

**A contextual constraint approach to studying melodic
expectation: behavioral and neural studies**

A dissertation

submitted by

Allison R. Fogel

In partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

in

Psychology

TUFTS UNIVERSITY

May, 2019

Committee Members:

Aniruddh D. Patel, Ph.D.

Gina R. Kuperberg, M.D., Ph.D.

Elizabeth Race, Ph.D.

L. Robert Slevc, Ph.D.

Abstract

Melodic expectation is a perennial topic in music cognition. A number of ERP studies of melodic expectation have been conducted, but they have so far been limited in their approach: they have manipulated the expectancy of a target note without investigating how the prior context contributes to forming expectations. We have developed a new approach to studying melodic expectation inspired by the paradigms used in studies of sentence processing: manipulating contextual constraint, or the extent to which a context constrains expectations to one particular continuation. Previously, we developed the melodic cloze paradigm, where participants were asked to sing continuations to short melodic contexts. Using this task, it is possible to develop melodies that either do or do not lead to a strong prediction for a certain note to come next, as well as to quantify participants' expectations for different continuations. Here, we assessed the reliability of the melodic cloze paradigm by replicating it with different participants and task instructions, demonstrating that it is indeed a reliable method for quantifying participants' expectations in melodic sequences.

We then used stimuli developed through this task to study the neural correlates of melodic expectation using a set of novel ERP paradigms. We first explored the ERP effects of violating predictions for a specific note with unexpected but in-key target notes, then repeated the paradigm using out-of-key target notes. When in-key target notes violated a prediction for a different note, they elicited a sustained frontal positivity. Interestingly, this result resembles the

frontal positivity elicited by plausible words that violate predictions for a different word in language experiments. In contrast, out-of-key target notes elicited a late posterior positivity (P600) compared to expected notes. Unexpected notes did not elicit an early anterior negativity- the effect most frequently reported by previous ERP studies of melodic expectation- in either experiment. Finally, we discuss additional applications of the contextual constraint approach to studying melodic expectation, such as evaluating the accuracy of the predictions generated by different computational models of melodic expectation.

Acknowledgements

I would first like to sincerely thank my advisor Ani Patel for his support, patience, and guidance along every step of this journey. I would also like to thank my dissertation committee members Gina Kuperberg, Elizabeth Race, and Bob Slevc for their insights and feedback. I am grateful to the members of the Kuperberg Lab and Anjali Nair for their assistance in data collection and analysis, and especially to Emily Morgan for her invaluable contributions to these studies. Finally, thank you to my friends, family, and Buddy for their support and encouragement.

Table of Contents

Chapter 1. General Introduction	1
Chapter 2. Reliability of the cloze probability method for measuring melodic expectation	20
Introduction.....	20
Method	21
Experiment 1	21
Experiment 2	23
Results.....	24
Experiment 1 Results	24
Experiment 2 Results	29
Discussion	37
Chapter 3. Neural correlates of processing in-key melodic prediction violations.....	41
Introduction.....	41
Methods.....	45
Stimulus Construction.....	45
ERP Stimuli	47
Participants.....	48
Procedure	49
ERP data acquisition and processing	49
Analysis.....	51
Results.....	53
Behavioral	53
ERP	54
Regression.....	58
Discussion	62
Chapter 4. Neural correlates of processing out-of-key melodic prediction violations.....	65
Introduction.....	65
Methods.....	69
ERP Stimuli	69
Participants.....	71
Procedure	72

ERP data acquisition and processing	72
Analysis.....	73
Results.....	74
Behavioral	74
ERP	75
Regression.....	80
Experiment Comparison	82
Discussion	83
Chapter 5. Discussion, Applications, and Future Directions	86
ERP Discussion.....	86
Current Applications	90
Future Directions	94
Bibliography	99
Appendix A. MMN studies of melodic expectation	117
Appendix B. Notation for melodic stems used in Chapter 2.	121

List of Tables

Table 1.1. Previous ERP studies using unfamiliar melodies. Full citations are given in the bibliography.	9
Table 2.1. Participant characteristics.....	24
Table 3.1. Significant clusters in mass univariate analyses of pairwise comparisons.	57
Table 3.2. amplitude ~ constraint + cloze	59
Table 3.3. amplitude ~ condition + cloze.....	61
Table 3.4. amplitude ~ condition + constraint + cloze.....	62
Table 4.1. Size of the interval between pre-target notes and unexpected target notes in Experiment 2 as compared to Experiment 1.	71
Table 4.2. Significant clusters in mass univariate analyses of pairwise comparisons.	80
Table 4.3. amplitude ~ constraint + cloze	80
Table 4.4. amplitude ~ condition + cloze.....	82
Table 4.5. amplitude ~ condition + constraint + cloze.....	82

List of Figures

Figure 1.1. Sample stimuli from Margulis and Levine (2006).	4
Figure 1.2. Example stimuli from Koelsch et al. (2007).....	5
Figure 1.3. Results from Koelsch et al. (2007), showing more negative ERPs elicited by irregular chords compared to regular chords.	5
Figure 1.4. Sample stimuli from Brattico et al. (2006), showing versions of the same melody containing a congruous or out-of-key target note.....	7
Figure 1.5. Sample stimuli from Halpern et al. (2017)	7
Figure 1.6. Stimuli used in Granot and Donchin (2004).....	8
Figure 1.7. Results from Kuperberg et al. (2018): grand-averaged ERP waveforms at electrode FPz showing the late frontal positivity	13
Figure 1.8. Experimental design used in Hsu et al. (2015).....	15
Figure 1.9. Results from Hsu et al. (2015).....	15
Figure 1.10. Sample melodic stem pair used in Fogel et al. (2015).....	17
Figure 1.11. Distribution of participants' sung continuations for the melodic stem pair pictured in Figure 1.10.	18
Figure 2.1. Scale degree distribution of participants' sung continuations after the same melodic stems in the original Fogel et al. (2015) melodic cloze probability experiment and in the present replication.	25
Figure 2.2. The constraints of individual melodic stems in the original melodic cloze probability experiment compared to the present replication.	26
Figure 2.3. The distance between the cloze probability distributions for the 40 stems studied in Fogel et al. (2015) and in the replication experiment	27
Figure 2.4. Participants' responses to selected melodic stems in the two experiments.	29
Figure 2.5. Lengths of responses in each condition of Experiment 2.....	30
Figure 2.6. Average length of responses to the AC and NC versions of each melodic stem pair.....	30
Figure 2.7. Scale degree distribution of participants' sung responses after melodic stems in the original melodic cloze probability study and in the present experiment where they sang longer continuations.....	32
Figure 2.8. The constraints of individual melodic stems in the original melodic cloze probability experiment compared to their constraints in the present experiment where participants sang longer continuations..	33
Figure 2.9. Predictive constraint (the percentage of participants who sang the most common response) of matched stem pairs across the two studies.	34

Figure 2.10. The distance between cloze probability distributions in the two studies for each individual item..	35
Figure 2.11. Participants' responses to selected melodic stems in the two experiments.	37
Figure 3.1. Sample melodic stem pair.....	46
Figure 3.2. Distribution of participants' sung continuations for the melodic stem pair pictured in Figure 3.1.	46
Figure 3.3. Example ERP stimuli.....	48
Figure 3.4. Electrode layout used in Chapters 3 and 4.	50
Figure 3.5. For purposes of ERP analysis, melodies were divided into higher- and lower-constraint groups according to their empirical constraint based on the results of the melodic cloze probability paradigm.	53
Figure 3.6. Grand average ERPs for all conditions over frontal electrode sites. .	55
Figure 3.7. Grand average ERPs at electrode F7 to unexpected target notes in constraining melodies and unexpected target notes in non-constraining melodies	55
Figure 3.8. Grand average ERPs at electrode F3 to expected and unexpected target notes in constraining melodies.....	56
Figure 3.9. Topographic distributions of significant clusters revealed by mass univariate analyses of the main effects of Constraint and Cloze Probability.	58
Figure 3.10. Amplitude across frontal electrodes from 300-600 ms vs. contextual constraint for items in the HC-sixth and LC-sixth conditions.	60
Figure 4.1. Results from Kuperberg et al. (2018): grand-averaged ERP waveforms at electrode Oz showing the late posterior positivity	69
Figure 4.2. Example ERP stimuli in Experiment 2.	71
Figure 4.3. Grand average ERPs for all conditions at selected electrode sites. ...	76
Figure 4.4. Grand average ERPs at electrode F4 to expected and out-of-key target notes in constraining melodies.....	77
Figure 4.5. Grand average ERPs at electrode P4 to expected and out-of-key target notes in constraining melodies.....	78
Figure 4.6. Grand average ERPs at electrode AF4 to out-of-key target notes in constraining and non-constraining melodies	79
Figure 4.7. Grand average ERPs at electrode P4 to out-of-key target notes in constraining and non-constraining melodies	79
Figure 4.8. Amplitude across frontocentral electrodes from 100-250 ms vs. contextual constraint for items in the HC-unexpected and LC-unexpected conditions.....	81

Figure 4.9. Topographic distributions of the difference between expected and unexpected notes in constraining contexts between 500 and 900 ms in Experiments 1 and 2.....	83
Figure 5.1. Figure from Morgan et al. (2019). Human data from the melodic cloze experiments (top row) and raw predictions from both models.....	93

Chapter 1. General Introduction

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, a theory that is still the subject of much research today (Egermann, Pearce, Wiggins, & McAdams, 2013; Huron, 2006; Koelsch, 2012, 2014; Salimpoor, Zald, Zatorre, Dagher, & McIntosh, 2014; Steinbeis, Koelsch, & Sloboda, 2006). In recent years, theories of musical expectation have been brought into a modern cognitive science framework, and expectation has been studied empirically with both behavioral and neural methods (e.g., Huron, 2006; Huron & Margulis, 2010; Koelsch, Vuust, & Friston, 2019; Margulis, 2005; Patel & Morgan, 2016; Pearce, Ruiz, Kapasi, Wiggins, & Bhattacharya, 2010; Pearce & Wiggins, 2006, 2012). 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 expectations are generated for multiple aspects of music, such as harmony, rhythm, timbre, and meter (Rohrmeier & Koelsch, 2012; Tillmann, Poulin-Charronnat, & Bigand, 2014). Such expectations are thought to be automatically generated by enculturated listeners, who have been passively exposed to music in their environments (Koelsch, 2012; Koelsch, Gunter, Friederici, & Schröger, 2000). Here, we focus on melodic prediction, specifically on expectations for what upcoming notes will be that are generated based on implicit knowledge of the melodic and harmonic structures of Western tonal music (Tillmann, Bharucha, & Bigand, 2000); we use “prediction” to refer to a

process by which representations are pre-activated before bottom-up input is available (Kuperberg & Jaeger, 2016).

In Western tonal music, the 7 tones belonging to a given key form a scale. The notes in a scale vary in their degree of structural stability and how related they are to each other, with more tonally stable notes being rated as more related (Krumhansl, 1979; Krumhansl & Cuddy, 2010; Krumhansl & Shepard, 1979). Within the context of a given key, the tonic (or base of the scale) is typically the most stable note and the most related to other notes in the key, and is therefore the most activated in activation-based models of the mental representation of music notes and their relations (e.g., Tillmann et al., 2000). Other notes that are harmonically related to the tonic (e.g., the fifth scale degree) are next most activated, then other diatonic (in-key) notes, with non-diatonic (out-of-key) notes being the least activated (Krumhansl, 1979). This same pattern applies to the chords based on these notes as well. Just as in language, where ‘cat’ and ‘dog’ are more related than ‘cat’ and ‘water’, the tonic and fifth are more related to each other than the tonic and a note that is out-of-key. These relationships are learned implicitly from environmental exposure; children without musical training can accurately make explicit judgments about whether notes belong in a key by five years old and exhibit neural sensitivity to key membership by four years old (Corrigall & Trainor, 2014; Garnier, Poulin-Charronnat, Schellenberg, Stevens, & Bigand, 2005; Jentschke, Friederici, & Koelsch, 2014; Schellenberg, Adachi, Purdy, & McKinnon, 2002; Trainor & Trehub, 1994).

These key-based tonal relationships can influence expectations for notes in melodies. In music cognition, studies have generally investigated melodic expectancy by manipulating the predictability or congruity of different target notes in a given context. For example, “probe-tone” paradigms have been used. In this task, participants hear the beginning of a musical context before being presented with a test tone and rating how well it “fits” as a continuation (Krumhansl & Shepard, 1979). For instance, Schmuckler (1989) presented participants with full musical contexts, which were halted and continued with different test (“probe”) tones at different points. Participants rated how well each final note “fit their expectations of what was to come next”. Using much shorter contexts, Cuddy and Lunney (1995) presented participants with two note intervals and asked them to rate how well a third test tone continued the melody. Similarly, Schellenberg (1996) used a tone-rating paradigm to evaluate the predictions of the Implication-Realization (IR) model, a gestalt-based of melodic expectancy (Narmour, 1990); participants listened to short unfamiliar melodic fragments, then rated how well test tones continued them.

Additional behavioral methods have been used to investigate influences on melodic expectancy as well. For instance, “priming” studies measure the speed or accuracy of participants’ responses to a target item. Margulis and Levine (2006) played participants short melodic contexts, after which they had to perform a timbre discrimination task on a test tone; timbre identification improved when a pitch was close in frequency to the preceding pitch and harmonically congruent with the preceding context (Figure 1.1). In Kim, Large, Gwon, and Ashley (2018),

participants heard unfamiliar tonal melodies ending with target tones, which they responded to by singing the target back as quickly as possible. Here, sing-back reaction times were shorter when target tones implied an expected harmony. A “betting” paradigm has also been used where participants place bets on a set of possible continuations for a musical passage, which can be distributed across multiple possible outcomes; this task has the advantage of providing a measure of the strength of an expectation for a range of specific notes, but it requires participants to make post hoc judgments and is rather time consuming (Huron, 2006).

Melody 1 and probes



Melody 2 and probes



Figure 1.1. Sample stimuli from Margulis and Levine (2006). Stemmed notes (the melody) were played, followed by one of the unstemmed notes (the probes) played in one of two test timbres.

Studies of expectancy in music have also made use of neural measures, mainly event-related potentials (ERPs, the focus of Chapters 3 and 4). The majority of ERP studies that have been conducted using musical stimuli have manipulated the relationship of the last chord of a sequence to its prior context (Koelsch, 2009; Koelsch et al., 2000; Koelsch, Jentschke, Sammler, & Mietchen, 2007; Rohrmeier & Koelsch, 2012). For example, (Koelsch et al., 2007) measured ERPs in response to “regular” and “irregular” chords. Here, regular chords were the tonic (the base of the scale), while irregular chords were either a double dominant (containing out-of-key notes) or a supertonic chord (syntactically

unexpected, but containing only in-key notes; see Figure 1.2). As compared to regular chords, irregular chords elicited a right-lateralized, frontal negative deflection in the ERP that was maximal about 150 ms after the onset of a chord (see Figure 1.3). This effect has been termed the early right anterior negativity, or ERAN, and has been the most frequent ERP component reported in ERP studies of musical expectations (Koelsch et al., 2019; Rohrmeier & Koelsch, 2012).



Figure 1.2. Example stimuli from Koelsch et al. (2007): chord sequences in D major, ending either on a tonic chord (regular, A), on a double dominant (irregular, B), or on a supertonic (irregular, C).

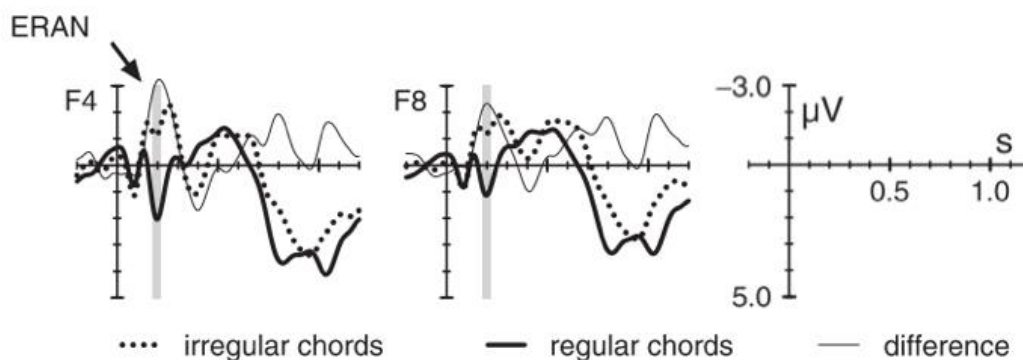


Figure 1.3. Results from Koelsch et al. (2007), showing more negative ERPs elicited by irregular chords compared to regular chords.

Fewer ERP studies of music have been conducted using monophonic melodic stimuli. Table 1.1 contains an overview of the existing ERP studies of melodic expectation that have used novel melodies (without immediate repetition during the experiment), unimodal auditory presentation, and typical healthy adult or child participants. (Additional relevant studies that did not meet these

parameters are listed in Appendix A.) The paradigms used in these studies vary significantly, differing in participant population, experimental task, and the placement and task relevance of target notes. However, the majority of melody studies have used the same fundamental experimental manipulation as most studies using chords: static melodic contexts are manipulated to contain “expected” or “unexpected” notes. In many cases, unexpected notes have simply been out-of-key (Besson & Faïta, 1995; Brattico, Tervaniemi, Näätänen, & Peretz, 2006; Choi et al., 2014; Demorest & Osterhout, 2012; Fiveash, Thompson, Badcock, & McArthur, 2018; Habibi, Wirantana, & Starr, 2013; Kalda & Minati, 2012; Kuriki, Isahai, & Ohtsuka, 2005; Lagrois, Peretz, & Zendel, 2018; Miranda & Ullman, 2007; Vuvan, Zendel, & Peretz, 2018; Zendel, Lagrois, Robitaille, & Peretz, 2015). For instance, Brattico et al. (2006) used unfamiliar melodies containing mid-sequence target notes that were either congruous, out-of-key, or out-of-tune (Figure 1.4). Others have used stimuli containing notes that are not out-of-key, but are still melodically unexpected in other ways; in-key notes can be melodically unexpected when they are low in the tonal hierarchy, imply an unlikely harmony, or create an unexpected melodic leap (Halpern, Martin, &

Reed, 2008; Halpern et al., 2017; Koelsch & Jentschke, 2010; Marmel, Perrin, & Tillmann, 2011; see Figure 1.5 for an example).

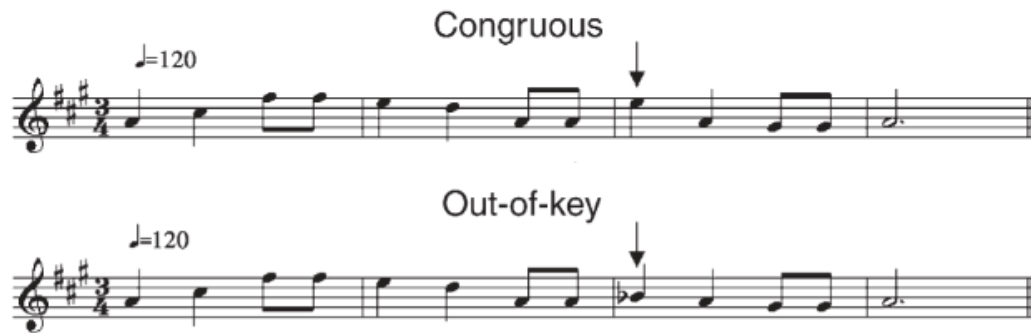


Figure 1.4. Sample stimuli from Brattico et al. (2006), showing versions of the same melody containing a congruous or out-of-key target note.



Figure 1.5. Sample stimuli from Halpern et al. (2017), showing (A) a melody with an expected completion and (B) a melody with an in-key completion that is tonally unexpected (low in the tonal hierarchy).

Some ERP studies have used other paradigms. Granot and Donchin (2004) constructed stimuli using a cloze probability-like task where participants wrote down musical notation for the note they expected to complete a sequence; melodies where more than 75% of participants wrote the same note were categorized as “strongly constraining,” while those where fewer than 50% of participants agreed on a continuation were categorized as “weakly constraining.” The resulting experimental melodies were either strongly or weakly constraining

and concluded with either “congruent” or “incongruent” (out-of-key) target notes (Figure 1.6). However, they still compared ERPs in response to highly expected “congruent” to out-of-key “incongruent” target notes. Pearce et al. (2010) used a computational model (IDyOM model: Pearce, 2005) to quantify the probabilities of different notes appearing in hymn melodies. Compared to high probability notes, low probability notes elicited a right-lateralized frontal positivity lasting from 280-600 ms, as well as a centroparietal negativity from 400-550 ms. While the stimuli were presented very slowly and without any rhythmic structure (each note had a duration of 1 second), the fact that they consisted of “real” melodies suggests that this paradigm was likely more natural than previous studies.



Figure 1.6. Stimuli used in Granot and Donchin (2004). Strongly or weakly constraining melodies were completed with congruent or incongruent (out-of-key) notes.

Table 1.1. Previous ERP studies using unfamiliar melodies. Full citations are given in the bibliography.

Citation	Participants	Unexpected notes	Task	Main neural response to unexpected vs. expected note
Besson and Faïta (1995)	musicians, non-musicians	out-of-key, final	recognition & congruity judgments	• LPC (400-800 ms)
Brattico et al. (2006)	non-musicians	out-of-key, mid-sequence	passive listening, congruity judgments	• frontal negativity (180-280 ms) • P600 (480-580 ms)
Choi et al. (2014)	non-musicians	out-of-key, mid-sequence (Markov seqs)	passive listening	• ↑ P2 (150-230 ms)
Corrigall and Trainor (2014)	4-year-old children, adults (suppl. info)	out-of-key, final	passive listening while watching silent movie	• children: none for melodies • adults: ERAN (135-195 ms)
Demorest and Osterhout (2012)	mixed college students	out-of-key, mid-sequence	congruity rating	• P600 (500-800 ms)
Fiveash et al. (2018)	mixed college students	out-of-key, mid-sequence	congruity judgments	• ERAN (150-250 ms)
Granot and Donchin (2004)	musicians, non-musicians	out-of-key, final	congruity judgments	• P300, ↑ in constraining contexts (300-700 ms)
Halpern et al. (2017)	older & younger adults	in-key, final	congruity rating	• ↑ N1 (80-130 ms) • ↑ P2 left (150-250 ms) • LPC (500-800 ms)
Halpern et al. (2008)	musicians, non-musicians	in-key, mid-sequence	major/minor classification	• LPC (400-700 ms)
Kalda and Minati (2012)	musicians, non-musicians	out-of-key, mid-sequence	timbre anomaly detection	• ERAN (230-290 ms)
Koelsch and Jentschke (2010)	non-musicians	in- and out-of-key, final	timbre anomaly detection	• ↑ negativity 90-150 ms

Kuriki et al. (2005)	non-musicians	out-of-key, final	congruity judgments	• MEG: ↑ N1/P2 (120-160 ms)
Lagrois et al. (2018)	older & younger non-musicians	out-of-key, mid-sequence	anomaly detection & congruity judgments	• ERAN (64-185 ms) • frontal/central negativity 555-613 ms • congruity task: P600 (472-635 ms)
Marmel et al. (2011)	musicians	in-key, mid-sequence	out-of-tune note detection	• more negative (0-100 ms) • more negative (280-380 ms)
Miranda and Ullman (2007)	musicians, non-musicians	out-of-key, mid-sequence	timbre anomaly detection	• ERAN (150-270 ms) • anterior negativity ("N5") (500-700 ms) • P600 (500-700 ms)
Omigie, Pearce, Williamson, and Stewart (2013)	non-musicians (controls)	in-key, mid-sequence	timbre anomaly detection	• ↑ N1/ERAN (96-116 ms)
Pearce et al. (2010)	non-musicians	in-key, mid-sequence	passive listening	• negativity 90-160 ms • right frontal positivity 280-600 ms • centroparietal negativity 450-550 ms
Peretz, Brattico, Järvenpää, and Tervaniemi (2009)	non-musicians (controls)	out-of-key, mid-sequence	anomaly detection	• ↑ N200/ERAN (180-300 ms) • P600 (600-780 ms)
Sun et al. (2018)	non-musicians (controls)	out-of-key, final	timbre anomaly detection	• ERAN (130-250 ms) • N5 (500-650 ms)
Vuvan et al. (2018)	non-musicians	out-of-key, mid-sequence	congruity judgments	• ERAN (100-250 ms) • P300 (350-525 ms) • P600 (525-700 ms)
Zendel et al. (2015)	non-musicians (controls)	out-of-key, mid-sequence	click detection, pitch anomaly detection	• click detection task: ERAN (139-283 ms) • pitch detection task: P600 (439-637 ms)

The results of ERP studies of melodic processing have been somewhat heterogeneous (see Table 1.1), but the most frequent finding has been an early frontal negativity elicited by unexpected notes compared to expected notes, usually interpreted as an “ERAN”; it has been reported to start as early as 64 ms after the presentation of a target note, but more typically appears around 150 ms (e.g., Fiveash et al., 2018; Koelsch & Jentschke, 2010; Lagrois et al., 2018; Miranda & Ullman, 2007; Vuhan et al., 2018). Late posterior positivities around 500-700 ms have often been observed as well. These are described in a variety of terms (late positive component, LPC), but most likely represent the P600 (e.g., Besson & Faïta, 1995; Brattico et al., 2006; Demorest & Osterhout, 2012; Miranda & Ullman, 2007).

