical expectation. *NeuroImage*, *50*(1), 302–313. https://doi.org/10.1016/j.neuroimage.2009.12.019

Pearce, M. T., & Wiggins, G. a. (2006). Expectation in Melody: The Influence of Context and Learning. *Music Perception*, 377–405. Retrieved from http://www.jstor.org/stable/10.1525/mp.2006.23.5.377

Peretz, I. (1993). Auditory atonalia for melodies. *Cognitive Neuropsychology*, *10*(April 2015), 21–56. https://doi.org/10.1080/02643299308253455

Perruchet, P., & Poulin-Charronnat, B. (2013). Challenging prior evidence for a shared syntactic processor for language and music. *Psychonomic Bulletin & Review*, *20*(2), 310–7. https://doi.org/10.3758/s13423-012-0344-5

Pickering, M. J., & Garrod, S. (2007). Do people use language production to make predictions during comprehension? *Trends in Cognitive Sciences*, *11*(3), 105–110. https://doi.org/10.1016/j.tics.2006.12.002

Povel, D. (1996). Exploring the elementary harmonic forces in the tonal system. *Psychological Research*, *58*(June 1995), 274–283. Retrieved from http://link.springer.com/article/10.1007/BF00447073

Rohrmeier, M. a, & Koelsch, S. (2012). Predictive information processing in music cognition. A critical review. *International Journal of Psychophysiology*, *83*(2), 164–75. https://doi.org/10.1016/j.ijpsycho.2011.12.010

Rohrmeier, M., & Rebuschat, P. (2012). Implicit learning and acquisition of music. *Topics in Cognitive Science*, *4*(4), 525–53. https://doi.org/10.1111/j.1756-8765.2012.01223.x

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*. https://doi.org/10.1126/science.274.5294.1926

Schellenberg, E. (1996). Expectancy in melody: tests of the implication-realization model. *Cognition*, *58*(1), 75–125. https://doi.org/10.1016/0010-0277(95)00665-6

Schellenberg, E. G. (1997). Simplifying the Implication-Realization Model of Melodic Expectancy. *Music Perception*, *14*(3), 295–318.

Schellenberg, E. G., Adachi, M., Purdy, K. T., & McKinnon, M. C. (2002). Expectancy in melody: Tests of children and adults. *Journal of Experimental Psychology: General*, *131*(4), 511–537. https://doi.org/10.1037//0096-3445.131.4.511

Schmuckler, M. (1989). Expectation in Music: Investigation of Melodic and Harmonic Processes. *Music Perception*, *7*(2), 109–150. Retrieved from http://www.jstor.org/stable/40285454

Schmuckler, M. (1990). The Performance of Global Expectations. *Psychomusicology: A Journal of Research in …*, *9*(2). Retrieved from http://psycnet.apa.org/journals/pmu/9/2/122/

Shook, A., Marian, V., Bartolotti, J., & Schroeder, S. R. (2013). Musical experience influences statistical learning of a novel language. *American Journal of Psychology*, *126*(1), 95–104. https://doi.org/10.5406/amerjpsyc.126.1.0095

Skoe, E., Krizman, J., Spitzer, E., & Kraus, N. (2013). The auditory brainstem is a barometer of rapid auditory learning. *Neuroscience*, *243*, 104–114. https://doi.org/10.1016/j.neuroscience.2013.03.009

Slater, J., Skoe, E., Strait, D. L., O'Connell, S., Thompson, E., & Kraus, N. (2015). Music training improves speech-in-noise perception: Longitudinal evidence from a community-based music program. *Behavioural Brain Research*, *291*, 244–252. https://doi.org/10.1016/j.bbr.2015.05.026

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*, 637–652. https://doi.org/10.3758/s13423-014-0712-4

Slevc, L. R., & Patel, A. D. (2011). Meaning in music and language: Three key differences Comment on "Towards a neural basis of processing musical semantics" by Stefan Koelsch. *Physics of Life Reviews*, *8*(2), 110-111-128. https://doi.org/10.1016/j.plrev.2011.05.003

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), 374–81. https://doi.org/10.3758/16.2.374

Steinbeis, N., Koelsch, S., & Sloboda, J. a. (2006). The role of harmonic expectancy violations in musical emotions: evidence from subjective, physiological, and neural responses. *Journal of Cognitive Neuroscience*, *18*(8), 1380–93. https://doi.org/10.1162/jocn.2006.18.8.1380

Swaminathan, J., Mason, C. R., Streeter, T. M., Best, V., Kidd, G., & Patel, A. D. (2015). Musical training, individual differences and the cocktail party problem. *Scientific Reports*, *5*(11628), 1–10. https://doi.org/10.1038/srep11628

Tanenhaus, M. K., Spivey-knowlton, M. J., Eberhard, K. M., Sedivy, J. C., Tanenhaus, M. K., Spivey-knowlton, M. J., … Sedivy, J. C. (2016). Integration of Visual and Linguistic Information in Spoken Language Comprehension., *268*(5217), 1632–1634.