Most ERP studies of melody processing have focused on detecting easily-perceivable incongruities. Paradigms of this kind have been successfully used to explore a number of research questions. For example, they have been used to investigate differences between music processing in older and younger adults (Halpern et al., 2017; Lagrois et al., 2018), interactions between music and language processing (Fiveash, Thompson, et al., 2018), and pitch processing in amusia (Zendel et al., 2015). However, paradigms containing obvious incongruities are less suited to studying how listeners process “normal” music. When musical notes are “implausible,” neural responses may actually reflect processing related to incongruity or error detection. Additionally, it is difficult to quantify *how* unexpected incongruent target notes are (or how expected congruent notes are) when they are chosen through application of some general rule (e.g.,

being out-of-key) without the use of an objective task of some kind to measure the strength of participants' expectations.

The present approach to studying melodic expectation is inspired by current research in language processing. In psycholinguistics, ERPs are often used to measure the neural correlates of prediction in language. 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 & Federmeier, 2011; Kutas & Hillyard, 1980, 1984). 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 related to the prior sentence context and thus less expected.

In recent language ERP research, *violations* of specific lexical predictions with other plausible words have 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; Kuperberg, Brothers, & Wlotko, 2018; Van Petten & Luka, 2012a). 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...'*). It is often elicited when predictions have been formed and then violated, while the N400 is just as large to unexpected words appearing in constraining or in non-constraining contexts where no prediction could have been made. The precise significance of this late frontal positivity is not yet entirely clear; it may index the costs of inhibiting incorrect predictions and updating one's internal representation of context in order to generate more accurate predictions in the future (see Kuperberg, 2013; Kuperberg et al., 2018; Kuperberg & Jaeger, 2016 for discussion).

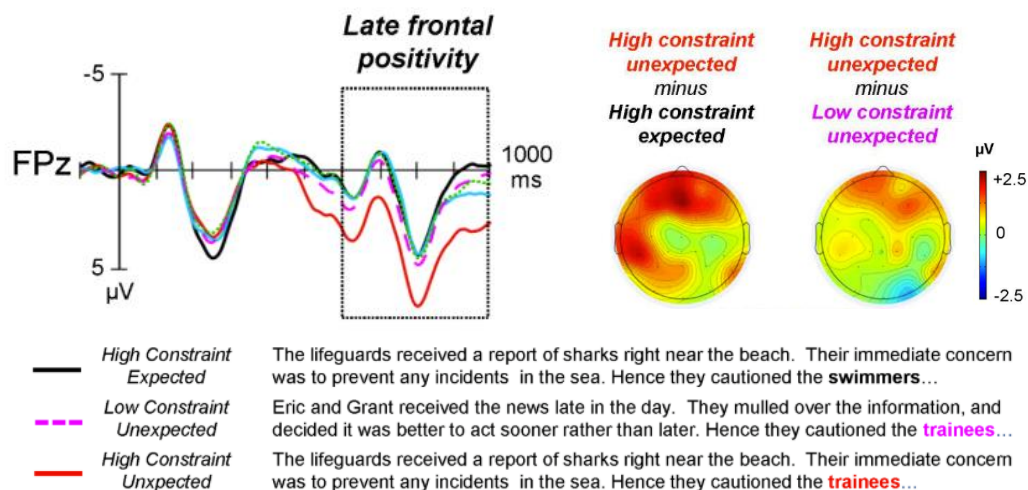


Figure 1.7. Results from Kuperberg et al. (2018): grand-averaged ERP waveforms at electrode FPz showing the late frontal positivity, which is larger to unexpected words in high constraint contexts than to unexpected words in low

constraint contexts, and voltage maps showing the scalp topographies of the effect averaged across the 600-1000ms time window.

Very little previous research using tonal auditory stimuli has distinguished between notes that are unexpected because they violate a prediction for a different item and notes that are unexpected because their context did not lead to the formation of any predictions. One exception is Hsu, Le Bars, Hämäläinen, and Waszak (2015), who used very simple tone sequences to differentiate between predicted, mispredicted (i.e., incorrectly predicted), and unpredicted tones (i.e., no prediction took place at all). Participants were presented with either a series of four notes in a scale that would be expected to continue in a pattern (e.g., the rising note pattern D-E-F-G is expected to be followed by A) or a random series of tones while performing a volume-monitoring task. 75% of trials consisted of tones in a scale pattern that did continue with the expected fifth note, 12.5% were tones in a scale pattern that would lead participants to expect the fifth tone of the series but instead continued with an *unexpected* (mispredicted) note, and 12.5% were random series of tones that did not lead to a prediction (Figure 1.8). Differences in neural responses were observed on the N1 ERP component, with its amplitude reduced for unpredicted tones and enhanced for mispredicted tones compared to predicted tones (Figure 1.9). This increased response to mispredicted stimuli may index the prediction error when a strong prediction for a specific note proves to be inaccurate, demonstrating a consequence of failed predictions in a musical context. However, while this paradigm does rely on participants' implicit knowledge of the structure of a musical scale, the stimuli differed significantly from naturalistic music. Also, it is impossible to fully differentiate the influence

of participants' prior implicit musical knowledge from the effect of adaptation to the statistics of the local context over the course of the experiment, as the majority of trials consisted of scales leading to a correctly predicted fifth note.¹

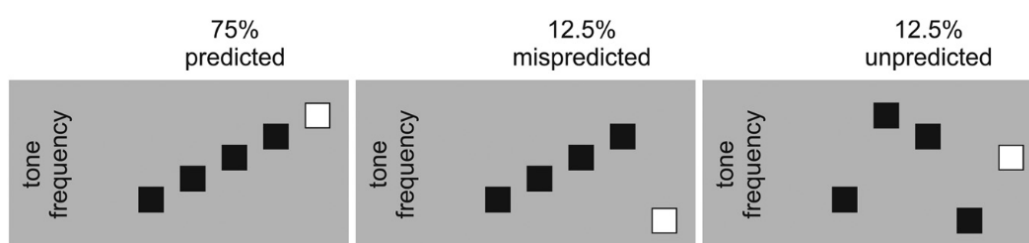


Figure 1.8. Experimental design used in Hsu et al. (2015). 75% of trials led participants to generate a correct prediction, 12.5% led to an incorrect prediction (misprediction), and 12.5% did not lead to any prediction.

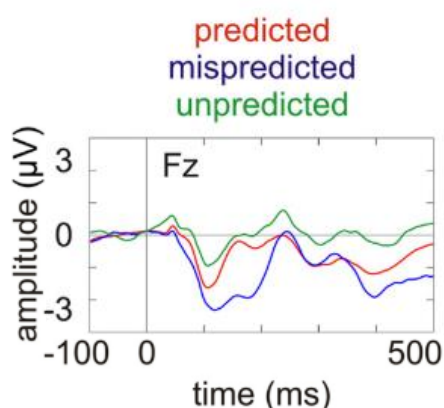


Figure 1.9. Results from Hsu et al. (2015); note that negative voltage is plotted down.

No ERP studies of melodic expectation to date have used designs comparable to those used in modern research on prediction in language, where contextual constraint is manipulated in naturalistic sequences in order to study responses to an unexpected target note when there is vs. is not a strong prediction

¹ For further discussion see Ross and Hansen (2016), who suggest it may be possible to differentiate the bases for participants' predictions by repeating the experiment using an unfamiliar tonal context such as the Bohlen-Pierce Scale (Loui, Wessel, & Kam, 2010).

for a different note. Designing ERP studies in this way can facilitate the comparative study of prediction in music and language, particularly if unexpected targets are not incongruous, just as unexpected linguistic target words are not incongruous in many modern neural studies of prediction in language (cf. Figure 1.7). Thus, we have developed a new approach to studying melodic expectation inspired by language research: manipulating contextual constraint, or the extent to which a context constrains expectations to one particular continuation. We have previously created a melodic cloze probability task as a new way of measuring expectations in melodies (Fogel, Rosenberg, Lehman, Kuperberg, & Patel, 2015). In the commonly used linguistic cloze probability task, participants complete sentences with the word they think comes next. In the musical analogue of this paradigm, participants are presented with the beginnings of novel tonal melodies (“melodic stems”) and asked to sing the note they expect to come next. The percentage of participants singing a certain response for a given melody is taken as the “cloze probability” of that response.

Monophonic melodies were composed for this task in pairs ($n = 45$), where they were matched in their length, rhythm, melodic contour, and key strength but differed in their underlying harmonic structure. Stems were 5-9 notes long (mean 8.38 notes), were played at 120 BPM, and had note durations that varied from eighth notes (250 ms) to half notes (1000 ms). They had no rests, variations in loudness, or non-diatonic pitches (see Fogel et al., 2015 for further details). One member of each pair had an implied harmonic structure designed to constrain participants’ expectations for a certain note (these stems ended with an

implied authentic cadence, which conventionally creates an expectation for the tonic) while the other ended with an implied IV, vi, or ii harmony, designed to not create an expectation for any particular continuation. The former stems formed the AC (“authentic cadence”) condition, while the latter formed the NC (“non cadence”) condition (see Figure 1.10 for an example melodic stem pair). As expected, participants’ responses were significantly more consistent following AC stems than NC stems (see Figure 1.11 for results for the melodies pictured in Figure 1.10). This shows that participants formed different predictions for what would come next based on different underlying harmonies in these melodies, establishing that naturalistic musical contexts can be designed to constrain participants’ expectations for a certain continuation.



Figure 1.10. Sample melodic stem pair used in Fogel et al. (2015). One stem in each pair (AC condition) had an underlying harmonic structure designed to constrain expectations for the next note to a specific continuation (the tonic) by ending with an implied authentic cadence; the other (NC condition) was matched in terms of length, rhythm, melodic contour, and key, but ended with an implied IV, vi, or ii harmony and was not designed to constrain expectations. Each participant heard only one version of each stem pair.

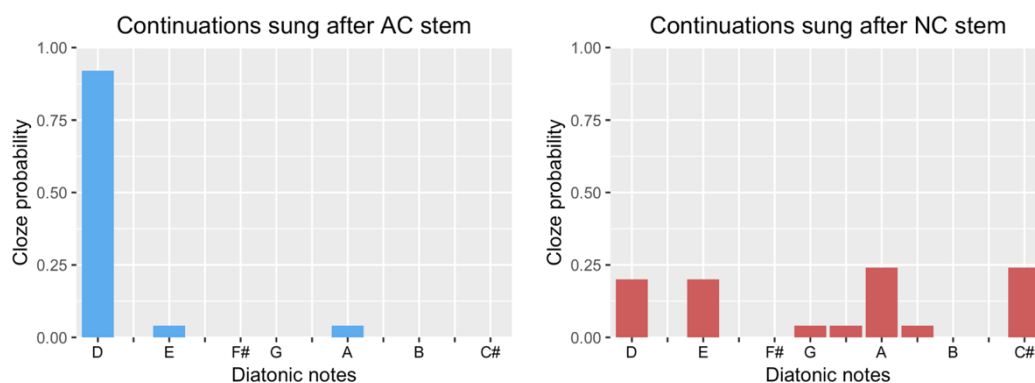


Figure 1.11. Distribution of participants’ sung continuations for the melodic stem pair pictured in Figure 1.10. Participants’ responses were significantly more consistent following this AC stem, where 92% of participants sang the tonic (D), than following the NC version, where no more than 24% of participants sang any one continuation.

Unlike other behavioral measures that have been used in previous studies of melodic expectation, this task produces quantitative measurements of the ‘predictive constraint’ of each melodic context; the larger the proportion of participants providing a given continuation for a melody, the more it has constrained expectations for that particular continuation. Using this method, we can develop melodic contexts that either do or do not lead to a strong prediction for a certain note to come next as well as quantify the degree to which they do so. Crucially, when a context leads to an expectation for one particular note, a different note can violate this prediction even while still being in-key and congruous.

Chapter 2 assesses the reliability of the melodic cloze paradigm as a method of quantifying melodic expectancy by examining the extent to which melodic cloze probabilities depend on participants and task instructions. In Chapter 3, stimuli developed through the cloze task are used to study the neural

correlates of melodic expectation in a novel way, by exploring the ERP effects of violating predictions in naturalistic melodies with in-key target notes. These notes are compared to the same notes in non-constraining melodies, where they do not violate any predictions. Chapter 4 extends the previous experiment using out-of-key target notes, facilitating comparisons to previous ERP studies of melodies. Chapter 5 discusses additional directions using this approach to studying melodic expectation, such as evaluating the accuracy of the predictions of different computational models of melodic expectation and examining brain activity during the buildup of melodic expectation.

Chapter 2. Reliability of the cloze probability method for measuring melodic expectation

Introduction

We have previously developed a melodic cloze probability task to facilitate studying expectation in melodies (Fogel et al., 2015). This method is modeled on the well-known linguistic cloze probability task, where participants complete sentences with the word they think comes next. In the musical paradigm, participants are presented with the beginnings of novel tonal melodies (“melodic stems”) and asked to sing the note they expect to come next; the percentage of participants singing a certain response for a given melody is taken as the “cloze probability” of that response. While a few previous studies have used sung continuations to probe melodic expectation, these have used very short contexts (e.g., 2 notes: Carlsen, 1981) or just a few melodies, and (unlike in language studies) the constraint of melodic stems has not been systematically quantified.

As explained in Chapter 1, for our melodic cloze task melodies were composed in pairs that differed in their underlying harmonic structure: one member of each pair had a harmonic structure designed to constrain participants’ expectations for a certain note (containing an implied authentic cadence, which conventionally creates an expectation for the tonic; AC condition) while the other ended in an implied harmony designed to not create an expectation for any particular continuation (NC condition). Participants formed different predictions for what would come next based on different underlying harmonies in these

melodies, establishing that naturalistic musical contexts can be designed to constrain participants' expectations for a certain continuation (see Figures 1.10 and 1.11, and associated text).

Chapters 3-5 of this thesis provide examples of some of the ways that the melodic cloze probability task can be used to conduct further experiments.

However, in order to be able to rely on this method as the basis for future studies, it is necessary to establish its reliability and validity. While the initial participants in Fogel et al. (2015) were remarkably consistent in their responses to AC melodies, there was much more variation in responses to NC melodies, which could potentially change given a different group of individual participants.

Additionally, while participants were instructed to *continue* the melody with a single note, they may have been naturally inclined to try to *complete* it instead by singing a “continuation” note that provided a sense of closure. Here, the original melodic cloze probability experiment (Fogel et al., 2015) was replicated 1) using a subset of identical melodic stems with a new group of participants and 2) with participants asked to sing multi-note continuations.

Method

Experiment 1

First, the original melodic cloze probability task (Fogel et al., 2015) was repeated, with participants asked to sing the note they expect to come next (for the full set of melodic stems, see Appendix B). In order to norm stimuli for further experiments, this experiment had an additional 12 melodic stem pairs added to the

original 48 pairs². Additionally, a number of melodic stems were edited by the original composer in minor ways (such as altering rhythm) that did not change their fundamental harmonic structure; these changes were intended to enhance the constraint differences between AC and NC stems. However, out of the 120 total melodic stems participants heard in this experiment, 40 stems remained completely identical to the versions used in the original task. These 23 AC stems and 17 NC stems were items that had sufficiently high or low constraints in the original experiment and did not require editing.

54 participants took part in this experiment; all were amateur musicians recruited from online postings or flyers aimed at the Tufts University community. As in the original melodic cloze study of Fogel et al. (2015), all participants had at least 5 years of musical experience within the past 10 years (playing an instrument, singing, or musical training). The mean age of participants was 23.7 years ($SD = 13.4$), and they reported having an average of 8.6 years ($SD = 1.9$) of musical experience within the last 10 years and a mean of 9.1 years ($SD = 6.8$) of total formal musical training. In Fogel et al. (2015) there were 50 participants (29 female, 21 male), with a mean age of 20.3 years, who had a mean of 9 years of formal musical training. Thus the participants in the two studies were comparable in age, mix of genders, and musical training.

As in the original study, participants were instructed that they would hear the beginning of a melody and would need to “sing the note you think comes

² Three pairs were removed from analysis due to differences in the melodic contours of the two stems, hence data from 45 pairs was analyzed in Chapter 1.

next” to continue (not necessarily complete) the melody. The mean fundamental frequency of each sung response was measured in Praat (Boersma, 2001) and rounded to the nearest semitone for analysis.

Experiment 2

Experiment 2 used the exact set of 45 melodic stem pairs used in Fogel et al. (2015). However, in this experiment, participants were instructed to *complete* the melody by singing *multiple* notes (“complete it by singing up to a few notes”). Cloze probabilities were computed using only the note sung most immediately after the end of a stem. As this was not the last note a participant sang in response to a melody, it was less likely to be affected by the natural tendency to provide a response that would create a sense of closure.

50 participants took part in Experiment 2; all were amateur musicians recruited from online postings or flyers aimed at the Tufts University community. Once again, all participants had at least 5 years of musical experience within the past 10 years (playing an instrument, singing, or musical training). Participants (mean age = 20.7 years, $SD = 1.5$) reported having an average of 9.1 years ($SD = 1.5$) of musical experience within the last 10 years and a mean of 9.3 years ($SD = 4.9$) of total formal musical training (see Table 2.1). They were instructed that they would hear the beginning of a melody and would need to “sing up to a few notes” to complete it. For each response, the mean fundamental frequency of the note sung most immediately after the end of the stem was measured using a custom MATLAB script and rounded to the nearest semitone for analysis.

Aside from the changes noted above, all other aspects of the design and procedure of Experiments 1 and 2 remained identical to Fogel et al. (2015).

Table 2.1. Participant characteristics.

	Total n =	Female/ Male n =	Mean age (years)	Mean years of formal musical training	Vocal performance experience
Original: Fogel et al., 2015	50	29/21	20.3	9.0 (<i>SD</i> = 4.8)	22 (44%)
Replication: Experiment 2	50	26/24	20.7	9.3 (<i>SD</i> = 4.9)	36 (72%)

Results

Experiment 1 Results

Participants' sung responses to the 40 stems that were identical to Fogel et al. (2015) were very close to those from the original study; AC stems were most frequently continued with the tonic (see Figure 2.1 for scale degree distributions), while sung responses to NC stems varied much more across in-key scale degrees. No significant differences were observed between experiments for any scale degree in either condition (all *ps* > 0.09).

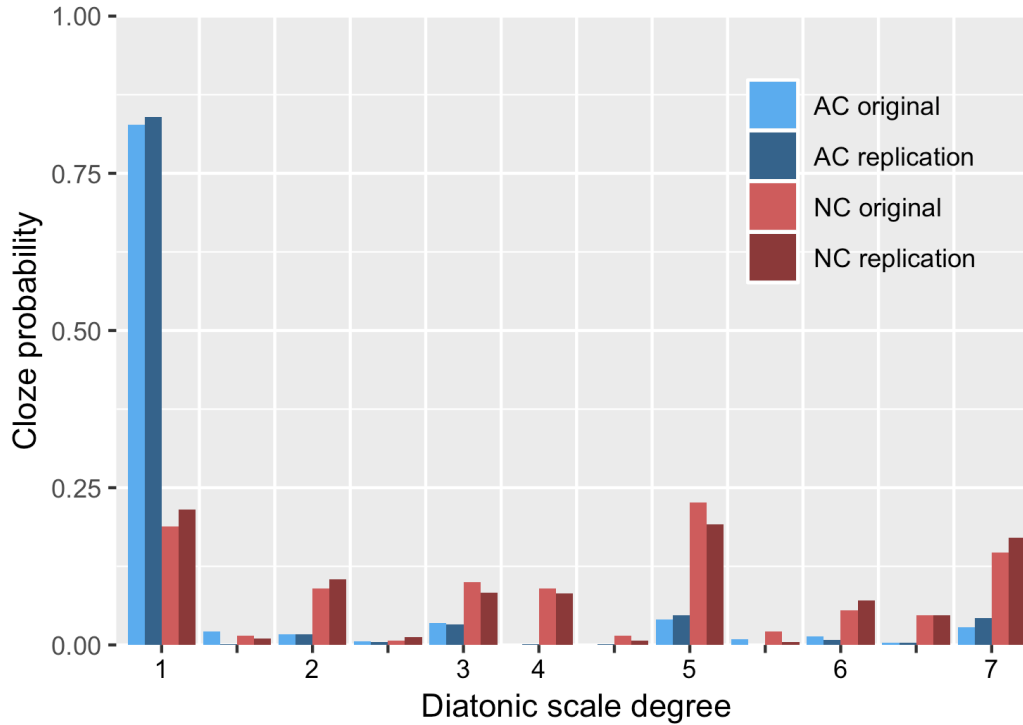


Figure 2.1. Scale degree distribution of participants' sung continuations after the same melodic stems in the original Fogel et al. (2015) melodic cloze probability experiment and in the present replication. No significant differences were observed between experiments for any scale degree in either condition (all p s > 0.09).

The 'predictive constraint' of a stem was defined as the percentage of participants who provided the most common response for a given stem; the more a melody restricts participants' expectations to a certain continuation, the higher its constraint. Overall, the constraints of individual melodic stems were very strongly positively correlated between studies, $r(38) = 0.94$, $p < 0.001$ (see Figure 2.2). The mean constraint of AC stems ($M = 0.84$, $SD = 0.09$) did not differ from the original study ($M = 0.83$, $SD = 0.09$), $t(22) = 0.56$, $p = 0.58$. The average constraint of NC stems ($M = 0.33$, $SD = 0.07$) was slightly higher than in the original study ($M = 0.28$, $SD = 0.07$), but this was not a significant difference, $t(16) = 2.03$, $p = 0.06$.

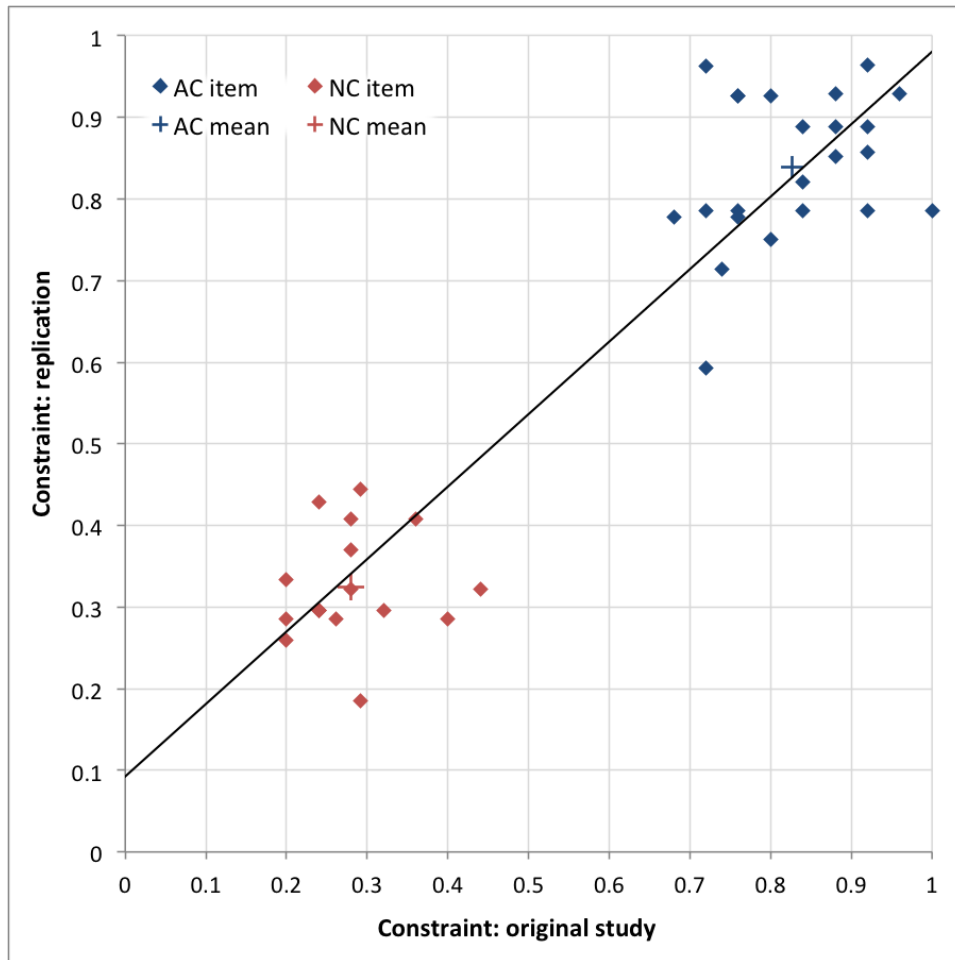


Figure 2.2. The constraints of individual melodic stems in the original melodic cloze probability experiment compared to the present replication. A very strong positive correlation was observed, $r(38) = 0.94$, $p < 0.001$.

A distance metric was created for each melodic stem to quantify how much its cloze probability distribution differed between studies. This was computed by taking the absolute value of the difference in cloze probability for each possible response, summing these values, and dividing by two. This results in a value between 0 and 1 for each item, representing the amount of probability mass that would need to be moved in order to turn one distribution into the other (where higher numbers represent a greater difference in probability distributions). Across all items, the average distance was 0.21, indicating that on average about

21% of the cloze probability distribution for each stem differed between studies. The average distance was smaller for AC stems ($M = 0.15$) than for NC stems ($M = 0.28$), demonstrating that participants' responses to AC stems remained more consistent between experiments, $t(44) = 2.95$, $p = 0.005$ (see Figure 2.3 for distances of individual items and Figure 2.4 for examples of response distributions for individual stems).

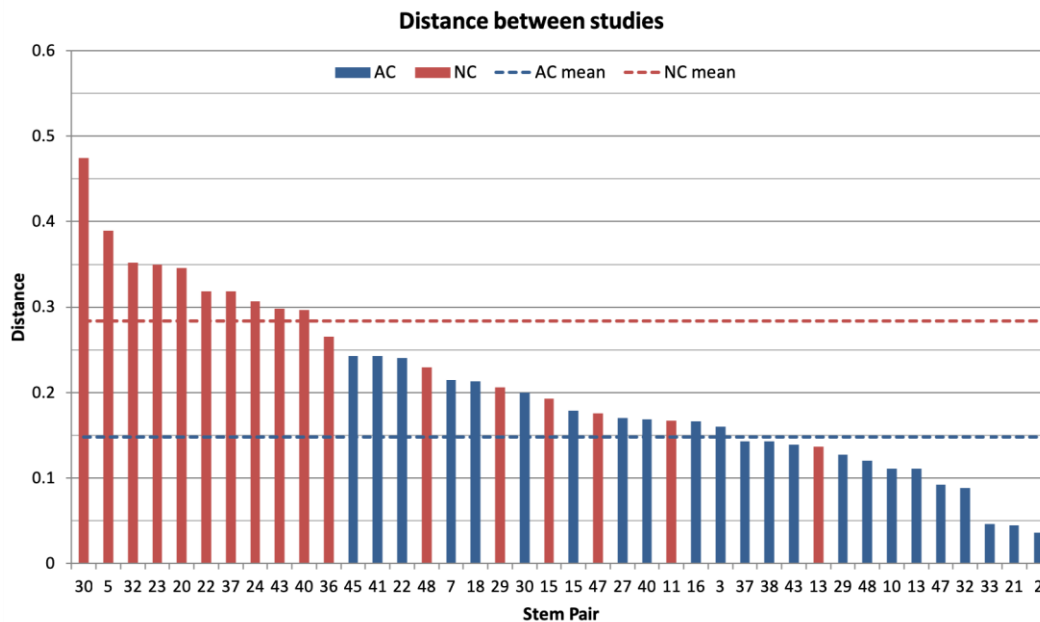


Figure 2.3. The distance between the cloze probability distributions for the 40 stems studied in Fogel et al. (2015) and in the replication experiment, sorted in order of decreasing distance. On average, responses to AC stems differed less between experiments ($M = 0.15$) than responses to NC stems ($M = 0.28$). Melodies are labeled by number as they appear in Appendix B.

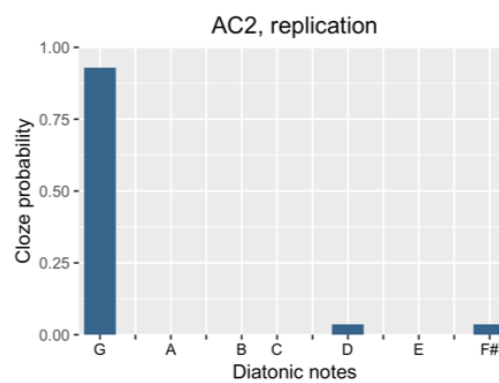
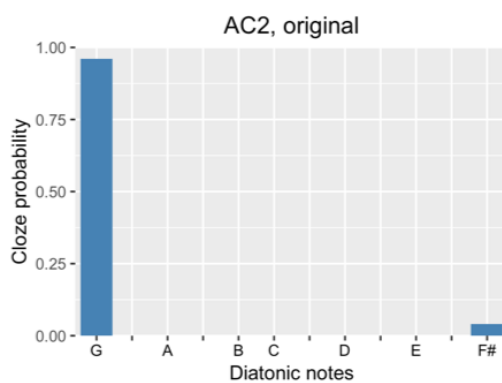
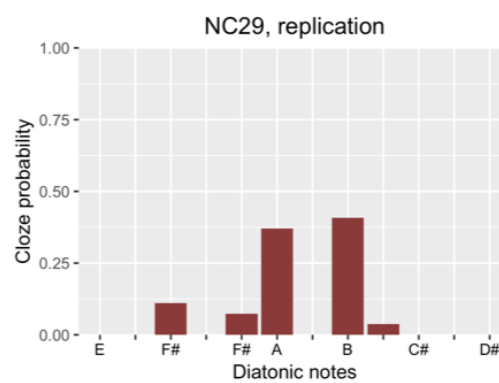
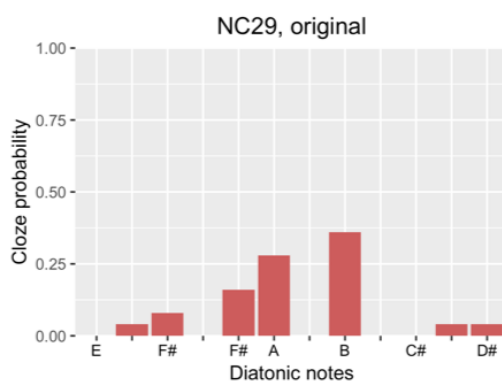
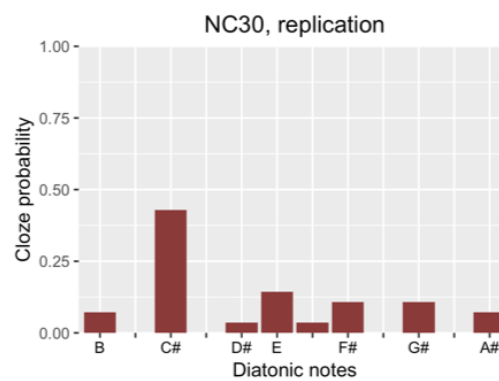
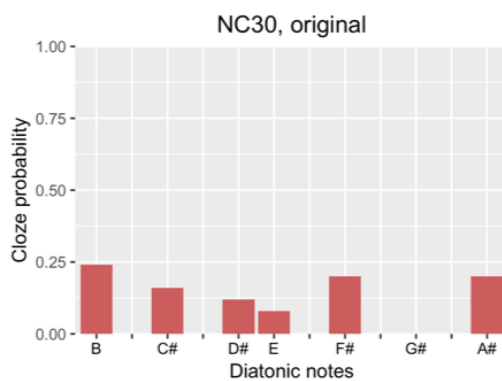


Figure 2.4. Participants' responses to selected melodic stems in the two experiments.

- a. NC melodic stem #30 had the greatest distance between its two response distributions (0.47).
- b. The distance between response distributions for NC stem #29 (0.21) was closest to the mean for all melodic stems.
- c. AC stem #2 had the least distance between its two response distributions (0.04).

Experiment 2 Results

In this experiment, which used the same 45 stem pairs as Fogel et al. (2015), participants were asked to complete the melody by singing *multiple* notes (“complete it by singing up to a few notes”). Participants sang 1-16 notes after the end of melodic stems; responses were an average of 4.44 notes ($SD = 2.73$) long. Responses consisting of a single sung note were the most common length for AC stems (31.0% of responses), while they were only 13.4% of responses to NC stems (see Figure 2.5). In contrast, the most common length of responses to NC stems was five notes (16.3% of responses). When the average lengths of responses to each stem were compared, responses to AC stems ($M = 4.10$ notes, $SD = 1.06$) were on average shorter than responses to NC stems ($M = 4.78$, $SD = 1.00$), $t(44) = 7.34$, $p < 0.0001$. The lengths of the sung continuations for AC and NC versions of a stem pair were positively correlated, $r(44) = 0.818$, $p < 0.0001$; that is, 66.9% of the variance in the mean length of the responses to an item was explained by the identity of the matched pair (see Figure 2.6).

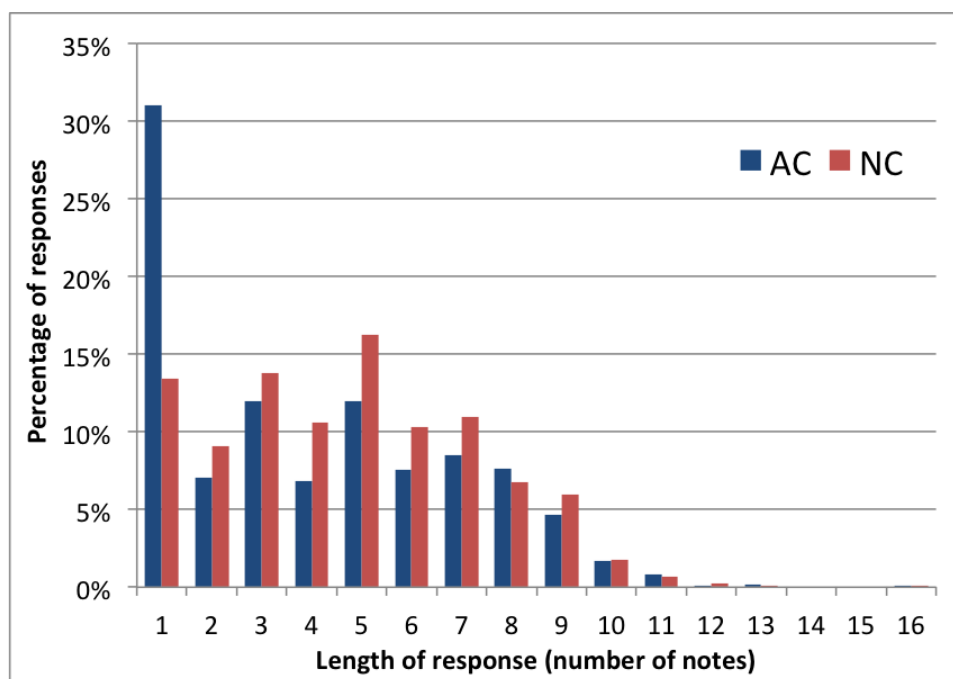


Figure 2.5. Lengths of responses in each condition of Experiment 2. More AC responses consisted of a single note (31.0%) than NC responses (13.4%); the most common length of responses to NC stems was five notes (16.3% of responses).

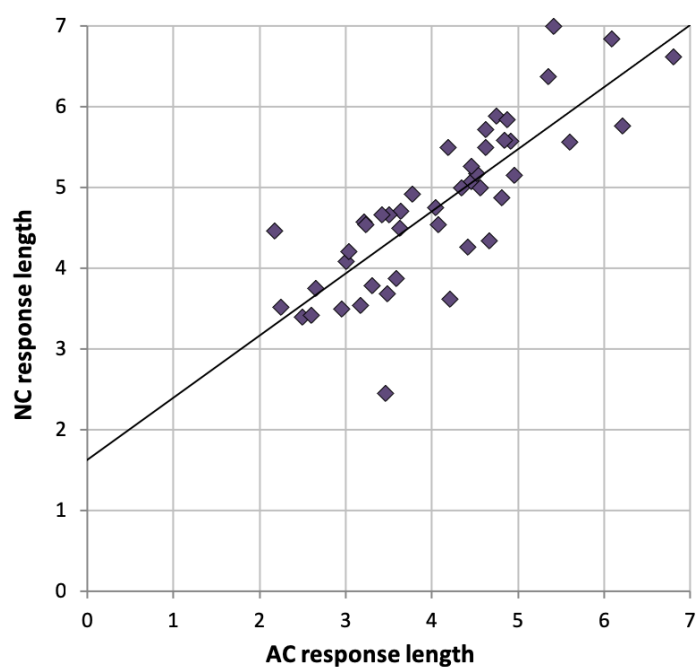


Figure 2.6. Average length of responses to the AC and NC versions of each melodic stem pair. Responses to AC stems ($M = 4.10$ notes) were on average

shorter than responses to NC stems ($M = 4.78$), but the average lengths of the AC and NC versions of a stem pair were positively correlated, $r(44) = 0.818$, $p < 0.0001$.

The distribution of participants' sung responses remained generally similar to the original task. AC stems were again most frequently continued with the tonic ($M = 0.62$, $SD = 0.18$), while responses to NC stems were more widely distributed across in-key scale degrees (tonic $M = 0.13$, $SD = 0.10$; see Figure 2.7 for scale degree distributions). For AC stems, the cloze probabilities of scale degrees 1, 2, 3, and 7 differed significantly from their values in the original study; the cloze probability of the tonic was lower in this task ($M = 0.62$, $SD = 0.18$) than in the original ($M = 0.69$, $SD = 0.17$; $t(44) = 3.87$, $p < 0.001$), while the probabilities of scale degrees 2, 3, and 7 were slightly higher ($t(44) = 2.89$, $p = 0.006$; $t(44) = 2.43$, $p = 0.02$; $t(44) = 2.39$, $p = 0.02$). No significant differences were observed between experiments for any NC scale degree (all $ps > 0.06$).

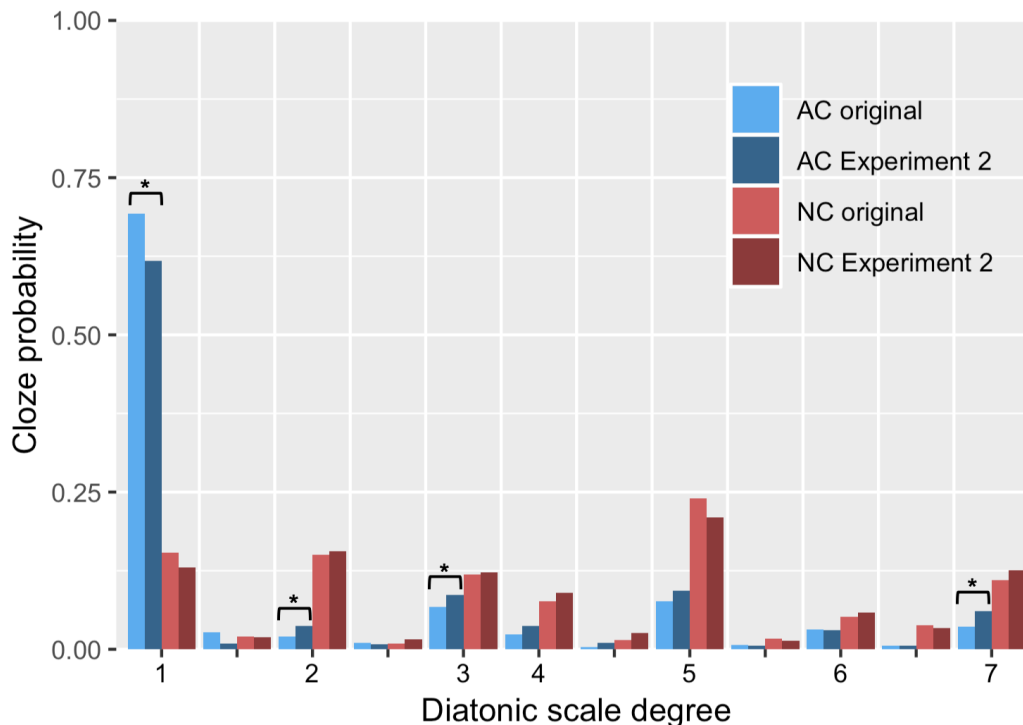


Figure 2.7. Scale degree distribution of participants' sung responses after melodic stems in the original melodic cloze probability study and in the present experiment where they sang longer continuations. For AC stems, the cloze probabilities of scale degrees 1, 2, 3, and 7 differed significantly between experiments (significant differences marked with asterisk), while no significant differences were observed between experiments for NC stems.

The constraints of individual melodic stems were strongly positively correlated between studies, $r(88) = 0.82$, $p < .001$ (see Figure 2.8). However, the mean constraint of AC stems was significantly lower in this experiment ($M = 0.63$, $SD = 0.16$) than in the original study ($M = 0.69$, $SD = 0.17$), $t(44) = 3.48$, $p = 0.001$. In contrast, the mean constraint of NC stems ($M = 0.40$, $SD = 0.14$) did not differ from the initial study ($M = 0.42$, $SD = 0.15$), $t(44) = 0.68$, $p = 0.50$. As in the original study, there was considerable variation among matched stem pairs in the degree to which the AC and NC versions of a given pair differed in constraint (see Figure 2.9). Though the constraint of AC stems went down overall in this experiment, many pairs display similar patterns when their constraints are compared; for example, the NC version of stem pair #27 has an unusually high constraint compared to the AC version in both experiments.

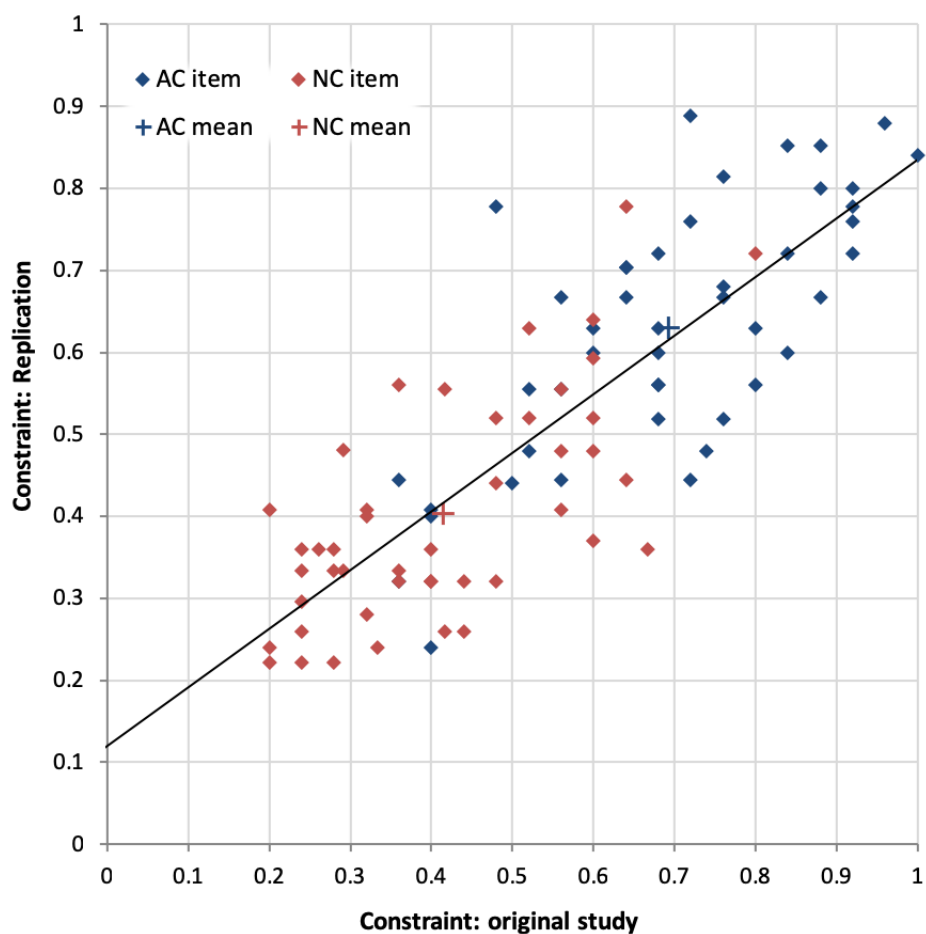


Figure 2.8. The constraints of individual melodic stems in the original melodic cloze probability experiment compared to their constraints in the present experiment where participants sang longer continuations. A strong positive correlation was observed, $r(88) = 0.82$, $p < .001$.

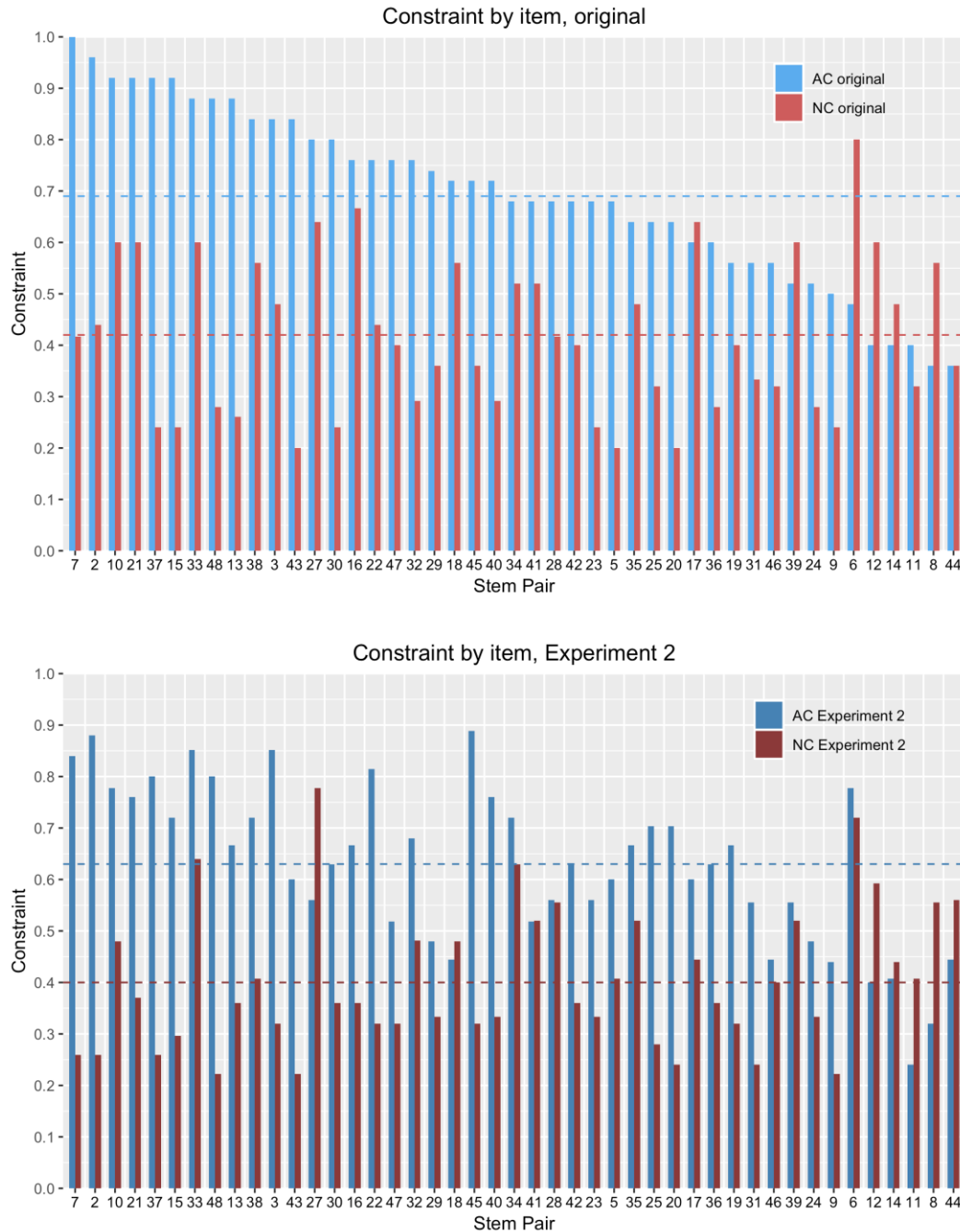


Figure 2.9. Predictive constraint (the percentage of participants who sang the most common response) of matched stem pairs across the two studies. The mean constraint of AC stems ($M = 0.63$, $SD = 0.16$) was significantly lower than in the original study ($M = 0.69$, $SD = 0.17$; $t(44) = 3.48$, $p = 0.001$), while the mean constraint of NC stems ($M = 0.40$, $SD = 0.14$) did not differ from the original ($M = 0.42$, $SD = 0.15$), $t(44) = 0.68$, $p = 0.50$. In the upper graph, data have been sorted by descending order of constraint in AC members of each pair. The resulting order of stems is preserved in the lower figure.

A distance metric was computed as described for Experiment 1, resulting in a value between 0 and 1 for each item representing the amount of probability mass that would need to be moved in order to turn one distribution into the other. Across all items, the average distance was 0.25, indicating that on average about 25% of the cloze probability distribution for each stem differed between studies. Participants' responses to AC stems remained more consistent between experiments; the average distance was smaller for AC stems ($M = 0.22$) than for NC stems ($M = 0.28$), $t(44) = 2.95$, $p = .005$. See Figure 2.10 for values for individual items and Figure 2.11 for examples of response distributions for individual stems.