Taylor, W. L. (1953). "Cloze procedure": a new tool for measuring readability. *Journalism Quarterly*, *30*, 415–433. Retrieved from http://search.ebscohost.com/login.aspx?direct=true&db=psyh&AN=1955-00850-001&site=ehost-live

Temperley, D. (2008). A probabilistic model of melody perception. *Cognitive Science*, *32*(2), 418–44. https://doi.org/10.1080/03640210701864089

Thompson, W., Cuddy, L., & Plaus, C. (1997). Expectancies generated by melodic intervals: Evaluation of principles of melodic implication in a melody-completion task. *Perception & Psychophysics*, *59*(7), 1069–1076. Retrieved from http://link.springer.com/article/10.3758/BF03205521

Thompson, W. F., Schellenberg, E. G., & Husain, G. (2004). Decoding speech prosody: do music lessons help? *Emotion (Washington, D.C.)*, *4*(1), 46–64. https://doi.org/10.1037/1528-3542.4.1.46

Tillmann, B., Bharucha, J. J., & Bigand, E. (2000). Implicit learning of tonality: A self-organizing approach. *Psychological Review*, *107*(4), 885–913. https://doi.org/10.1037//0033-295X.107.4.885

Tillmann, B., & Bigand, E. (2015). Response: A Commentary on: "Neural overlap in processing music and speech." *Frontiers in Human Neuroscience*, *9*(June). https://doi.org/10.3389/fnhum.2015.00330

Tillmann, B., Janata, P., & Bharucha, J. J. (2003). Activation of the inferior frontal cortex in musical priming. *Cognitive Brain Research*, *16*(2), 145–161. https://doi.org/10.1016/S0926-6410(02)00245-8

Tillmann, B., Janata, P., Birk, J., & Bharucha, J. J. (2003). The costs and benefits of tonal centers for chord processing. *Journal of Experimental Psychology: Human Perception and Performance*, *29*(2), 470–482. https://doi.org/10.1037/0096-1523.29.2.470

Tillmann, B., Poulin-Charronnat, B., & Bigand, E. (2013). The role of expectation in music: from the score to emotions and the brain. *Wiley Interdisciplinary Reviews: Cognitive Science*, n/a-n/a. https://doi.org/10.1002/wcs.1262

Traxler, M. J., Williams, R. S., Blozis, S. A., & Morris, R. K. (2005). Working memory, animacy, and verb class in the processing of relative clauses. *Journal of Memory and Language*, *53*(2), 204–224. https://doi.org/10.1016/j.jml.2005.02.010

Unsworth, N., Heitz, R. P., Schrock, J. C., & Engle, R. W. (2005). An automated version of the operation span task. *Behavior Research Methods*, *37*(3), 498–505. https://doi.org/10.3758/bf03192720

Unyk, A., & Carlsen, J. (1987). The influence of expectancy on melodic perception. *Psychomusicology: Music, Mind & Brain*. Retrieved from http://www.psycontent.com/index/3841843323148W37.pdf

Van Petten, C., & Luka, B. J. (2012a). Prediction during language comprehension: Benefits, costs, and ERP components. *International Journal of Psychophysiology*, *83*(2), 176–190. https://doi.org/10.1016/j.ijpsycho.2011.09.015

Van Petten, C., & Luka, B. J. (2012b). Prediction during language comprehension: Benefits, costs, and ERP components. *International Journal of Psychophysiology*. Elsevier B.V. https://doi.org/10.1016/j.ijpsycho.2011.09.015

Van Petten, C., Weckerly, J., McIsaac, H. K., & Kutas, M. (1997). Working memory capacity dissociated lexical and sentental context effects. *Psychological Science*, *8*(3), 238–242.

Vassena, E., Kochman, K., Latomme, J., & Verguts, T. (2016). Unimodal and cross-modal prediction is enhanced in musicians. *Nature Scientific Reports*, *6*(April), 25225. https://doi.org/10.1038/srep25225

Wallentin, M., Nielsen, A. H., Friis-Olivarius, M., Vuust, C., & Vuust, P. (2010). The Musical Ear Test, a new reliable test for measuring musical competence. *Learning and Individual Differences*, *20*(3), 188–196. https://doi.org/10.1016/j.lindif.2010.02.004

Warker, J., & Halpern, A. (2005). Musical stem completion: Humming that note. *The American Journal of Psychology*, *118*(4), 567–585. Retrieved from http://www.jstor.org/stable/30039086

Wicha, N. Y. Y., Moreno, E. M., & Kutas, M. (2004). Anticipating words and their gender: an event-related brain potential study of semantic integration, gender expectancy, and gender agreement in Spanish sentence reading. *Journal of Cognitive Neuroscience*, *16*(7), 1272–88. https://doi.org/10.1162/0898929041920487

Wlotko, E. W., Federmeier, K. D., & Kutas, M. (2012). To predict or not to predict: age-related differences in the use of sentential context. *Psychology and Aging*, *27*(4), 975–88. https://doi.org/10.1037/a0029206

Zioga, I., Di Bernardi Luft, C., & Bhattacharya, J. (2016). Musical training shapes neural responses to melodic and prosodic expectation. *Brain Research*, *1650*, 267–282. https://doi.org/10.1016/j.brainres.2016.09.015

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