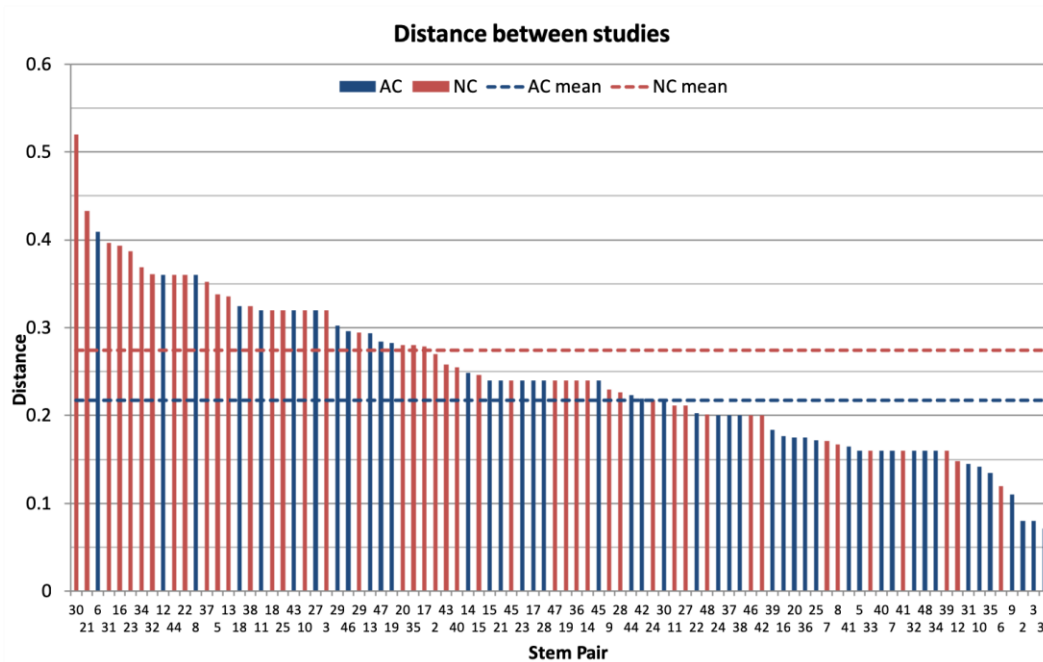


Figure 2.10. The distance between cloze probability distributions in the two studies for each individual item. On average, responses to AC stems differed less ($M = 0.22$) than responses to NC stems ($M = 0.28$).

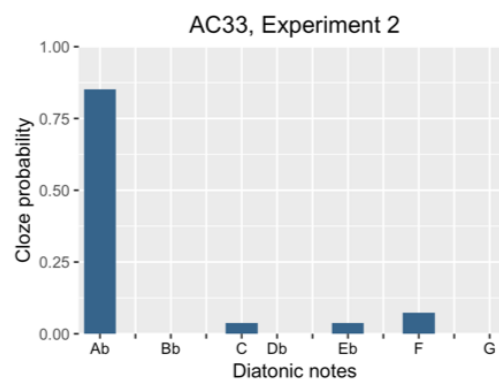
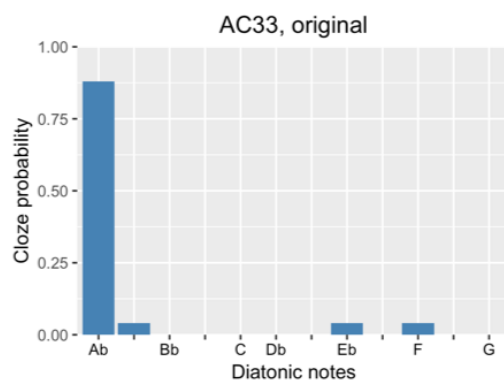
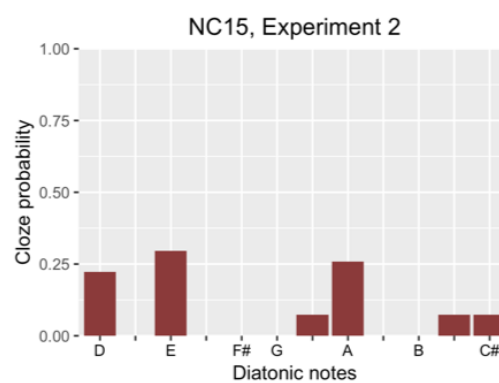
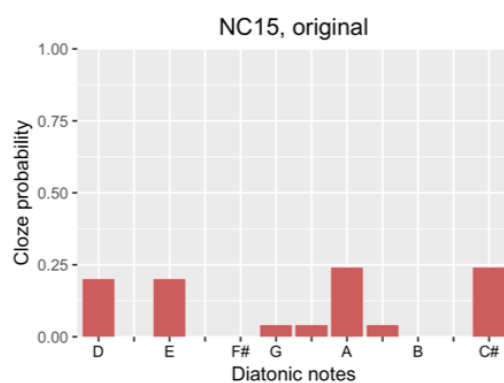
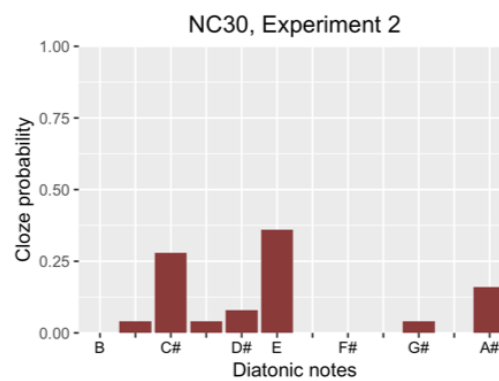
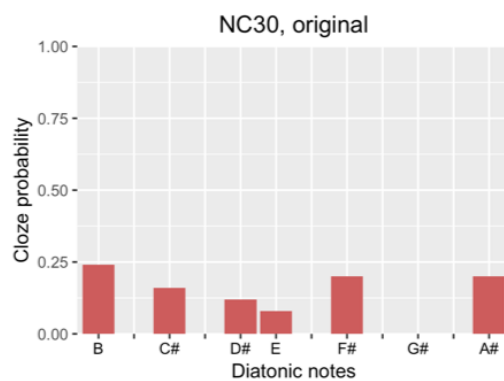


Figure 2.11. Participants' responses to selected melodic stems in the two experiments.

- a. NC melodic stem #30 had the greatest distance between its two response distributions (0.52).
- b. The distance between response distributions for NC stem #15 (0.25) was closest to the mean for all melodic stems.
- c. AC stem #33 had the least distance between its two response distributions (0.07).

Discussion

In general, cloze probabilities remained similar between the original cloze probability task and each of the two experiments. In Experiment 1, the original experiment was replicated exactly for a subset of 40 melodic stems. As expected, a very strong positive correlation was observed between items' constraints in the two experiments, and no significant differences were observed between experiments for the average cloze probability of any scale degree in either condition. However, participants' responses to AC stems were more consistent between experiments than NC stems; the average distance between AC response distributions was smaller, and the mean constraint of AC stems remained the same. AC stems were originally designed to constrain expectations to a particular continuation (the tonic), and all 40 melodic stems used for this replication were selected because they already had high or low constraints as intended. It is therefore unsurprising that melodies that had previously constrained participants' expectations well continued to do so here. NC stems were designed to not lead participants to expect any particular continuation, so responses naturally varied more. The larger distance between NC response distributions in the two

experiments demonstrates that they truly did not constrain expectations for any particular continuation. It is possible in theory for NC items to have a fairly low constraint and still lead participants to expect two or three particular notes, but if this were the case participants' responses would likely be more consistent between experiments.

In Experiment 2, participants sung longer responses to complete melodic stems instead of being restricted to singing one-note continuations. When the results of this task were compared to the original study, the average constraint of AC items (as well as the cloze probability of the tonic) went down, while the average constraint of NC items was unchanged. It seems that some of the participants' responses in the initial study were indeed influenced by a natural tendency to try to complete a melody, which was no longer a factor when the note sung immediately after the end of a stem did not need to be the last note of the entire sequence. Despite these differences, responses were still quite similar in general: the reduction in the mean constraint of AC stems (from 69% to 63%) was not dramatic, and the overall correlation between items' constraints in the two experiments remained strong ($r = 0.82$). Significant differences in the cloze probabilities of scale degrees other than the tonic (2, 3, and 7) were likely a result of the reduction in the cloze probability of the tonic, as responses shifted to other notes. Even for the scale degrees where significant differences were found in experiment 2, the numerical differences in constraint were not dramatic.

Though the average constraint of AC melodic stems differed more between Fogel et al. (2015) and Experiment 2 than did the average constraint of

NC stems, the mean distance between items' response distributions in the two experiments was higher for NC items. As in Experiment 1, it appears that NC stems truly did not constrain expectations for any particular continuation. It is notable that this pattern was still observed in Experiment 2, which used the full set of 45 pairs of melodic stems as opposed to the subset that had been selected as good examples of their conditions for Experiment 1. In fact, the average distance between NC stems' distributions in the original study and Experiment 1 (0.284) was nearly identical to that between the original and Experiment 2 (0.275). Though the mean constraint of AC items was lower in Experiment 2 than in the original study, these stems continued to largely constrain participants' responses to the tonic, so the distribution of responses across scale degrees remained relatively consistent.

Overall, the results of Experiments 1 and 2 demonstrate the reliability of the melodic cloze probability task as a method for determining participants' expectations in melodic sequences. The most notable difference observed here was the reduction in the constraint of AC stems from 69% to 63% when participants were able to sing longer continuations in Experiment 2, implying that about 7% of responses in the original task may have been influenced by a tendency to try to complete the melody instead of reflecting unbiased expectations. However, even when participants were given the opportunity to sing multiple notes, 31% of responses to AC stems and 13% of responses to NC stems still consisted of a single note, demonstrating that in many cases participants truly expected the melodies to end at this point. This likely accounts for the fact that the

two sets of results did not differ dramatically. When choosing which form of the melodic cloze probability task to use, researchers should select the version that is most relevant for their purposes. The task in Experiment 2 left participants free to sing single or multiple note responses, but when norming data for different future experiments it may be desirable to obtain only expectations for mid-sequence notes (requiring participants to always sing multiple notes) or for sequence-final notes (requiring participants to always sing a single note response).

There are numerous uses for the melodic cloze probability task in studying melodic expectation, some of which are demonstrated in this thesis (see Fogel et al., 2015 for further discussion). For example, cloze probabilities obtained from this task can be used to test the performance of computational models of melodic expectation in terms of their ability to predict human responses. Another important use of the task is in designing stimuli for neural studies of melodic expectation that are more analogous to current research on prediction in language. The ERP studies of melodic expectation in Chapters 3 and 4 use stimuli normed using the cloze probability paradigm.

Chapter 3. Neural correlates of processing in-key melodic prediction violations

Introduction

ERPs have been an important tool for the study of prediction in language. By manipulating the predictive constraint of sentences, language researchers have been able to observe the neural effects of violating strong lexical predictions with plausible but unexpected (low-probability) words. As discussed in Chapter 1, when the brain response to such words is compared to the brain response to similar low-probability words in sentences that do not lead to a strong lexical prediction for another word, these violations elicit a late anteriorly distributed positive component, typically peaking around 600-900 ms after the presentation of a critical item (Federmeier et al., 2007; Van Petten & Luka, 2012b).

While a number of ERP studies of music processing have examined the effects of violating melodic expectations, manipulations have mainly used unexpected target notes in a fixed melodic context. In contrast to the plausible and coherent violations of specific predictions used in language studies investigating the frontal positivity, the target notes in music studies have generally been out-of-key or otherwise harmonically incongruous in some noticeable way (e.g., Koelsch, Gunter, Friederici, & Schröger, 2000; Koelsch, Jentschke, Sammler, & Mietchen, 2007); such violations can sound unnatural (like a “wrong note”) and may trigger processes concerned with error detection (like frank anomalies in language). Meanwhile, in ERP studies where target notes have been in-key, target notes have been described simply as low- or high-probability without regard for

the constraint of the prior context (e.g., Pearce, Ruiz, Kapasi, Wiggins, & Bhattacharya, 2010). No ERP study has yet investigated the effects of prediction in music processing by manipulating the contextual constraint of sequences and using plausible, in-key target notes as targets.

The present study was modeled after the ERP paradigms used in language experiments (e.g., Federmeier et al., 2007) where contextual constraint can be manipulated independently of the expectedness of a target item. By using the melodic cloze paradigm to measure the predictive constraint of melodies, stimuli were developed that vary in the extent to which they lead participants to expect a certain target note. ERPs were used to examine neural responses to target notes that are equally unexpected (i.e. equally unlikely given the prior melodic context) in high- and low-constraint contexts, but still in-key and melodically coherent. In high-constraint contexts, strong expectations are generated for a certain continuation, so these unexpected (low-probability) notes violate a strong prediction. In contrast, low constraint contexts do not lead a listener to expect any specific note, so unexpected (low-probability) notes do not violate any particular expectation. Measuring the difference between the ERPs in response to unexpected target notes in different constraint conditions can reveal the neural correlates of violating a strong expectation for a certain note in a melodic context.

Previous neural studies of melodic expectation have typically compared unexpected (usually out-of-key) notes to expected notes in a fixed melodic context, and have frequently reported that this contrast elicits an early right anterior negativity (ERAN) in various time windows from 100-300 ms after the

presentation of a target note (see Table 1.1). However, because contextual constraint has not been measured or manipulated in these experiments, it cannot be determined from their results whether the early anterior negativity represents the effect of *violating* a prediction for a specific note with a different note (like the frontal positivity effect in language) or the effect of processing a low probability note regardless of prediction (like the N400 in language). Nevertheless, current evidence suggests that the early anterior negativity probably does not represent a misprediction effect. It has been shown to not change over repeated presentations of the same stimuli or based on participants' veridical knowledge (Koelsch et al., 2019); its amplitude also remains the same when participants are told in advance whether a music-syntactic irregularity will occur, though its peak latency may be slightly earlier (by ~10 ms) in this circumstance (Guo & Koelsch, 2016; Koelsch et al., 2019). Additionally, Vuvan, Zendel, and Peretz (2018) reported that the amplitude of the ERAN *increased* when participants were instructed to make anomaly judgements and provided with random feedback about the accuracy of their performance, which lowered their confidence in their predictions. It therefore seems unlikely that unexpected target notes in constraining conditions, which violate a prediction, will elicit an early anterior negativity relative to unexpected notes in non-constraining conditions.

Instead, we predicted that this comparison would result in a frontal positivity, as it does in language stimuli. This prediction was based on the idea that melodies, like sentences, involve structured sequences than can lead one to commit to a particular upcoming representation (i.e., form a strong prediction for

a particular note, just as a sentence can lead one to strongly predict a particular upcoming word). If a different, but plausible element appears, then activity associated with the original prediction (e.g., preactivation of the predicted element) must be suppressed in favor of the actual element delivered, or one's internal model of the sequence must be updated. Such suppression or updating is not needed when the same unexpected (low-probability) element occurs after a context that does not lead to a strong prediction for a different element. In other words, the cognitive and neural mechanisms involved in handling a plausible but mispredicted element in a sequence could overlap for music and language, even though the underlying representations on which these operations act might be distinct. This idea of music and language engaging overlapping processing mechanisms operating on separable and domain-specific representations is part of the “resource sharing” theoretical framework for music and language processing (Patel, 2012, 2013), which has been extensively explored in the case of syntactic processing (Fedorenko, Patel, Casasanto, Winawer, & Gibson, 2009; Kunert, Willems, Casasanto, Patel, & Hagoort, 2015; Patel, 2003; Slevc, Rosenberg, & Patel, 2009). While the current study does not focus on similarities between music and linguistic syntactic processing, we use the same “resource sharing” framework to predict that the ERP response to mispredicted notes in high-constraint melodies compared to the same notes in low-constraint melodic stems will resemble the frontal positivity seen in language studies of similar design.

It is more difficult to precisely predict what effects will be seen in the comparison between expected and unexpected target notes within constraining

contexts. Based on language studies, we anticipate that unexpected (mispredicted) notes might elicit a frontal positivity relative to expected target notes (see Figure 1.7). Additionally, Pearce et al. (2010) reported that a frontal positivity was elicited by (in-key) low-probability notes compared to high-probability notes; however, we cannot determine if their low-probability notes were mispredicted (in constraining contexts) or unpredicted (in non-constraining contexts). Despite the methodological differences between this and previous ERP studies of melodies (e.g. task and violation type), it is also possible that this contrast might elicit the early anterior negativity that has frequently been reported to be elicited by unexpected compared to expected notes.

Methods

Stimulus Construction

Stimuli were developed using the melodic cloze probability paradigm, in which participants are presented with the beginnings of novel tonal melodies (“melodic stems”) and asked to sing the note they expect to come next; the percentage of participants singing a certain response for a given melody is taken as the “cloze probability” of that response, while the percentage of participants singing the most common response for a given melody is taken as the “constraint” of that melody (Fogel et al., 2015). Melodic stems were composed in pairs, where they were matched in their length, rhythm, melodic contour, and key strength but differed in their underlying harmonic structure: one member of each pair had a harmonic structure designed to constrain participants’ expectations for a certain note (containing an implied authentic cadence, AC condition) while the other was

designed to not create an expectation for any particular continuation (ending with an implied IV, vi, or ii harmony, NC condition); see **Figure 3.1** for an example melodic stem pair and **Figure 3.2** for an example of participants' responses for this item. The 60 pairs of 5-9 note melodic stems used to construct stimuli for the present study were normed by 54 participants following the procedure described in Fogel et al. (2015); participants' sung responses were recorded, then the mean fundamental frequency of each sung note was extracted and rounded to the nearest semitone.



Figure 3.1. Sample melodic stem pair. One stem in each pair (AC condition) has an underlying harmonic structure designed to constrain expectations for the next note to a specific continuation (the tonic) by ending with an implied authentic cadence; the other (NC condition) is matched in terms of length, rhythm, melodic contour, and key, but ends with an implied IV, vi, or ii harmony and is not designed to constrain expectations. Each participant heard only one version of each stem pair.

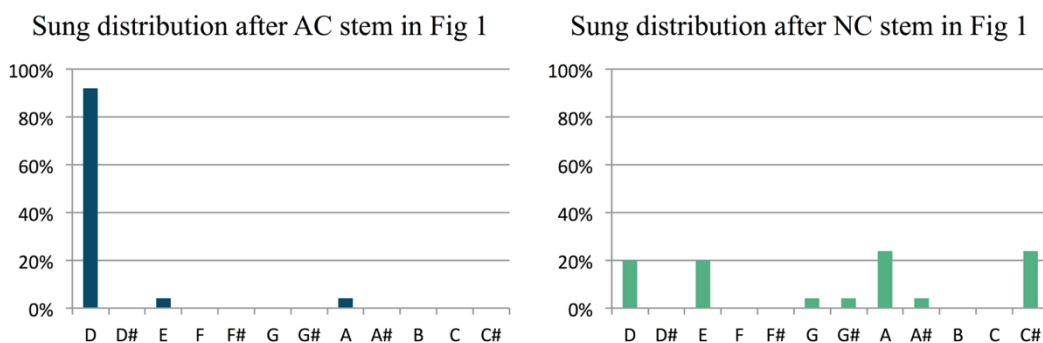


Figure 3.2. Distribution of participants' sung continuations for the melodic stem pair pictured in Figure 3.1. Participants' responses were significantly more

consistent following this AC stem, where 92% of participants sang the tonic (D), than following the NC version, where no more than 24% of participants sang any one continuation.

ERP Stimuli

The stimuli used in the present experiment consisted of sixty pairs of novel 11-16 note melodies ($M = 14.0$ notes), which were in various major keys and contained no out-of-key notes or other melodic or harmonic incongruities. These full melodies continued after the “melodic stem” portions used for norming and were created by the same composer (Jason Rosenberg). All melodies were presented using a piano timbre generated by Finale software at 120 BPM and lasted 4.0-10.7 seconds ($M = 7.4$ seconds, $SD = 1.5$). Note durations ranged from eighth notes (250 ms) to half notes (1000 ms) throughout the melodies, with all critical target notes being quarter notes (500 ms); no melody contained any rests, intensity variations, or articulation variations (e.g., staccato vs. legato).

Each melodic pair had one version that constrained participants' expectations, leading to a strong mid-sequence expectation for the tonic, which was the highly expected target note. For the less expected condition, the target notes in these same melodies were replaced with the sixth scale degree, which, while not particularly expected, is not incongruous and sounds melodically plausible. (In the case where the 6th scale degree occurs at the end of an AC stem, this is similar to a “deceptive cadence” in tonal music, where a vi chord follows a cadential V chord). In the versions of these melodies that do not constrain expectations to a strong degree, the non-expected target note was the sixth scale degree as well (see **Figure 3.3** for example stimuli). In each of the two blocks of

the experiment, an individual participant heard a given melody pair in only one of these three conditions (i.e., 20 melodies per condition per subject), as well as 30 filler melodies of a comparable length and style that did not contain a mid-sequence cadence, for a total of 90 melodies in each section. No experimental melody was repeated in the same condition between blocks.

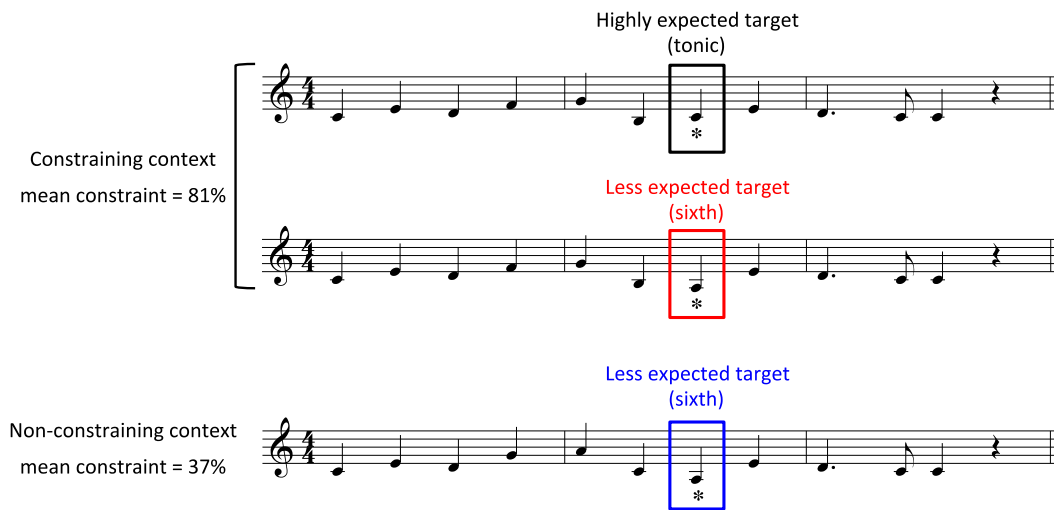


Figure 3.3. Example ERP stimuli. The constraining melody (AC stem) constrains participants' expectations for a certain continuation, the tonic, while the non-constraining melody (NC stem) does not. Both stems have the sixth scale degree (A) at the target position as the unexpected note (shown with boxes). This note is equally unexpected after both stems (with 0% cloze probability in both conditions for this melodic pair), but is in-key and perfectly “coherent” in both cases. Because the constraining stem constrains expectations for the tonic, the sixth scale degree in this condition violates this expectation, but the same note does not violate any expectations in the non-constraining condition.

Participants

Participants were recruited from online postings aimed at the Tufts University community. As our melodies were normed using musicians in the melodic cloze probability task, thirty-six self-identified musicians took part in the experiment and were included in the data analysis (sample size was determined

based on a power analysis using prior language data from Fogel, Wlotko, Kuperberg, & Patel, 2017; an additional 18 participants were excluded due to excessive artifact). Participants had at least five years of active musical experience within the last ten years ($M = 8.8$ years, $SD = 1.7$) and an average of 8.2 years of lifetime formal musical training ($SD = 3.9$, ranging from 0-14 years). All participants were between the ages of 18 and 35 ($M = 19.9$, $SD = 3.3$), right-handed, had normal hearing, and reported no history of neurological disorders or head trauma. They provided informed consent in accordance with the procedures of the Institutional Review Board of Tufts University and were compensated for their time.

Procedure

Participants were seated in a comfortable chair in a dimly lit room and provided with headphones. Trials were self-paced; each melody played only once a participant pressed a button to advance. Forty of the 90 melodies (44%) were followed by a memory test, in which participants heard a short excerpt (1-4 seconds) from a melody and indicated on a keyboard if the excerpt had appeared in the melody they just heard. Thus the behavioral task was orthogonal to the manipulation of interest (i.e., the attention of participants was not drawn toward unexpected notes, nor were such notes task relevant.)

ERP data 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

(**Figure 3.4**). 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. 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.

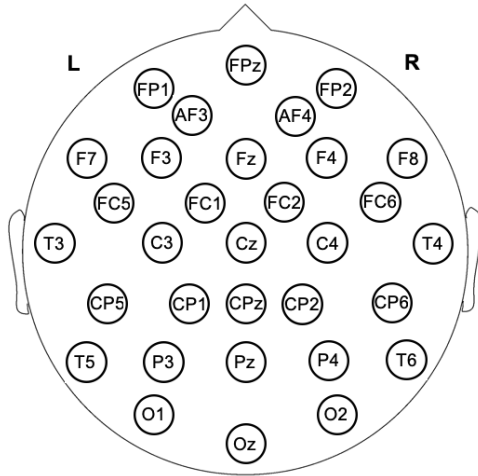


Figure 3.4. Electrode layout used in Chapters 3 and 4.

EEG and ERP data processing was conducted in EEGLAB v14.1.2 and ERPLAB v7.0.0.0. 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. It was then segmented into epochs from 300 ms before to 1100 ms after the onset of each target note, using the interval from -250 to 0 ms for baseline correction. Algorithms in ERPLAB were used to detect trials containing blinks, eye movements, and other artifacts within the time window used for analysis (-250 to 1000 ms); after artifact rejection, an average of 14.6% of critical trials were excluded per participant ($SD = 8.6$).

Analysis

For categorical analysis, melodies were divided into higher- and lower-constraint groups according to their empirical constraint based on the results of the melodic cloze probability paradigm, rather than by their original theoretical construction as AC or NC stems as discussed above (see **Figure 3.5**).

Additionally, five melodies where the sixth scale degree was in fact the most expected note (based on melodic cloze results) were excluded from analysis. After being divided, the mean constraint of the 58 melodic stems in the high constraint group was 80.6% ($SD = 10.4$), while that of the 57 stems in the low constraint group was 36.9% ($SD = 8.8$). In the highly constraining melodies, expected target notes (the tonic) had a mean cloze probability of 78.8% ($SD = 16.4$), while less expected target notes (the sixth) had a mean cloze probability of 0.9% ($SD = 1.6$). In low constraint melodies, the mean cloze probability of the less expected sixth scale degree was 7.4% ($SD = 7.5$); the mean cloze probability of the sixth differed significantly between high and low constraint melodies, $t(113) = 6.42$, $p < .0001$. While it would have been ideal for the high and low constraint melodies to be matched in the cloze probability of the unexpected target note, this is difficult to achieve in practice because there are only 7 in-key scale degrees. Thus in non-constraining melodies, some probability mass is likely to fall on each of these degrees (indeed, if all 7 were equally likely, each would have a close probability of ~14.3%). As noted later, to ascertain if ERP differences between the 6th scale degree after high vs. low-constrain contexts were due to differences in the cloze probability of this note, a regression-based approach was used in data analysis.

For each subject, averaged ERPs were calculated to target notes in each of the three conditions from trials free of artifact after subtraction of the 250 ms prestimulus baseline. Statistical analyses were performed on the averaged ERPs for each subject in each condition. Based on our prediction that ERPs in this study would resemble the frontal positivities found in language studies of comparable design, repeated measures ANOVAs were conducted within a priori spatial and temporal regions of interest. In addition, we used a mass univariate approach, in which tests are performed at multiple time points and electrode sites and a permutation-based cluster mass test is used to account for multiple comparisons; for further information on this approach, see Fields and Kuperberg (2019). In order to detect any novel ERP effects, an omnibus analysis as well as pairwise comparisons between conditions were carried out across a wide time window (100-900 ms) and spatial region (all electrodes excluding temporal sites) with factors of Condition and Block, using the Mass Univariate ERP Toolbox (Groppe, Urbach, & Kutas, 2011) and Factorial Mass Univariate ERP Toolbox (Fields, 2017).

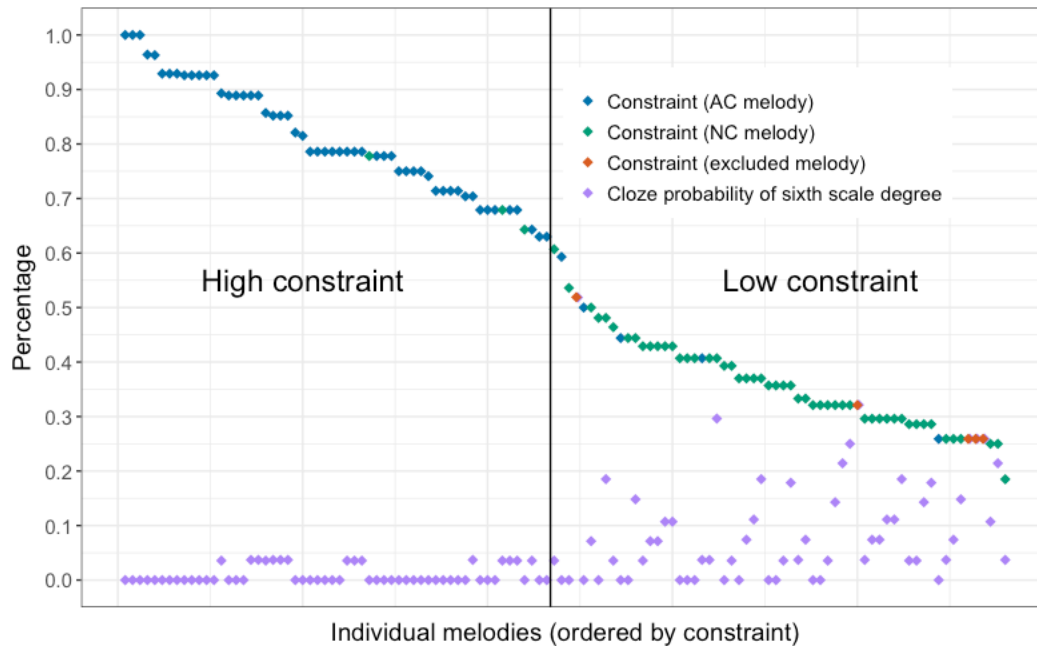


Figure 3.5. For purposes of ERP analysis, melodies were divided into higher- and lower-constraint groups according to their empirical constraint based on the results of the melodic cloze probability paradigm. Melodies originally designed as AC melodic stems are plotted in blue, while melodies originally designed as NC melodic stems are plotted in green; the cloze probability of the sixth scale degree (the non-expected target note) for each of these melodies is plotted below in purple. Five melodies plotted in red were excluded because the sixth scale degree was the most expected note and therefore would not have been non-expected. The mean constraint of the melodic stems used in the high constraint group ($n = 58$) was 80.6%, while that of the stems used in the low constraint group ($n = 57$) was 36.9%.

Results

Behavioral

Participants' performance on the memory recognition task was generally accurate. Behavioral responses were not available for seven participants; the remaining 29 participants correctly indicated whether or not an excerpt had appeared in a melody an average of 84.9% of the time ($SD = 5.3\%$), and all participants' accuracies were over 70%.

ERP

Grand average waveforms for each of the three critical conditions are shown in **Figure 3.6**. The P1, N1 and P2 components elicited by stimulus onset are visible in all conditions; these repeat when the next note of the melody begins at 500 ms after the critical note. The omnibus mass univariate analysis revealed clusters showing a significant main effect of Condition between 314-768 ms ($p < .001$) and 736-900 ms ($p = .02$); there was no significant interaction between Condition and Block in this or any follow-up analysis. This was followed up with separate pairwise analyses of the effects of Constraint (unexpected target notes in constraining melodies vs. unexpected target notes in non-constraining melodies) and Cloze Probability (expected target notes vs. unexpected target notes in constraining melodies).

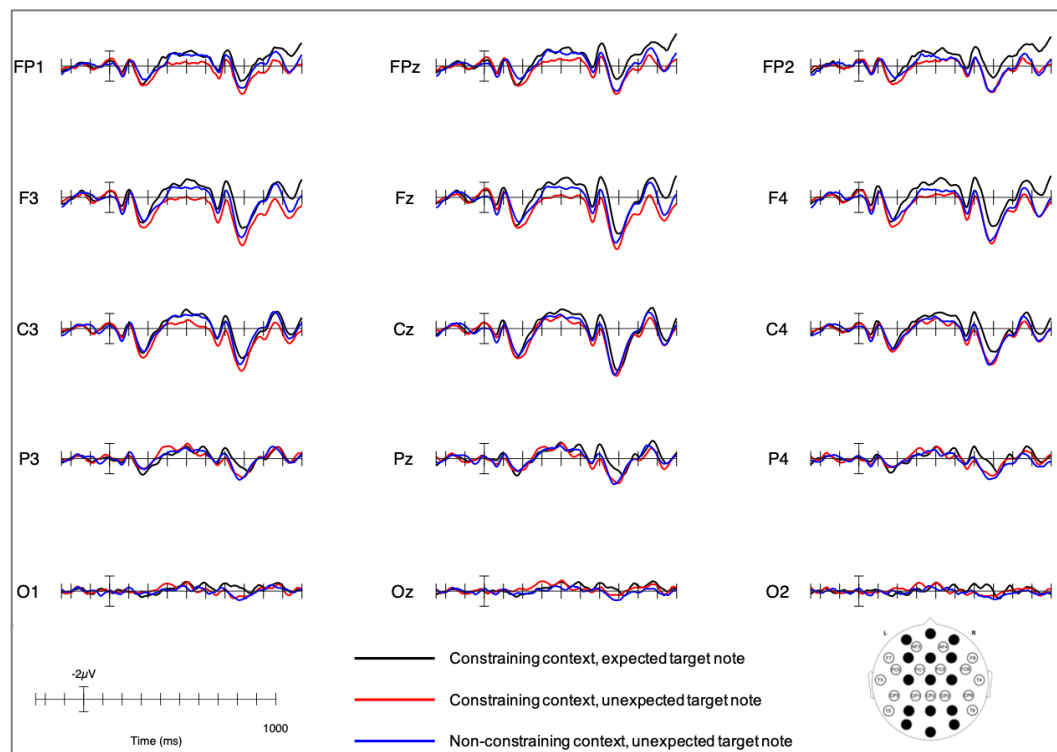


Figure 3.6. Grand average ERPs for all conditions over frontal electrode sites. All ERP waveforms in figures have been low-pass filtered at 20 Hz for purposes of visualization.

Constraint. In line with prior studies in the language literature (e.g., Federmeier et al., 2007; Fogel, Wlotko, Kuperberg, & Patel, 2017), 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 spatiotemporal window, unexpected target notes in constraining melodies elicited a marginally significant left-lateralized frontal positivity relative to the same unexpected target notes in non-constraining melodies, $F(1,35) = 3.13$, $p = 0.085$ (**Figure 3.7**). Within the left side of this region (FP1, FPz, AF3, F7, F3, Fz), the contrast was significant, $F(1,35) = 7.06$, $p = 0.012$. Furthermore in earlier time windows, the cluster analysis for Constraint revealed a significant cluster that showed a main effect between 307-596 ms (see **Table 3.1** and **Figure 3.9**).

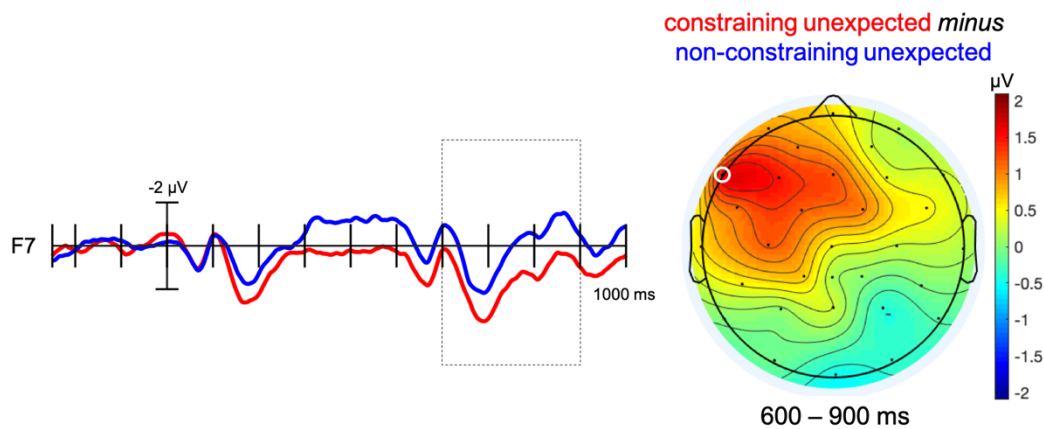


Figure 3.7. Grand average ERPs at electrode F7 to unexpected target notes in constraining melodies and unexpected target notes in non-constraining melodies; topographic distribution of the difference between their mean amplitudes between 600 and 900 ms (time window indicated in the box on the waveform).

Cloze probability. Previous studies have reported ERP effects of violating musical expectations at anterior sites within in the range of 150-300 ms (e.g., Koelsch et al., 2000, 2007). To evaluate this effect, ERPs were measured in this time window at ten anterior electrode sites (FP1, FPz, AF3, F7, F3, Fz, F4, F8, AF4, FP2). Here, there was no evidence of the early right anterior negativity (ERAN) that has previously been associated with musical expectancy violations; instead, unexpected target notes elicited an bilateral anterior *positivity* relative to expected target notes within constraining melodies in the 150-300 ms window, $F(1,35) = 6.02, p = .019$ (**Figure 3.8**).

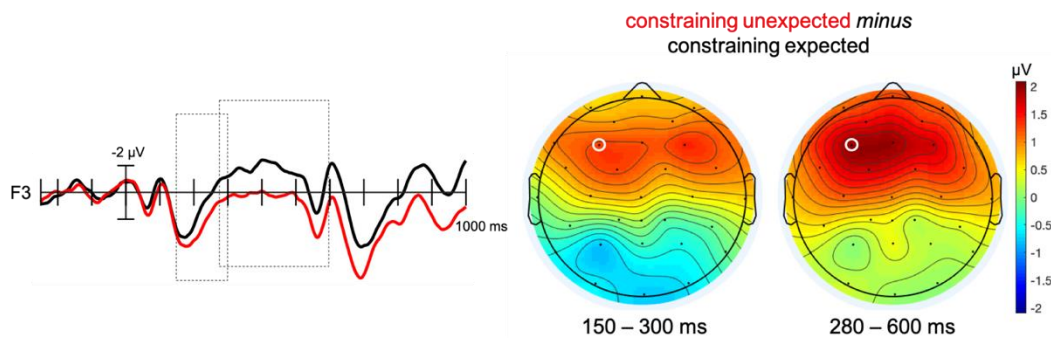


Figure 3.8. Grand average ERPs at electrode F3 to expected and unexpected target notes in constraining melodies; topographic distribution of the difference between their mean amplitudes between 150-300 and 280-600 ms.

Using melodic stimuli, Pearce et al. (2010) reported a more positive amplitude for low-probability than for high-probability notes in frontal regions from 280 to 600 ms. In anterior electrodes, we observed a similar effect in the same 280-600 ms time window: in constraining contexts, the amplitude elicited by unexpected notes was more positive than that elicited by expected notes, $F(1,35) = 13.78, p < .001$ (**Figure 3.8**). Additionally, the cluster analysis for

Cloze Probability revealed two significant clusters with a main effect of Condition between 150-721 and 736-900 ms; see **Table 3.1** and **Figure 3.9**.

Table 3.1. Significant clusters in mass univariate analyses of pairwise comparisons.

Comparison	Cluster <i>p</i> -value	Spatial extent	Temporal extent	Spatial cluster mass peak	Temporal cluster mass peak
HC unexpected vs. LC unexpected (Effect of constraint)	0.045	FP1, FPz, AF3, F7, F3, FC1, FC5, C3, CP1, FC2, Fz	307-596 ms	F7	455 ms
	0.079	AF3, F7, F3, FC1, FC5, C3, CP1, CP5, FC2, Fz	674-900 ms	F7	846 ms
HC expected vs. HC unexpected (Effect of cloze probability)	0.0001	FP1, FPz, AF3, F7, F3, FC1, FC5, C3, CP1, CP5, T5, P3, Pz, O1, Oz, O2, P4, T6, CP6, CP2, CPz, Cz, C4, FC6, FC2, Fz, F4, F8, AF4, FP2	150-721 ms	F3	627 ms
	0.011	FP1, FPz, AF3, F7, F3, FC1, FC5, C3, CP1, CP5, CPz, Cz, C4, FC6, FC2, Fz, F4, F8, AF4, FP2	736-900 ms	F3	893 ms
HC expected vs. LC unexpected	0.0045	FP1, FPz, AF3, F3, FC1, FC5, C3, CP1, CP5, T5, P3, Pz, O1, Oz, O2, P4, T6, CP6, CP2, CPz, Cz, C4, FC6, FC2, Fz, F4, F8, AF4, FP2	572-830 ms	FP2	666 ms

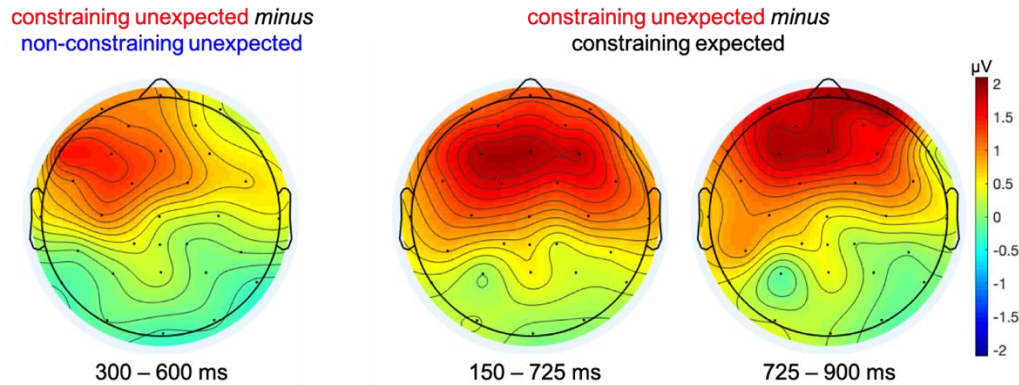


Figure 3.9. Topographic distributions of significant clusters revealed by mass univariate analyses of the main effects of Constraint and Cloze Probability.

Regression

As noted in the *Analysis* section above, the mean cloze probability of unexpected target notes differed significantly between high and low constraint melodies. This confounds the comparison between unexpected target notes in constraining melodies and unexpected target notes in non-constraining melodies, as they differed in both contextual constraint *and* cloze probability. This problem is somewhat inevitable when working with melodic stimuli, as there are a limited number of “plausible” continuations in music. As constraint goes down and fewer participants respond with the tonic in the cloze probability task, more responses will be distributed to each of the other plausible notes, and their probabilities will rise (see **Figure 3.2** for an illustration).

To fully explore the relationship between target note cloze probability, contextual constraint, and amplitude, we ran a mixed-effects regression using data from the HC-sixth and LC-sixth conditions (i.e. the two conditions in which the target note was *not* the most likely continuation). The dependent variables were z-transformed target note cloze probability and contextual constraint. The

independent variable was amplitude averaged across electrodes FP1, FPz, AF3, F7, F3, FC1, FC5, C3, CP1, FC2, and Fz from 300-600 ms (based on the region where a significant categorical effect was found in the mass univariate analysis of this contrast). Following Barr, Levy, Scheepers, & Tily (2013) in this and all other regression analyses we included the maximal random effects structure justified by the design, here a by-subjects intercept, by-subjects random effects of cloze and constraint, a by-items intercept, and a by-items random effect of constraint (with an “item” defined as a matched pair of melodic contexts). Regression analyses were fit using the lme4 package (Bates, Mächler, Bolker, & Walker, 2015) or in Julia (Bezanson, Edelman, Karpinski, & Shah, 2017) using the JuliaCall package (Li, 2018) in R (R Core Team, 2018).

Table 3.2. amplitude ~ constraint + cloze

	Estimate	Std.Error	t value	P(> t)
(Intercept)	-0.493	0.209	-2.352	0.023
constraint	0.537	0.198	2.705	0.009
cloze	-0.160	0.199	-0.802	0.427

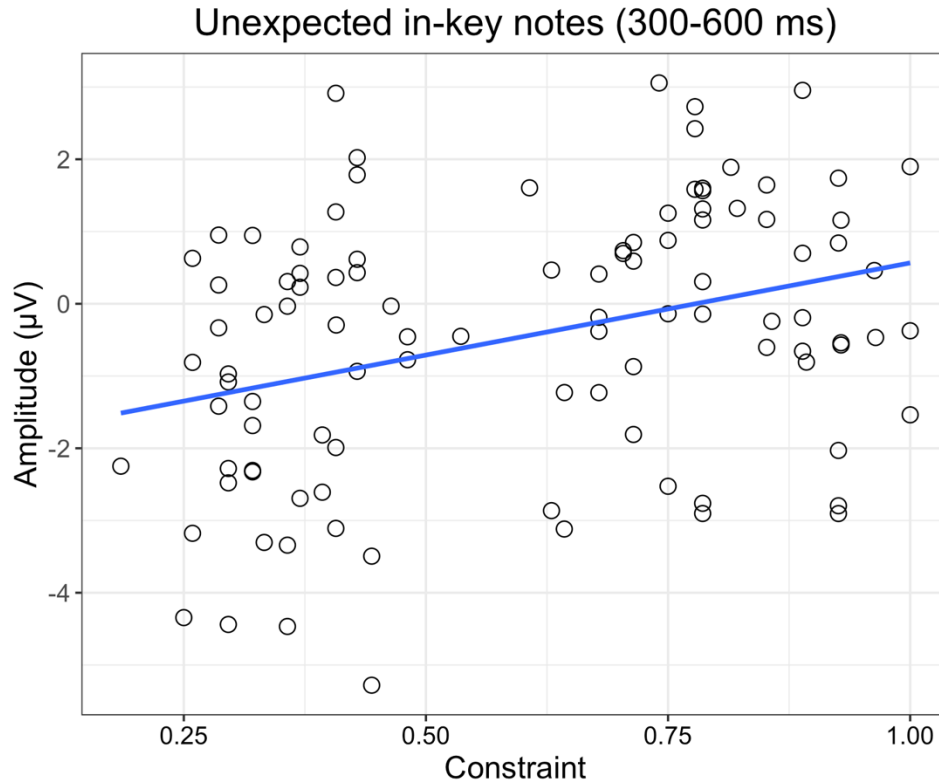


Figure 3.10. Amplitude across frontal electrodes from 300-600 ms vs. contextual constraint for items in the HC-sixth and LC-sixth conditions. Constraint has a significant effect on amplitude of low cloze items, such that the ERPs elicited by in-key unexpected notes are more positive in higher constraint contexts.

As seen in **Table 3.2**, constraint has a significant effect on amplitude of low cloze items, such that the ERPs elicited by in-key unexpected notes are more positive in higher constraint contexts (**Figure 3.10**). Cloze probability does not significantly modulate amplitude.

Is the effect of constraint seen here truly a continuous effect of constraint on amplitude, or could it be driven by the categorical difference between high constraint (HC) and low constraint (LC) contexts? (Note that these designations are based on empirically measured constraint of stems as shown in **Figure 3.5**, rather than their AC vs. NC status.) To test this, we ran two further analyses. First,

we ran a regression using cloze probability and categorical condition (treatment coded with HC as reference level) as predictors. As shown in **Table 3.3**, we find a significant effect of condition on amplitude, with LC contexts eliciting more negative amplitudes than HC contexts. Next, to directly compare continuous constraint and categorical condition as predictors of amplitude, we ran a regression including both constraint and condition, as well as cloze, as predictors. As shown in **Table 3.4**, constraint is marginally significant and condition does not reach significance in this analysis. From this we conclude that while constraint is a better predictor of amplitude than condition, there is not enough unique variance predicted by either one to determine which is truly driving the effect.³

We note that in none of our models does cloze probability of the target note approach significance, providing further evidence that the frontal positivity seen here specifically reflects the effect of violating of a prediction generated by the context, rather than the surprise of the target note per se (though we do note that because cloze probabilities of the target note in these conditions are low across the board, there would be little power to detect any effect of target note cloze).

Table 3.3. amplitude ~ condition + cloze

	Estimate	Std.Error	z value	P(> z)
(Intercept)	-0.052	0.273	-0.190	0.850
condition: LC	-0.861	0.403	-2.134	0.033
cloze	-0.164	0.197	-0.832	0.405

³ We also conducted versions of all analyses including the interaction of cloze with constraint and/or condition. In no case did the interaction reach significance; thus, we report here models fit without the interaction term.

Table 3.4. amplitude ~ condition + constraint + cloze

	Estimate	Std.Error	z value	P(> z)
(Intercept)	-0.725	0.436	-1.663	0.096
Constraint	0.749	0.411	1.822	0.069
condition: LC	0.486	0.791	0.615	0.539
Cloze	-0.167	0.200	-0.834	0.404

Discussion

Here, we examined neural correlates of in-key expectation violations in melodic sequences. Unexpected in-key notes in constraining contexts elicited a sustained frontal positivity when compared to unexpected notes in non-constraining contexts. This effect is statistically significant from around 300 – 900 ms following the onset of the target note, including in the time window where the frontal positivity seen in comparable studies of linguistic expectancy violations is also significant (600-900 ms). This supports the predictions of the resource-sharing framework for music and language processing: based on the observation of a frontal positivity in both domains, mispredicted notes in melodic sequences may be processed similarly to mispredicted words in sentences. This positivity may reflect how the brain handles prediction error, e.g., by suppressing activation of the mispredicted element or updating the evolving representation of the unfolding sequence. Thus while there are surely many differences in the precise neural processes at work when perceiving and predicting information in music and language, our findings suggest that there may be some similarities in the processing of violated predictions for specific upcoming items in both domains, as indexed by the frontal positivity. It is worth noting that this positivity becomes

significant around 300 ms in melodic sequences, earlier than typically seen in comparable language studies. This may reflect the fact that linguistic studies often use visual (rather than auditory) presentation, in addition to linguistic expectancy violations possibly taking more time to process due to the need to integrate semantic with structural information in lexical processing.

Turning to the ERP comparison between expected vs. unexpected notes in highly-constraining contexts, an unexpected and striking finding of this study is the lack of an early anterior negativity, the effect that has been elicited by unexpected compared to expected notes by the majority of previous studies of unexpected notes in melodies. Instead of a negativity, we found that unexpected notes in constraining contexts elicited a positivity compared to expected notes beginning around the same time window (as they also did in comparison to unexpected notes in non-constraining contexts). In contrast to most previous studies, which have generally used unexpected notes that were incongruous in some way (e.g., out-of-key), the unexpected in-key notes in the present experiment were in-key and congruous in their melodic contexts.

However, our finding of a frontal positivity elicited by unexpected compared to expected notes does appear to be similar to the results reported by Pearce et al. (2010). They also used stimuli that contained only in-key notes varying in probability and reported that low probability notes elicited a right-lateralized frontal positivity from 280-600 ms, a time window where we also observed a significant bilateral frontal positivity. Though Pearce et al. (2010) used stimuli with rhythmic structure removed and with a very slow note presentation

rate (1 note per second), their melodies were sourced from hymns thus otherwise resembled natural musical stimuli, and target notes were not task-relevant.

Taken together with the results of Pearce et al. (2010), our ERP findings for unexpected vs. expected notes in constraining melodies indicate that neural responses to unexpected notes in constraining melodies can be strikingly different from the early anterior negativity that has typically been reported in such cases. It is notable that we did not observe any effects that appear to be a function of the cloze probability of a note; instead, we only saw a frontal positivity elicited by notes that violate specific predictions. Our use of in-key target notes may be largely responsible for this difference. To explore this possibility and to make our study more comparable to prior work on melodic expectation, we next conducted a study using the same design as the current study but with out-of-key target notes, as described in the next chapter.

Chapter 4. Neural correlates of processing out-of-key melodic prediction violations

Introduction

The ERP study described in the previous chapter (Experiment 1) differed from the majority of previous ERP studies of melodic processing. Prior experiments have not controlled the predictive constraint of melodic contexts or not quantified the degree of expectancy violation represented by unexpected notes. Additionally, in prior studies unexpected notes have often been task-relevant (e.g., in incongruity detection tasks), which could influence the neural response to such notes. Even when the target note is not-task relevant, participants have sometimes been asked to detect other melodic tones that differ in timbre (e.g., Kalda & Minati, 2012; Koelsch & Jentschke, 2010; Miranda & Ullman, 2007). However, it has recently been demonstrated that including multiple timbres in an experiment can interfere with the processing of musical syntax in melodies, even when participants have been instructed to attend specifically to syntactic violations (Fiveash, McArthur, & Thompson, 2018; Fiveash, Thompson, et al., 2018). Given the numerous differences between our paradigm and other experimental paradigms, it is difficult to compare the effects of the in-key prediction violations used in Experiment 1 to the effects of the out-of-key violations used in previous studies; differences in results could be due to any number of aspects of the experiment beyond the nature of the melodic expectation violations themselves.

The primary finding of most ERP studies of melodic expectancy violations has been an anterior negativity elicited by unexpected target notes compared to expected notes, generally referred to as the Early Right Anterior Negativity or ERAN (Brattico et al., 2006; Koelsch & Jentschke, 2010; Lindsen, Pearce, Doyne, Wiggins, & Bhattacharya, 2010; Miranda & Ullman, 2007; Pearce et al., 2010; Schön & Besson, 2005). The ERAN is thought to reflect violations of syntactic structure in music (Kalda & Minati, 2012; Koelsch, 2009; Koelsch, Rohrmeier, Torrecuso, & Jentschke, 2013; Rohrmeier & Koelsch, 2012). The latency and location of this component are both reported to vary to some degree, possibly depending on a number of different factors including the temporal predictability of stimuli, but it is generally maximal in the right hemisphere between 150 and 250 ms after stimulus presentation (Koelsch, 2009). A subset of the paradigm used by Miranda and Ullman (2007) is perhaps the most relevant comparison to the present study: here, out-of-key target notes in the middle of novel tonal melodies elicited an anterior central negativity (interpreted as an ERAN) from 150-270 ms.

Later ERP effects of violating melodic expectations have been less consistently reported. For example, Miranda and Ullman (2007) found that out-of-key notes elicited a posterior positivity 500-700 ms after stimulus presentation (P600), and Besson & Faïta (1995) and Halpern et al. (2017) report similar results. However, other studies of unexpected notes in melodies have not reported effects of violating expectations on the P600 (e.g., Kalda & Minati, 2012; Koelsch & Jentschke, 2010), and still more have found that elicitation of the P600 depends

on the task performed by participants (Brattico et al., 2006; Lagrois et al., 2018; Zendel et al., 2015).

Hsu, Le Bars, Hämäläinen, and Waszak (2015) used sequences of five ascending tones to dissociate mispredicted (i.e., incorrectly predicted) stimuli from unpredicted (i.e. no prediction took place) stimuli (see Figure 1.8). While these sequences cannot truly be described as “melodies,” listeners still have to draw on their long term tonal knowledge to generate expectations for their continuations (see Ross & Hansen, 2016 for further discussion). The experimental manipulation was similar to the present study in that based on sequence context, a given note in a sequence was either predicted, mispredicted, or unpredicted. While target notes in Hsu et al. 2015 were sequence-final and within the C major scale, in contrast to our sequence-medial and out-of-key target notes, both mispredicted notes are comparable in that they violate a schema. Hsu et al. (2015) reported differences in the ERPs elicited by each category of target tones in the N1/P2 range. Specifically, mispredicted tones were more negative than predicted tones while unpredicted tones were less negative than predicted tones on the N1 component around 100-250 ms, demonstrating distinct neural responses elicited by these different types of unexpected stimuli.

In order to determine to what extent the differences between the results of Experiment 1 and prior ERP studies of melodic expectation were due to our use of in-key target notes, we conducted a second experiment (Experiment 2) that was identical to Experiment 1, except that unexpected notes were now out-of-key. As in Experiment 1, participants were asked to perform a melody recognition task,

which requires attentive listening but does not call to attention any particular aspect of the melodies and does not require the use of timbre deviants in any of the melodies. We predict that ERP responses in this experiment with out-of-key unexpected notes will be more comparable to those reported by other melody studies using out-of-key targets (e.g., Halpern et al., 2017; Miranda & Ullman, 2007): in constraining contexts, unexpected out-of-key notes will elicit an early anterior negativity and/or P600 compared to expected notes. Additionally, we might observe similar results to Hsu et al. (2015), such that out-of-key notes in constraining contexts will be more negative than expected notes in constraining contexts, which will themselves be more negative than out-of-key notes in non-constraining contexts between 100-250 ms.

A secondary prediction is derived from the language literature, where it has been demonstrated that the late posterior positivity (P600) elicited by syntactically or semantically anomalous words is larger when critical words appear in high constraint than in low constraint contexts (Gunter, Friederici, & Schriefers, 2000; Kuperberg et al., 2018) (**Figure 4.1**). Therefore, if out-of-key notes do elicit a P600 relative to expected notes, we predict that the P600 elicited by out-of-key notes will be larger when they are in constraining melodies than when they appear in non-constraining melodies.

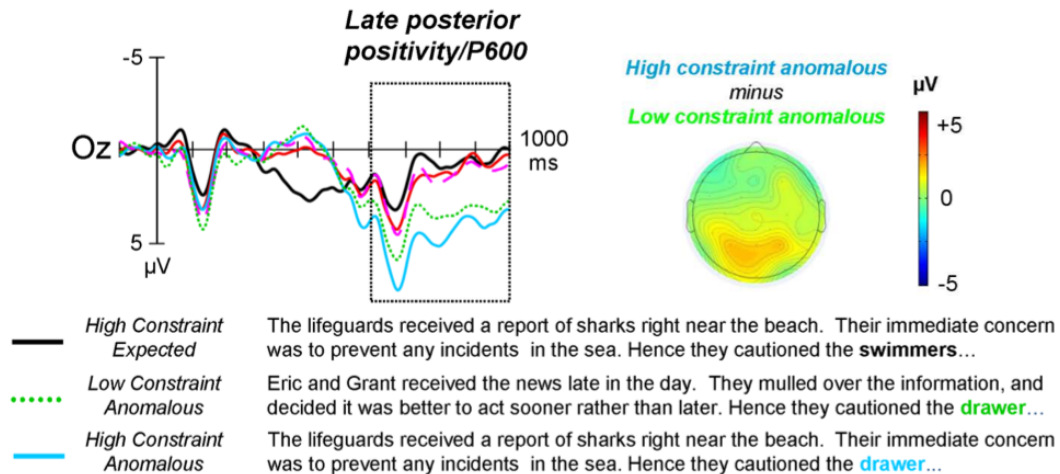


Figure 4.1. Results from Kuperberg et al. (2018): grand-averaged ERP waveforms at electrode Oz showing the late posterior positivity, which is larger to anomalous words in high constraint contexts than to anomalous words in low constraint contexts, and voltage maps showing the scalp topographies of the effect averaged across the 600-1000ms time window.

Methods

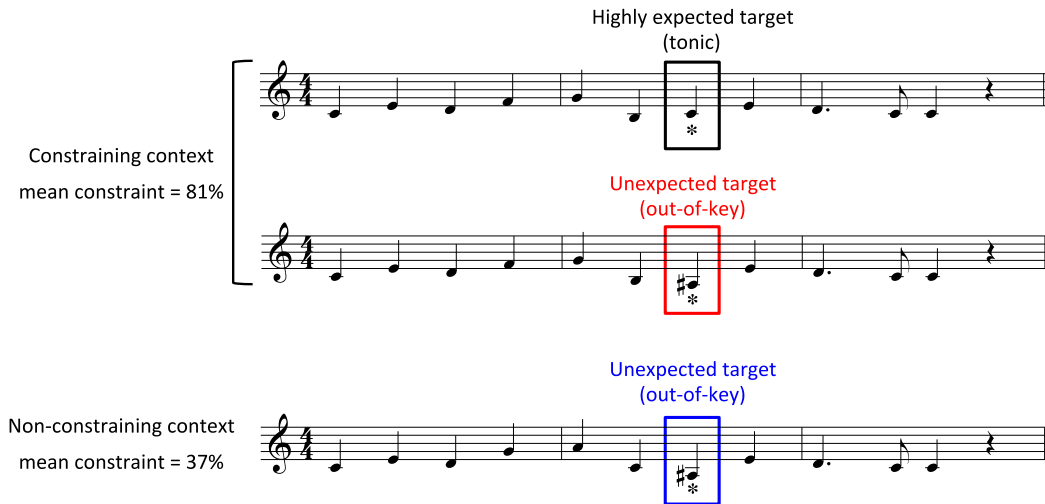
ERP Stimuli

The stimuli used in the present experiment consisted of the same sixty pairs of novel 11-16 note melodies ($M = 14.0$ notes) as Experiment 1 (see Chapter 3 for full details of stimuli development and characteristics). As in Experiment 1, each melodic pair had one version that constrained participants' expectations for the tonic and one that did not create a strong expectation for any particular note. Constraining contexts again contained either a highly expected mid-sequence target note or an unexpected target note; however, in Experiment 2, unexpected target notes in constraining and non-constraining contexts were out-of-key. For each pair, the same out-of-key target note was used in both contexts (see **Figure 4.2** for example stimuli). These out-of-key notes were chosen according to the following rules:

1. Use the note one semitone above or below the sixth diatonic scale degree, whichever is closer to the pre-target note in the constraining condition, unless the distance from this notes and the pre-target note in either condition would be a tritone (an augmented fourth or diminished fifth), in which case...
2. Use the note one semitone in the opposite direction or three semitones below the sixth, unless that would disrupt the melodic contour of the melody, in which case...
3. Use the tone five semitones above the sixth.

In following these rules, for most melodies the interval between the pre-target and target note was altered by just one semitone; see **Table 4.1** for characteristics of differences between experiments.

In each of the two blocks of the experiment, an individual participant heard a given melody pair in only one of these three conditions (i.e., 20 melodies per condition per subject), as well as 30 filler melodies of a comparable length and style that did not contain a mid-sequence cadence, for a total of 90 melodies in each section. No experimental melody was repeated in the same condition between blocks.



Kuperberg, & Patel, 2017; an additional 6 participants were excluded due to excessive artifact). Participants had at least five years of active musical experience within the last ten years ($M = 8.7$ years, $SD = 2.0$) and an average of 8.9 years of lifetime formal musical training ($SD = 3.6$, ranging from 0-15 years). All participants were between the ages of 18 and 35 ($M = 19.9$, $SD = 1.7$), right-handed, had normal hearing, and reported no history of neurological disorders or head trauma. They provided informed consent in accordance with the procedures of the Institutional Review Board of Tufts University and were compensated for their time.

Procedure

Participants were seated in a comfortable chair in a dimly lit room and provided with headphones. Trials were self-paced; each melody played only once a participant pressed a button to advance. Forty of the 90 melodies (44%) were followed by memory tests, where participants heard a 1-4 second excerpt from a melody and indicated on a keyboard if it had appeared in the melody they just heard.

ERP data 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. 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. 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 v14.1.2 and ERPLAB v7.0.0.0. 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. It was then segmented into epochs from 300 ms before to 1100 ms after the onset of each target note, using the interval from -250 to 0 ms for baseline correction. Algorithms in ERPLAB were used to detect trials containing blinks, eye movements, and other artifacts within the time window used for analysis (-250 to 1000 ms); after artifact rejection, an average of 12.2% of critical trials were excluded per participant ($SD = 8.8$).

Analysis

For purposes of group analysis, melodies were again divided into higher- and lower-constraint groups according to their empirical constraint based on the results of the melodic cloze probability paradigm as described in Chapter 3 (see Figure 3.5). After being divided, the mean constraint of the 58 melodic stems in the high constraint group was 80.6% ($SD = 10.4$), while that of the 57 stems in the low constraint group was 36.9% ($SD = 8.8$). In highly constraining melodies, expected target notes (the tonic) had a mean cloze probability of 78.8% ($SD = 16.4$), while unexpected out-of-key target notes had a mean cloze probability of 1.4% ($SD = 8.9$). In low constraint melodies, the mean cloze probability of unexpected out-of-key target notes was 3.0% ($SD = 7.0$). In this experiment, the

cloze probability of unexpected target notes did not significantly differ between constraint conditions, $t(113) = 1.07$, $p = 0.29$. For each subject, averaged ERPs were calculated for each condition from trials free of artifact after subtraction of the 250 ms prestimulus baseline.

Statistical analyses were performed on the averaged ERPs for each subject in each condition; repeated measures ANOVAs were conducted within a priori spatial and temporal regions of interest. In addition, we used a mass univariate approach, in which tests are performed at multiple time points and electrode sites and a permutation-based cluster mass test is used to account for multiple comparisons; for further information on this approach, see (Fields & Kuperberg, 2019). An omnibus analysis as well as pairwise comparisons between conditions were carried out across a wide time window (100-900 ms) and spatial region (all electrodes excluding temporal sites) with factors of Condition and Block, using the Mass Univariate ERP Toolbox (Groppe et al., 2011) and Factorial Mass Univariate ERP Toolbox (Fields, 2017).

Results

Behavioral

Participants' performance on the memory recognition task was generally accurate. Participants correctly indicated whether or not an excerpt had appeared in a melody an average of 85.6% of the time ($SD = 4.6\%$); all participants' accuracies were over 75%.

ERP

Grand average waveforms for each of the three critical conditions are shown in **Figure 4.3**. The P1, N1 and P2 components elicited by stimulus onset are visible in all conditions; these repeat when the next note of the melody begins at 500 ms after the critical note. The omnibus mass univariate analysis revealed a cluster showing a significant main effect of Condition between 424-774 ms ($p = .019$); there was no significant interaction between Condition and Block in this or any follow-up analysis. This was followed up with separate pairwise analyses of the effects of Cloze Probability (expected target notes vs. out-of-key target notes in constraining melodies) and Constraint (out-of-key target notes in constraining melodies vs. out-of-key target notes in non-constraining melodies).

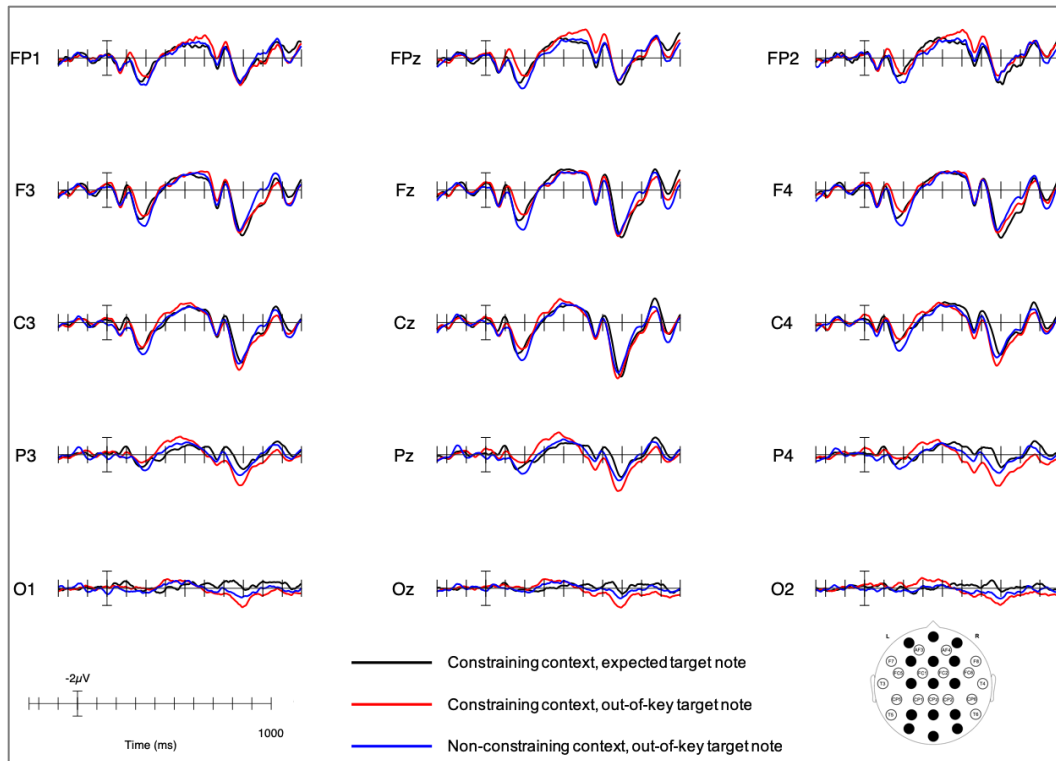


Figure 4.3. Grand average ERPs for all conditions at selected electrode sites. All ERP waveforms in figures have been low-pass filtered at 20 Hz for purposes of visualization.

Cloze probability. The effect of varying cloze probability was evaluated by comparing the ERPs elicited by expected vs. unexpected notes within constraining contexts. In order to assess the early anterior negativity reported by previous studies (e.g., Miranda & Ullman, 2007), ERPs were measured from 150-270 ms at ten anterior electrode sites (FP1, FPz, AF3, F7, F3, Fz, F4, F8, AF4, and FP2). Here, there was no evidence of the early right anterior negativity (ERAN) that has previously been associated with musical expectancy violations, $F(1,26) = 0.27$, $p = 0.61$. Across these electrodes, we also did not observe the early anterior difference that was reported by Hsu et al. (2015) between predicted

and mispredicted notes in the range of 100-250 ms, $F(1,26) = 1.44$, $p = 0.24$

(**Figure 4.4**).

Again following Miranda and Ullman (2007), ERPs were measured from 500-700 ms at posterior electrode sites (CP1, CP5, P3, Pz, O1, Oz, O2, P4, CP6, CP2, and CPz) to evaluate the presence of a late posterior positivity (P600). In this region, the amplitude of ERPs elicited by unexpected notes was more positive than the ERPs elicited by expected notes, $F(1,26) = 12.95$, $p = 0.001$ (**Figure 4.5**). The cluster analysis for Cloze Probability revealed a significant cluster corresponding to the P600 location and latency (**Table 4.2**).

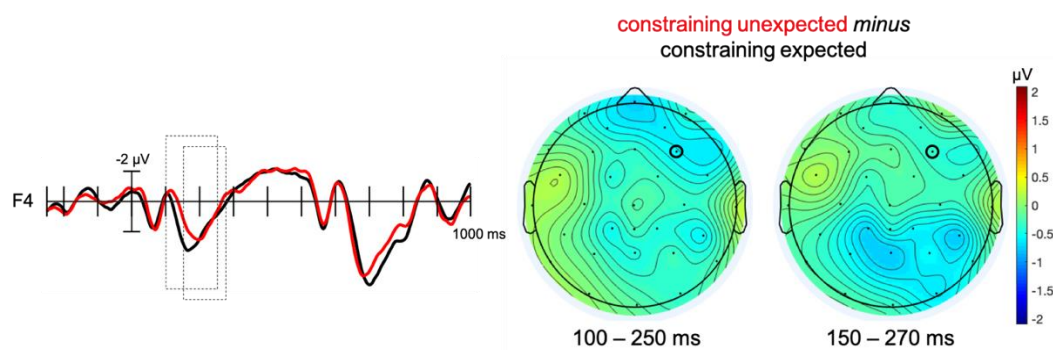


Figure 4.4. Grand average ERPs at electrode F4 to expected and out-of-key target notes in constraining melodies; topographic distribution of the difference between their mean amplitudes between 100-250 and 150-270 ms.

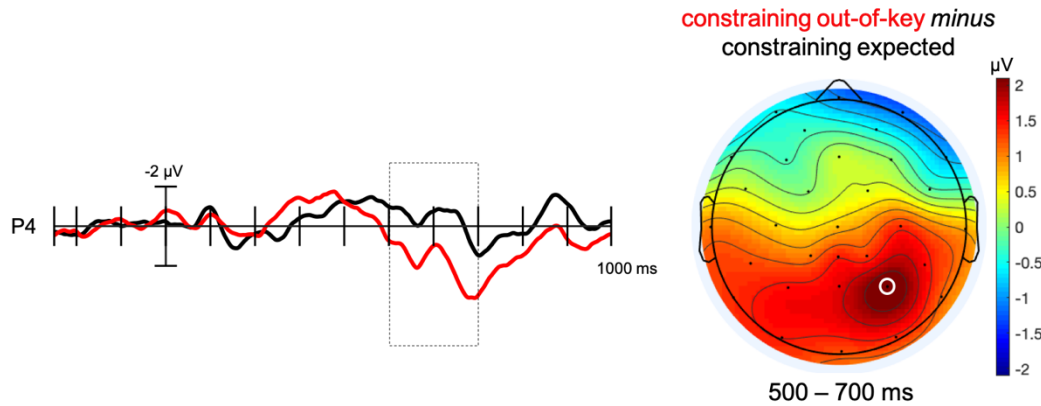


Figure 4.5. Grand average ERPs at electrode P4 to expected and out-of-key target notes in constraining melodies; topographic distribution of the difference between their mean amplitudes between 500 and 700 ms.

Constraint. The effect of predictive constraint was measured by comparing the ERPs elicited by out-of-key target notes when they were presented in constraining vs. non-constraining contexts. Previously, Hsu et al. (2015) found an early anterior difference between mispredicted and unpredicted stimuli in the N1/P2 time window, where mispredicted tones were more negative. To evaluate this effect, we measured the mean amplitude between 100 and 250 ms at ten anterior electrode sites (FP1, FPz, AF3, F7, F3, Fz, F4, F8, AF4, FP2). In this time window, out-of-key target notes in constraining melodies were more negative relative to the same out-of-key target notes in non-constraining melodies, $F(1,26) = 14.08$, $p < .001$ (**Figure 4.6**). The cluster analysis for Constraint revealed a significant cluster in this same time and region (**Table 4.2**).

ERPs were additionally measured from 500-700 ms at posterior electrode sites (CP1, CP5, P3, Pz, O1, Oz, O2, P4, CP6, CP2, and CPz) to evaluate the presence of a late posterior positivity (P600). In this region, the amplitude of ERPs elicited by unexpected notes in constraining contexts was more positive

than the ERPs elicited by unexpected notes in non-constraining contexts, $F(1,26) = 5.48, p = 0.027$ (**Figure 4.7**).

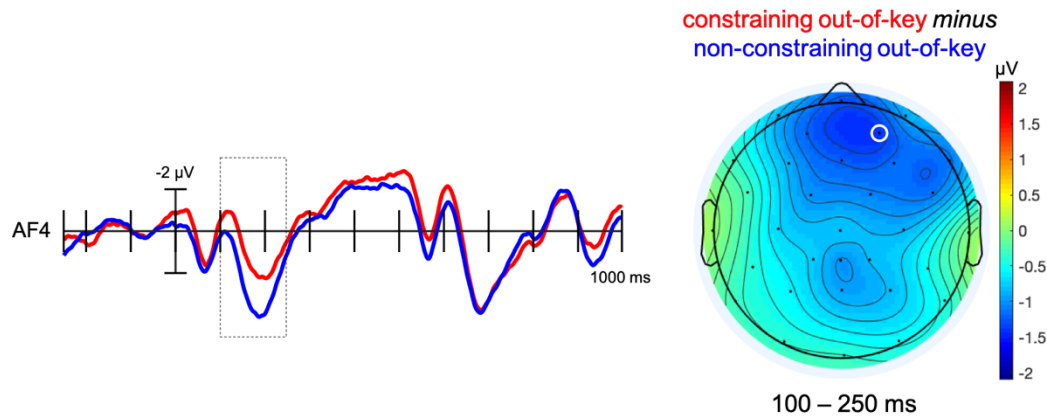


Figure 4.6. Grand average ERPs at electrode AF4 to out-of-key target notes in constraining and non-constraining melodies; topographic distribution of the difference between their mean amplitudes between 100 and 250 ms.

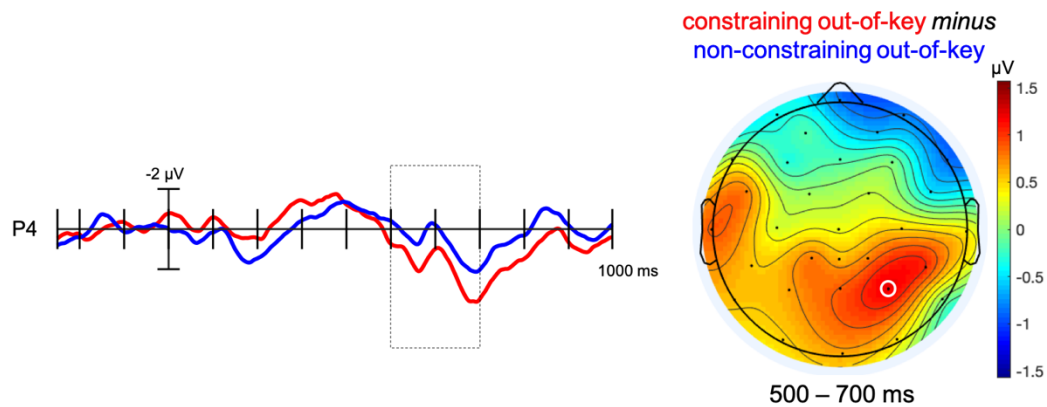


Figure 4.7. Grand average ERPs at electrode P4 to out-of-key target notes in constraining and non-constraining melodies; topographic distribution of the difference between their mean amplitudes between 500 and 700 ms.

Table 4.2. Significant clusters in mass univariate analyses of pairwise comparisons.

Effect	Cluster <i>p</i> -value	Spatial extent	Temporal extent	Spatial cluster mass peak	Temporal cluster mass peak
Cloze probability (constraining out-of-key vs. constraining expected)	0.015	FC1, FC5, C3, CP1, CP5, T5, P3, Pz, O1, Oz, O2, P4, T6, CP6, CP2, CPz, Cz, C4, FC2, Fz, F4	479-775 ms	P4	658 ms
Constraint (constraining out-of-key vs. non-constraining out-of-key)	0.018	FP1, FPz, AF3, F3, FC1, C3, CP1, CP5, T5, P3, Pz, P4, CP6, CP2, CPz, Cz, C4, FC6, FC2, Fz, F4, F8, AF4, FP2	104-244 ms	AF4	197 ms

Regression

We ran a mixed-effects regression using data from the high constraint (HC) out-of-key and low constraint (LC) out-of-key conditions, following the analysis in Chapter 3. The independent variable was amplitude averaged across the spatiotemporal region where a significant categorical effect was found in the mass univariate analysis of this contrast: electrodes FP1, FPz, AF3, F3, FC1, C3, CP1, CP5, T5, P3, Pz, P4, CP6, CP2, CPz, Cz, C4, FC6, FC2, Fz, F4, F8, AF4, and FP2, between 100-250 ms. All other parameters of this analysis remained the same as those used in Chapter 3. Regression analyses were fit using the `lme4` package (Bates et al., 2015) or in Julia (Bezanson et al., 2017) using the `JuliaCall` package (Li, 2018) in R (R Core Team, 2018).

Table 4.3. amplitude ~ constraint + cloze

	Estimate	Std.Error	t value	P(> t)
(Intercept)	1.311	0.317	4.129	<0.001
constraint	-0.420	0.167	-2.511	0.015
cloze	0.047	0.178	0.266	0.791

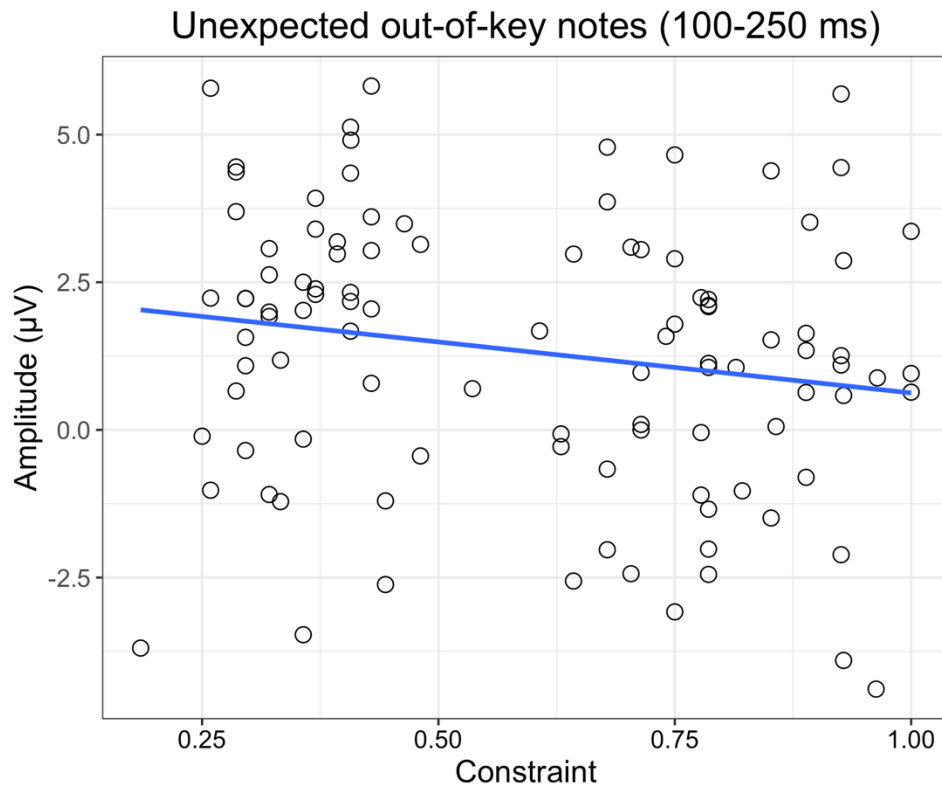


Figure 4.8. Amplitude across frontocentral electrodes from 100-250 ms vs. contextual constraint for items in the HC-unexpected and LC-unexpected conditions. Constraint has a significant effect on amplitude of out-of-key notes, such that the ERPs elicited by out-of-key notes are more negative in higher constraint contexts.

As seen in **Table 4.3**, constraint has a significant effect on amplitude of out-of-key notes, such that the ERPs elicited by out-of-key notes are more negative in higher constraint contexts than in lower constraint contexts (**Figure 4.8**). Cloze probability does not significantly modulate amplitude.

We ran two further analyses to test if the effect of constraint could be driven by the categorical difference between HC and LC contexts. In a regression using cloze probability and categorical condition (treatment coded with HC as reference level) as predictors, we find a significant effect of condition on

amplitude with LC contexts eliciting less negative amplitudes than HC contexts (**Table 4.4**). To directly compare continuous constraint and categorical condition as predictors, we ran a regression including both constraint and condition, as well as cloze. As shown in **Table 4.5**, condition was marginally significant and constraint did not reach significance in this analysis. Here, condition seems to be a slightly better predictor of amplitude than condition, but there is not enough unique variance predicted by either one to determine which is truly driving the effect.⁴

Table 4.4. amplitude ~ condition + cloze

	Estimate	Std.Error	t value	P(> t)
(Intercept)	0.823	0.359	2.295	0.026
condition: LC	0.973	0.328	2.962	0.004
cloze	0.057	0.175	0.326	0.745

Table 4.5. amplitude ~ condition + constraint + cloze

	Estimate	Std.Error	z value	P(> z)
(Intercept)	0.531	0.445	1.194	0.232
constraint	0.386	0.495	0.779	0.436
condition: LC	1.567	0.947	1.654	0.098
cloze	0.106	0.155	0.684	0.494

Experiment Comparison

In order to compare the spatial distributions of the difference between expected and unexpected notes in Experiments 1 and 2, we analyzed data from

⁴ We also conducted versions of all analyses including the interaction of cloze with constraint and/or condition. In no case did the interaction reach significance; thus, we report here models fit without the interaction term.

midline electrodes (FPz, Fz, Cz, CPz, Pz, Oz) in both experiments. We conducted a repeated measures ANOVA using within-subjects factors of Condition (expected vs. unexpected) and Electrode and a between-subjects factor of Violation Type (in-key vs. out-of-key), revealing a significant interaction between Condition, Electrode, and Violation Type, $F(5,305) = 13.55$, $p < 0.001$ (

Figure 4.9).

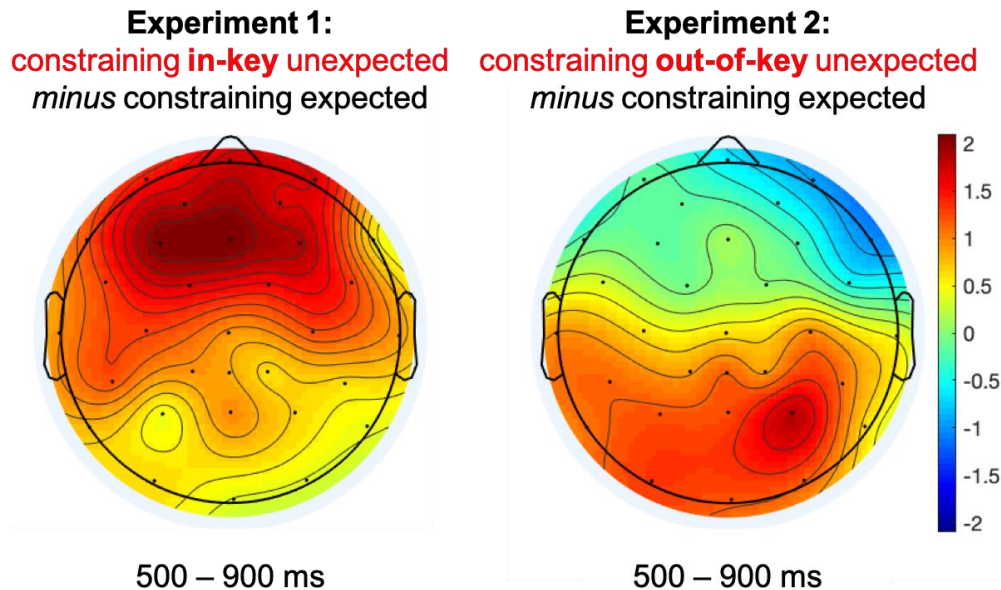


Figure 4.9. Topographic distributions of the difference between expected and unexpected notes in constraining contexts between 500 and 900 ms in Experiments 1 and 2.

Discussion

Our comparison of expected (tonic) vs. unexpected (out-of-key) notes in constraining contexts was similar to the manipulations of previous studies using melodic stimuli. However, in contrast to most prior reports, this contrast did *not*

elicit an early anterior negativity in the present study. Instead, the comparison elicited a posterior positivity around 500-700 ms (P600). Prior melodic expectation studies have sometimes reported a P600, but this finding has been inconsistent (Brattico et al., 2006; Halpern et al., 2017; Lagrois et al., 2018; Miranda & Ullman, 2007). It is notable that the presence of the P600 has sometimes been found to vary based on task instructions: Brattico et al. (2006) found that incongruous tones elicited a P600 when participants performed an incongruity detection task requiring that they actively pay attention to melodies, but not when they were instructed to ignore them, and Lagrois et al. (2018) also observed a P600 only when participants were given the task of detecting melodic incongruities. Here, our participants had to attend to melodies in order to perform the memory recognition task, though the task was designed to not call attention to the unexpected notes in particular. Given the widely varying tasks used in past melodic expectation experiments, differing task instructions may be part of the reason for inconsistencies between findings.

When comparing the effects of unexpected out-of-key notes in constraining vs. non-constraining contexts, we observed an early anterior negativity from 100-250 ms. Unexpected target notes were out-of-key in both of these conditions, and they were matched on cloze probability; the only difference between them was their prior melodic context. When comparing mispredicted to unpredicted tones, Hsu et al. (2015) also found that mispredicted tones were more negative in this time window; this resembles the same effect seen here. Unexpected out-of-key notes in constraining contexts also elicited a larger P600

compared to out-of-key notes in non-constraining contexts, a finding that has not previously been demonstrated in music.

Chapter 5. Discussion, Applications, and Future Directions

ERP Discussion

In Chapter 3 we used a novel manipulation, presenting participants with notes that were unexpected (but in-key and melodically congruous) in constraining or non-constraining contexts. Unexpected in-key notes in constraining contexts elicited a sustained frontal positivity compared to both expected notes in constraining contexts and unexpected notes in non-constraining contexts. When comparing ERPs to unexpected notes in constraining vs. non-constraining context, this effect began around 300 ms and was significant in a time window where a late frontal positivity has been observed in language (600-900 ms). When comparing ERPs to unexpected vs. expected notes in constraining contexts, this effect began early, emerging around 150 ms after stimulus presentation, and also remained significant through the time window where a late frontal positivity has been observed in language. In contrast to the results of previous studies (e.g., Fiveash, Thompson, et al., 2018; Kalda & Minati, 2012; Koelsch & Jentschke, 2010; Miranda & Ullman, 2007), unexpected notes did not elicit an early right anterior negativity (ERAN). Instead, unexpected notes elicited a bilateral frontal positivity, replicating the results seen by Pearce et al. (2010).

In Chapter 4 we repeated nearly the same paradigm, but used unexpected target notes that were *out-of-key*. Though this manipulation was more similar to most previous studies, we again did not observe an early anterior negativity when unexpected out-of-key notes were compared to expected notes in constraining contexts. However, this contrast did elicit a late posterior positivity (P600), which

has previously been reported to be elicited by melodic expectation violations (e.g., Halpern et al., 2017; Miranda & Ullman, 2007). Additionally, the P600 elicited by out-of-key notes was larger in constraining than in non-constraining contexts, despite the fact that these conditions did not differ significantly in cloze probability. In this experiment, we also observed an early anterior negativity from 100-250 ms when comparing the effects of unexpected out-of-key notes in constraining vs. non-constraining contexts. This effect resembles Hsu et al. (2015)'s finding of N1 enhancement for mispredicted stimuli, which they proposed to represent prediction error as sensory input fails to match with predictions. While we failed to replicate their effect for predicted vs unpredicted stimuli, they have since reported that that particular contrast appears to be particularly sensitive to manipulations of attention in a later study using a similar paradigm (Hsu, Hämäläinen, & Waszak, 2018).

In this set of experiments, we observed an early anterior negativity only when target notes that were both unexpected in constraining contexts *and* out-of-key were compared to notes that were unexpected and out-of-key in non-constraining contexts. While this early negativity is in the same temporal window as the effect referred to as the “ERAN” (and slightly right lateralized), it is unlikely to be the same effect. There is no sign of any early anterior negativity in this time window in the comparison between expected and unexpected notes in constraining contexts (**Figure 4.4**), which is the contrast that would typically be thought to elicit an ERAN. The melodic expectation violations used here may simply not have been incongruous enough to elicit an early anterior negativity,

even when they were out-of-key. The out-of-key unexpected target notes in Experiment 2 were more noticeable than the in-key unexpected notes in Experiment 1, but they may still have been perceived more as “unexpected” than as “wrong” notes. Additionally, participants were asked to perform a memory recognition task that was compatible with the normal experience of listening to music; it required them to attend to the melodies, but did not call any particular attention to prediction violations. Previous melody studies have used numerous different tasks, sometimes including overt acceptability judgements (e.g., Brattico et al., 2006; Halpern et al., 2017). Though there has not been any particularly consistent pattern of results among different studies using similar tasks, it has repeatedly been demonstrated that having participants perform different tasks can change the ERP effects elicited by melodic expectancy violations (Brattico et al., 2006; Lagrois et al., 2018; Zendel et al., 2015).

The P600 effect observed in Experiment 2 has a clear parallel in other domains, as it is known to be elicited by language, music, and visual sequence processing (Brattico et al., 2006; Cohn, Jackendoff, Holcomb, & Kuperberg, 2014; Cohn & Kutas, 2015; Demorest & Osterhout, 2012; Miranda & Ullman, 2007; Patel, Gibson, Ratner, Besson, & Holcomb, 1998; Vuvan et al., 2018). In language research, the P600 has classically been associated with syntax processing difficulty; it is elicited by syntactic errors as well as syntactically complex sentences (Osterhout & Holcomb, 1992; Van Petten & Luka, 2012b). More recently, it has also been observed in response to words that violate strong semantic constraints, which are anomalous or impossible in context and conflict

with the comprehender's current mental model (Kuperberg, 2007; Kuperberg, Sitnikova, Caplan, & Holcomb, 2003). Here, a P600 was elicited by the presentation of a syntactically unexpected out-of-key note. We also found that the P600 elicited by out-of-key notes was larger when they appear in constraining than in non-constraining contexts, an effect that has previously only been demonstrated in studies using language stimuli (Gunter et al., 2000; Kuperberg et al., 2018).

It is notable that the frontal positivity elicited by in-key unexpected notes in constraining melodic contexts in Experiment 1 appears to resemble the frontal positivity elicited by plausible yet unexpected words in constraining sentence contexts in language experiments (Federmeier et al., 2007; Kuperberg et al., 2018; Van Petten & Luka, 2012b; Wlotko, Federmeier, & Kutas, 2012). In language processing, predictions are thought to be made for specific lexical representations, which are clearly not a factor in music. However, there may be broader similarities in terms of anticipating that a certain representation will occur. While any number of differences exist between music and language, it is possible that some similar mechanisms of prediction operate in both domains. As such, in both language and music, frontal positivities may serve as an index of prediction error after encountering a plausible yet unexpected item, while posterior positivities may be elicited by items that are incompatible with the syntactic or semantic constraints of the prior context.

Current Applications

The melodic constraint approach is valuable for studying numerous aspects of melodic expectancy, some of which have been discussed thus far. Recently, this approach has additionally been used to evaluate the accuracy of the predictions of different computational models of melodic expectation (Morgan, Fogel, Nair, & Patel, 2019). There is a long history of using quantitative modeling as a method of investigating melodic expectancy. It has been used to investigate what factors contribute to forming predictions; for example, to quantify to what extent expectations are formed based on local contexts vs. larger scale structures (Krumhansl, Louhivuori, Toivainen, & Eerola, 1999; Margulis, 2005; Pearce et al., 2010; Schellenberg, 1996).

One category of models propose that expectations are governed by a number of simple rule-like principles, analogous to Gestalt principles in vision. Narmour's (1990) implication-realization (IR) model is probably the best-known example of a model of this type. More recently, Temperley (2008) developed the Probabilistic Model of Melody Perception, which uses three Gestalt-like principles to calculate the probability of a note given the context. These are the *central pitch tendency* ("melodies tend to remain within a narrow pitch range"), the *pitch proximity principle* ("note-to-note intervals within a melody tend to be small"), and the *key profile* ("notes tend to conform to a distribution (or key profile) that depends on the key"). While this model does utilize information about the key of a melody, it does not contain any factors directly based on harmonic or hierarchical structure. Temperley (2008) evaluated the ability of his

model to predict participants' responses using a dataset from Cuddy and Lunney (1995), where participants were played a context of two notes and were asked to rate how well a third note served as a continuation. The model performed relatively well on this rating data, providing some evidence that listeners' expectations are influenced by Gestalt-like principles. However, this dataset was far from a naturalistic representation of melodic expectation; contexts were only two notes long, and participants heard each context multiple times with several possible continuation notes.

A second category of models propose that listeners form expectations based on statistical learning of melodic structure. One example is Pearce's (2005) Information Dynamics Of Music (IDyOM) model, which posits that expectations are learned from statistical tendencies in both a listener's previous musical experience and within the current melody. In this model framework, probabilities are computed for every possible continuation note over multiple musical features. IDyOM incorporates numerous melodic parameters, but it does not include any factors explicitly based on harmonic structure. This model has also been validated using human data, successfully predicting behavioral ratings of continuation fit (Pearce & Wiggins, 2006) and uncertainty (Hansen & Pearce, 2014) as well as neural responses (Pearce et al., 2010). However, the melodies used in these studies did not include measurements or manipulations of predictive constraint, which largely relies on longer-scale harmonic expectations.

We used multinomial discrete-choice logit modeling (Agresti, 2002) to evaluate the ability of these two models to predict participants' sung responses on

the longer and more complex melodic stems used in the melodic cloze probability task, which varied in constraint (Morgan et al., 2019). The models were evaluated on melodic cloze data from (Fogel et al., 2015) and from the follow-up experiment described in Chapter 2. Both models successfully predicted some proportion of the human data, and each contributed significantly above and beyond the other. As such, it appears that both Gestalt-like principles *and* statistical learning contribute separately to the formation of melodic expectations. However, both models notably underpredicted the tonic in highly constraining melodies (**Figure 5.1**). These melodies were created using a manipulation of underlying harmony, ending with an implied authentic cadence. This suggests that implicit harmonic structure, which is not incorporated by either model, is a crucial part of how listeners form melodic expectations.

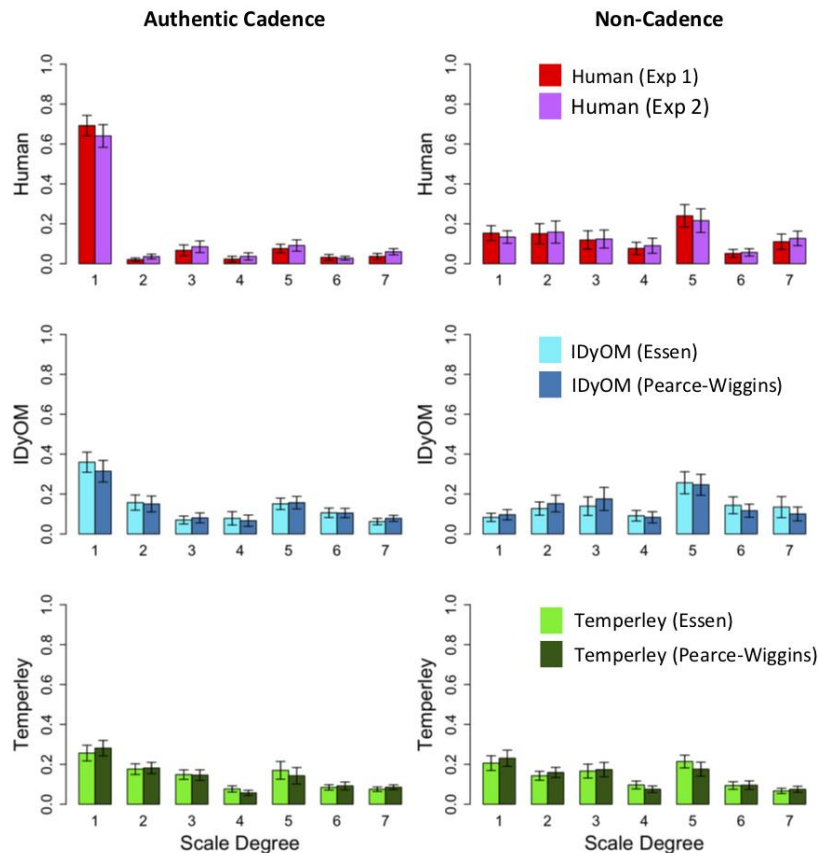


Figure 5.1. Figure from Morgan et al. (2019). Human data from the melodic cloze experiments (top row) and raw predictions from both models. Error bars show ± 2 standard errors (computed over melodies).

The melodic constraint approach is also being used to conduct research on cognitive processing in children with autism spectrum disorders (ASD). It has been demonstrated that children with ASD show a local processing bias in some tasks, focusing on local details over global coherence (the central coherence hypothesis)(Au-Yeung, Kaakinen, Liversedge, & Benson, 2018; Booth & Happé, 2010; Happé & Frith, 2006). Using a sentence completion task like the cloze probability task in language (Arcuri, Rabe-Hesketh, Morris, & McGuire, 2001), Booth and Happé (2010) found that children with ASD were more likely than typically developing (TD) children to provide completions that were based on the

immediate context of the end of the sentence instead of its global meaning. For example, participants might be asked to complete the sentence “*You can go hunting with a knife and*” Participants responding based on the global meaning of the sentence might respond with “*catch a bear,*” while those attending to its local details (i.e., the word “*knife*”) might respond with “*fork.*” The melodic cloze probability task will be used to test if this same pattern of cognitive processing styles is observable in other domains. According to the theory of central coherence, children with ASD will be more likely than TD children to provide responses that are based on the local context of the end of a melody instead of its larger-scale harmonic structure. As such, they may be less affected by the constraint manipulation used in the melodic cloze probability task (which relies on underlying harmony) and more likely to instead provide responses that adhere to basic melodic principles like pitch proximity in both conditions.

Future Directions

The ERP analyses presented in Chapters 3 and 4 investigated prediction by looking at the neural effects of *violating* predictions after a target note is presented. However, this existing ERP data could also be reanalyzed to examine brain activity *prior* to the target note in different constraint conditions in order to more directly observe the neural correlates of the buildup of expectations during melodic processing. Neural oscillations, which are increasingly being used to study language, can provide insight into this activity (Hauk, Giraud, & Clarke, 2017; Lewis, Wang, & Bastiaansen, 2015). In language, alpha oscillations are decreased prior to the target word in highly constraining sentence contexts

compared to low constraint contexts; this may represent attentional preparation for the upcoming item or engagement of the language network during prediction (Rommers, Dickson, Norton, Wlotko, & Federmeier, 2017; Wang, Hagoort, & Jensen, 2018). This effect may not be domain-specific, as it has been also observed in visual sequence processing (Bidet-Caulet et al., 2012). Therefore, if we were to compare neural activity prior to the target note in our melodic sequences, we would expect to also observe decreased alpha activity in highly constraining melodies.

Future studies using variations on the ERP paradigms in Chapters 3 and 4 could address questions yet to be answered by the present results. In the present set of studies, multiple conditions were altered between our first and second experiments: unexpected target notes in both high constraint *and* low constraint conditions changed from in-key to out-of-key. This change of context could potentially have confounded the comparison between experiments of the difference between expected and unexpected notes in highly constraining melodies. In Experiment 1, no melodies contained any out-of-key notes, while in Experiment 2 participants heard 40 melodies with out-of-key notes and 50 without in each block. The effects of unexpected out-of-key notes could be examined while minimizing changes to context by conducting an experiment using high constraint expected (Experiments 1&2), high constraint out-of-key (Experiment 2), and low constraint *in*-key (Experiment 1) conditions. In this structure, only 20 melodies would contain out-of-key notes while the remaining 70 would be entirely in-key. Unexpected out-of-key notes in high constraint

contexts would most likely elicit a P600 as they did in Experiment 2, but the effect may be stronger due to having a smaller proportion of out-of-key notes in the experiment overall.

Currently, the possibilities for future ERP studies of melodic processing are limited by our finite number of available stimuli: 60 matched pairs of AC and NC melodies. Putting aside the issue of obtaining more candidate melodies in the first place, the fact that melodic stems must be normed using the melodic cloze probability paradigm places a significant limitation on the development of additional stimuli. One way this could be addressed is by running the melodic cloze probability task over the internet, through a service such as Amazon's Mechanical Turk (this is already technologically feasible, but has not been implemented). If additional stimuli are obtained in this way, it may be possible to address some of the methodological differences between our ERP studies using melodies and the sentence studies they were modeled after. Though stimuli vary from experiment to experiment, the mean constraint of our low constraint contexts (36.9%) was somewhat higher than that of contexts typically classified as low constraint in language experiments, which have been around 20-30% (Federmeier et al., 2007; Thornhill & Van Petten, 2012; Wang et al., 2018; Wlotko & Federmeier, 2012). It is inherently more difficult to create sequences with low constraint in music than in language; there are only seven diatonic scale degrees (plausible responses) in an octave, so the lowest-constraint sequence possible with an even distribution of responses over all plausible notes would have a constraint of 14%. In the stimuli used here, the single lowest measurement of constraint

obtained for any one melodic stem was 18.5%, though the lowest value that was obtained for multiple stems was 25% (see **Figure 3.5**). Given enough candidate melodies to test using the cloze paradigm, it may be possible to construct an experiment using “low constraint” melodies that have a mean closer to this value. Additionally, with more stimuli, all five conditions used across Experiments 1 and 2 could be included in one experiment as has been done in language studies.

In addition to norming stimuli, running the melodic cloze probability task online could be a useful method of answering questions about melodic constraint itself. For example, we could investigate the contribution of rhythm to melodic constraint by using isochronous versions of the same melodies. As AC melodic stems were originally composed to end preceding a strong beat within the meter, they might be less constraining when presented without metric structure (Fogel et al., 2015). (Neither of the computational models discussed in the beginning of this chapter included a rhythmic parameter, so they might perform better at predicting participants’ responses in this form of the task.) Furthermore, we could test non-musicians’ performance in the melodic cloze probability task. Musicians were originally chosen as participants for the cloze task largely due to the necessity of having participants who can sing on pitch, but they might also form melodic predictions more readily or strongly than non-musicians due to their greater experience doing so. However, non-musicians still display knowledge of the underlying harmonic structures that guide the constraint manipulation, so their responses in this task would likely be similar to musicians’ (with more noise in the data). Cloze probabilities obtained using non-musicians or any other group

could then be used to construct stimuli for ERP experiments in that population as well.

Bibliography

- Agresti, A. (2002). *Categorical Data Analysis*. Wiley-Interscience.
- Arcuri, S. M., Rabe-Hesketh, S., Morris, R. G., & McGuire, P. K. (2001).
Regional variation of cloze probabilities for sentence contexts. *Behavior
Research Methods, Instruments, and Computers*, 33(1), 80–90.
doi:10.3758/BF03195350
- Au-Yeung, S. K., Kaakinen, J. K., Liversedge, S. P., & Benson, V. (2018). Would
adults with autism be less likely to bury the survivors?: An eye movement
study of anomalous text reading. *Quarterly Journal of Experimental
Psychology*, 71(1 Special Issue), 280–290.
doi:10.1080/17470218.2017.1322621
- Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects
structure for confirmatory hypothesis testing: Keep it maximal. *Journal of
Memory and Language*, 68(3), 255–278. doi:10.1016/j.jml.2012.11.001
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-
Effects Models Using {lme4}. *Journal of Statistical Software*, 67(1), 1–48.
doi:10.18637/jss.v067.i01
- Besson, M., & Faïta, F. (1995). An Event-Related Potential (ERP) Study of
Musical Expectancy: Comparison of Musicians With Nonmusicians. *Journal
of Experimental Psychology: Human Perception and Performance*, 21(6),
1278–1296. doi:10.1037/0096-1523.21.6.1278

- Bezanson, J., Edelman, A., Karpinski, S., & Shah, V. B. (2017). Julia: A fresh approach to numerical computing. *SIAM Review*, 59(1), 65–98.
- Bidet-Caulet, A., Barbe, P.-G., Roux, S., Viswanath, H., Barthélémy, C., Bruneau, N., ... Bonnet-Brilhault, F. (2012). Dynamics of anticipatory mechanisms during predictive context processing. *European Journal of Neuroscience*, 36(7), 2996–3004. doi:10.1111/j.1460-9568.2012.08223.x
- Boersma, P. (2001). PRAAT, a system for doing phonetics by computer. *Glot International*, 5(9/10), 341–345. doi:10.1097/AUD.0b013e31821473f7
- Booth, R., & Happé, F. (2010). “Hunting with a knife and ... fork”: Examining central coherence in autism, attention deficit/hyperactivity disorder, and typical development with a linguistic task. *Journal of Experimental Child Psychology*, 107(4), 377–393. doi:10.1016/j.jecp.2010.06.003
- Brattico, E., Tervaniemi, M., Näätänen, R., & Peretz, I. (2006). Musical scale properties are automatically processed in the human auditory cortex. *Brain Research*, 1117(1), 162–74. doi:10.1016/j.brainres.2006.08.023
- Carlsen, J. C. (1981). Some factors which influence melodic expectancy. *Psychomusicology: A Journal of Research in Music Cognition*, 1(1), 12–29. doi:10.1037/h0094276
- Choi, I., Bharadwaj, H. M., Bressler, S., Loui, P., Lee, K., & Shinn-Cunningham, B. G. (2014). Automatic processing of abstract musical tonality. *Frontiers in Human Neuroscience*, 8(December), 988. doi:10.3389/fnhum.2014.00988

- Cohn, N., Jackendoff, R., Holcomb, P. J., & Kuperberg, G. R. (2014). The grammar of visual narrative: Neural evidence for constituent structure in sequential image comprehension. *Neuropsychologia*, 64C, 63–70. doi:10.1016/j.neuropsychologia.2014.09.018
- Cohn, N., & Kutas, M. (2015). Getting a cue before getting a clue: Event-related potentials to inference in visual narrative comprehension. *Neuropsychologia*, 77, 267–278. doi:10.1016/j.neuropsychologia.2015.08.026
- 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. doi:10.1111/desc.12100
- Cuddy, L. L., & Lunney, C. (1995). Expectancies generated by melodic intervals: perceptual judgments of melodic continuity. *Perception & Psychophysics*, 57(4), 451–462. doi:10.3758/BF03213071
- Demorest, S. M., & Osterhout, L. (2012). ERP responses to cross-cultural melodic expectancy violations. *Annals Of The New York Academy Of Sciences*, 1252, 152–7. doi:10.1111/j.1749-6632.2012.06464.x
- Egermann, H., Pearce, M. T., Wiggins, G. A., & McAdams, S. (2013). Probabilistic models of expectation violation predict psychophysiological emotional responses to live concert music. *Cognitive, Affective and Behavioral Neuroscience*, 13(3), 533–553. doi:10.3758/s13415-013-0161-y
- 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. doi:10.1016/j.brainres.2006.06.101
- 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. doi:10.3758/MC.37.1.1
- Fields, E. C. (2017). Factorial Mass Univariate ERP Toolbox.
- Fields, E. C., & Kuperberg, G. R. (2019). Having your cake and eating it too : Flexibility and power with mass univariate statistics for ERP data. doi:10.31234/osf.io/qfkgc
- Fiveash, A., McArthur, G., & Thompson, W. F. (2018). Syntactic and non-syntactic sources of interference by music on language processing. *Scientific Reports*, 8(17918). doi:10.1038/s41598-018-36076-x
- Fiveash, A., Thompson, W. F., Badcock, N. A., & McArthur, G. (2018). Syntactic processing in music and language: Effects of interrupting auditory streams with alternating timbres. *International Journal of Psychophysiology*, 129, 31–40. doi:10.1016/j.ijpsycho.2018.05.003
- 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. doi:10.3389/fpsyg.2015.01718

- Fogel, A. R., Wlotko, E. W., Kuperberg, G. R., & Patel, A. D. (2017). Comparing linguistic prediction in musicians and non-musicians. In *24th Annual Meeting of the Cognitive Neuroscience Society*.
- Garnier, C., Poulin-Charronnat, B., Schellenberg, E. G., Stevens, C., & Bigand, E. (2005). Children's implicit knowledge of harmony in Western music. *Developmental Science*, 8(6), 551–566. doi:10.1111/j.1467-7687.2005.00447.x
- Granot, R., & Donchin, E. (2004). Do Re Mi Fa Sol La Ti—Constraints, Congruity, and Musical Training: An Event-Related Brain Potentials Study of Musical Expectancies. *Music Perception*, 19(4), 487–528. doi:10.1525/mp.2002.19.4.487
- Groppe, D. M., Urbach, T. P., & Kutas, M. (2011). Mass univariate analysis of event-related brain potentials/fields I: A critical tutorial review. *Psychophysiology*, 48(12), 1711–1725. doi:10.1111/j.1469-8986.2011.01273.x
- Gunter, T. C., Friederici, A. D., & Schriefers, H. (2000). Syntactic Gender and Semantic Expectancy: ERPs Reveal Early Autonomy and Late Interaction. *Journal of Cognitive Neuroscience*, 12(4), 556–568. doi:10.1162/089892900562336
- Guo, S., & Koelsch, S. (2016). Effects of veridical expectations on syntax processing in music: Event-related potential evidence. *Scientific Reports*,

6(November 2015), 1–11. doi:10.1038/srep19064

Habibi, A., Wirantana, V., & Starr, A. (2013). Cortical Activity During Perception of Musical Pitch. *Music Perception: An Interdisciplinary Journal*, 30(5), 463–479. doi:10.1525/mp.2013.30.5.463

Halpern, A. R., Martin, J. S., & Reed, T. D. (2008). An ERP Study of Major-Minor Classification in Melodies. *Music Perception: An Interdisciplinary Journal*, 25(3), 181–191. doi:10.1525/mp.2008.25.3.181

Halpern, A. R., Zioga, I., Shankleman, M., Lindsen, J., Pearce, M. T., & Bhattacharya, J. (2017). That note sounds wrong! Age-related effects in processing of musical expectation. *Brain and Cognition*, 113, 1–9. doi:10.1016/j.bandc.2016.12.006

Hansen, N. C., & Pearce, M. T. (2014). Predictive uncertainty in auditory sequence processing. *Frontiers in Psychology*, 5(SEP), 1–17. doi:10.3389/fpsyg.2014.01052

Happé, F., & Frith, U. (2006). The weak coherence account: Detail-focused cognitive style in autism spectrum disorders. *Journal of Autism and Developmental Disorders*, 36(1), 5–25. doi:10.1007/s10803-005-0039-0

Hauk, O., Giraud, A.-L., & Clarke, A. (2017). Brain oscillations in language comprehension. *Language, Cognition and Neuroscience*, 32(5), 533–535. doi:10.1080/23273798.2017.1297842

- Hsu, Y.-F., Hämäläinen, J. A., & Waszak, F. (2018). The processing of mispredicted and unpredicted sensory inputs interact differently with attention. *Neuropsychologia*, 111(August 2017), 85–91. doi:10.1016/j.neuropsychologia.2018.01.034
- Hsu, Y.-F., Le Bars, S., Hämäläinen, J. A., & Waszak, F. (2015). Distinctive Representation of Mispredicted and Unpredicted Prediction Errors in Human Electroencephalography. *Journal of Neuroscience*, 35(43), 14653–14660. doi: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. doi: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.
- Jentschke, S., Friederici, A. D., & Koelsch, S. (2014). Neural correlates of music-syntactic processing in two-year old children. *Developmental Cognitive Neuroscience*, 9, 200–208. doi:10.1016/j.dcn.2014.04.005
- Kalda, T., & Minati, L. (2012). Detecting scale violations in absence of mismatch requires music-syntactic analysis: A further look at the early right anterior negativity (ERAN). *Brain Topography*, 25(3), 285–292.

doi:10.1007/s10548-011-0208-8

- Kim, J. N., Large, E. W., Gwon, Y., & Ashley, R. (2018). The Online Processing of Implied Harmony in the Perception of Tonal Melodies. *Music Perception: An Interdisciplinary Journal*, 35(5), 594–606. doi:10.1525/mp.2018.35.5.594
- Koelsch, S. (2009). Music-syntactic processing and auditory memory: similarities and differences between ERAN and MMN. *Psychophysiology*, 46(1), 179–190. doi:10.1111/j.1469-8986.2008.00752.x
- Koelsch, S. (2012). *Brain and Music*. *Brain and Music*. John Wiley & Sons.
- Koelsch, S. (2014). Brain correlates of music-evoked emotions. *Nature Reviews Neuroscience*, 15(3), 170–180. doi:10.1038/nrn3666
- Koelsch, S., Gunter, T. C., Friederici, A. D., & Schröger, E. (2000). Brain Indices of Music Processing: “Nonmusicians’” are Musical.” *Journal of Cognitive Neuroscience*, 12(3), 520–541.
- Koelsch, S., & Jentschke, S. (2010). Differences in electric brain responses to melodies and chords. *Journal of Cognitive Neuroscience*, 22(10), 2251–62. doi:10.1162/jocn.2009.21338
- Koelsch, S., Jentschke, S., Sammler, D., & Mietchen, D. (2007). Untangling syntactic and sensory processing: an ERP study of music perception. *Psychophysiology*, 44(3), 476–90. doi:10.1111/j.1469-8986.2007.00517.x
- 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.
doi:10.1073/pnas.1300272110

Koelsch, S., Vuust, P., & Friston, K. (2019). Predictive Processes and the Peculiar Case of Music. *Trends in Cognitive Sciences*, 23(1), 63–77.
doi:10.1016/j.tics.2018.10.006

Krumhansl, C. L. (1979). The psychological representation of musical pitch in a tonal context. *Cognitive Psychology*, 11(3), 346–374. doi:10.1016/0010-0285(79)90016-1

Krumhansl, C. L., & Cuddy, L. L. (2010). A Theory of Tonal Hierarchies in Music. In M. Riess Jones, R. R. Fay, & A. N. Popper (Eds.), *Music Perception* (Vol. 36, pp. 51–87). New York, NY: Springer New York.
doi:10.1007/978-1-4419-6114-3

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. doi:10.2307/40285890

Krumhansl, C. L., & Shepard, R. N. (1979). Quantification of the hierarchy of tonal functions within a diatonic context. *Journal of Experimental Psychology. Human Perception and Performance*, 5(4), 579–594.
doi:10.1037/0096-1523.5.4.579

- Kunert, R., Willems, R. M., Casasanto, D., Patel, A. D., & Hagoort, P. (2015). Music and language syntax interact in Broca's area: An fMRI study. *PLoS ONE*, 10(11), 1–16. doi:10.1371/journal.pone.0141069
- Kuperberg, G. R. (2007). Neural mechanisms of language comprehension: challenges to syntax. *Brain Research*, 1146, 23–49. doi: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., Brothers, T., & Wlotko, E. W. (2018). A Tale of Two Positivities (and the N400): Distinct neural signatures are evoked by confirmed and violated predictions at different levels of representation. *BioRxiv*, 404780. doi:10.1101/404780
- Kuperberg, G. R., & Jaeger, T. F. (2016). What do we mean by prediction in language comprehension? *Language Cognition & Neuroscience*, 31(1), 32–59. doi:10.1080/23273798.2015.1102299
- Kuperberg, G. R., Sitnikova, T., Caplan, D., & Holcomb, P. J. (2003). Electrophysiological distinctions in processing conceptual relationships within simple sentences. *Cognitive Brain Research*, 17(1), 117–129.

doi:10.1016/S0926-6410(03)00086-7

Kuriki, S., Isahai, N., & Ohtsuka, A. (2005). Spatiotemporal characteristics of the neural activities processing consonant/dissonant tones in melody.

Experimental Brain Research, 162(1), 46–55. doi:10.1007/s00221-004-2114-8

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.

doi:10.1146/annurev.psych.093008.131123

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

doi:10.1126/science.7350657

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

Lagrois, M.-E., Peretz, I., & Zendel, B. R. (2018). Neurophysiological and Behavioral Differences between Older and Younger Adults When Processing

Violations of Tonal Structure in Music. *Frontiers in Neuroscience*,

12(February). doi:10.3389/fnins.2018.00054

Lewis, A. G., Wang, L., & Bastiaansen, M. (2015). Fast oscillatory dynamics during language comprehension: Unification versus maintenance and

prediction? *Brain and Language*, 148, 1–13.

doi:10.1016/j.bandl.2015.01.003

Li, C. (2018). JuliaCall: Seamless Integration Between R and “Julia”. R package version 0.12.2.

Lindsen, J. P., Pearce, M. T., Doyne, M., Wiggins, G. A., & Bhattacharya, J. (2010). Implicit Brain Responses During Fulfillment of Melodic Expectations. In *Proceedings of the 12th International Conference on Music Perception and Cognition* (pp. 606–607).

Loui, P., Wessel, D., & Kam, C. (2010). Humans Rapidly Learn Grammatical Structure in a New Musical Scale. *Music Perception*, 377–388.

Margulis, E. H. (2005). A model of melodic expectation. *Music Perception*, 22(4), 663–714. doi: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.
doi:10.1080/09298210600835042

Marmel, F., Perrin, F., & Tillmann, B. (2011). Tonal expectations influence early pitch processing. *Journal of Cognitive Neuroscience*, 23(10), 3095–104.
doi:10.1162/jocn.2011.21632

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

Miranda, R. A., & Ullman, M. T. (2007). Double dissociation between rules and

memory in music: an event-related potential study. *NeuroImage*, 38(2), 331–45. doi:10.1016/j.neuroimage.2007.07.034

Morgan, E., Fogel, A., Nair, A., & Patel, A. D. (2019). Statistical learning and Gestalt-like principles predict melodic expectations. *Cognition*, 189(December 2018), 23–34. doi:10.1016/j.cognition.2018.12.015

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

Omigie, D., Pearce, M. T., Williamson, V. J., & Stewart, L. (2013). Electrophysiological correlates of melodic processing in congenital amusia. *Neuropsychologia*, 51(9), 1749–62. doi:10.1016/j.neuropsychologia.2013.05.010

Osterhout, L., & Holcomb, P. J. (1992). Event-related brain potentials elicited by syntactic anomaly. *Journal of Memory and Language*, 31(6), 785–806. doi:10.1016/0749-596X(92)90039-Z

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.

Patel, A. D. (2013). Sharing and Nonsharing of Brain Resources for Language

and Music. In M. A. Arbib (Ed.), *Language, music, and the brain: A mysterious relationship*. MIT Press.

Patel, A. D., Gibson, E., Ratner, J., Besson, M., & Holcomb, P. J. (1998).

Processing syntactic relations in language and music: An event-related potential study. *Journal of Cognitive Neuroscience*, 10(6), 717–733.

Patel, A. D., & Morgan, E. (2016). Exploring Cognitive Relations Between Prediction in Language and Music. *Cognitive Science*, 1–18.

doi:10.1111/cogs.12411

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

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.

doi: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.

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

4(4), 625–52. doi:10.1111/j.1756-8765.2012.01214.x

- Peretz, I., Brattico, E., Järvenpää, M., & Tervaniemi, M. (2009). The amusic brain: In tune, out of key, and unaware. *Brain*, 132(5), 1277–1286.
doi:10.1093/brain/awp055
- R Core Team. (2018). R: A Language and Environment for Statistical Computing. Vienna, Austria. Retrieved from <https://www.r-project.org/>
- Rohrmeier, M., & Koelsch, S. (2012). Predictive information processing in music cognition. A critical review. *International Journal of Psychophysiology*, 83(2), 164–75. doi:10.1016/j.ijpsycho.2011.12.010
- Rommers, J., Dickson, D. S., Norton, J. J. S. S., Wlotko, E. W., & Federmeier, K. D. (2017). Alpha and theta band dynamics related to sentential constraint and word expectancy. *Language, Cognition and Neuroscience*, 32(5), 576–589.
doi:10.1080/23273798.2016.1183799
- Ross, S., & Hansen, N. C. (2016). Dissociating Prediction Failure: Considerations from Music Perception. *Journal of Neuroscience*, 36(11), 3103–3105.
doi:10.1523/JNEUROSCI.0053-16.2016
- Salimpoor, V. N., Zald, D. H., Zatorre, R. J., Dagher, A., & McIntosh, A. R. (2014). Predictions and the brain: how musical sounds become rewarding. *Trends in Cognitive Sciences*, 19(2), 86–91. doi:10.1016/j.tics.2014.12.001
- Schellenberg, E. G. (1996). Expectancy in melody: tests of the implication-realization model. *Cognition*, 58(1), 75–125. doi:10.1016/0010-0277(95)00665-6

- 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. doi:10.1037/0096-3445.131.4.511
- Schmuckler, M. A. (1989). Expectation in Music: Investigation of Melodic and Harmonic Processes. *Music Perception*, 7(2), 109–150.
- Schön, D., & Besson, M. (2005). Visually Induced Auditory Expectancy in Music Reading: A Behavioral and Electrophysiological Study. *Journal of Cognitive Neuroscience*, 17(4), 694–705. doi:10.1162/0898929053467532
- 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. doi: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. doi:10.1162/jocn.2006.18.8.1380
- Sun, Y., Lu, X., Ho, H. T., Johnson, B. W., Sammler, D., & Thompson, W. F. (2018). Syntactic processing in music and language: Parallel abnormalities observed in congenital amusia. *NeuroImage: Clinical*, 19(December 2017), 640–651. doi:10.1016/j.nicl.2018.05.032
- Temperley, D. (2008). A probabilistic model of melody perception. *Cognitive*

Science, 32(2), 418–44. doi:10.1080/03640210701864089

Thornhill, D. E., & Van Petten, C. (2012). Lexical versus conceptual anticipation during sentence processing: frontal positivity and N400 ERP components.

International Journal of Psychophysiology : Official Journal of the International Organization of Psychophysiology, 83(3), 382–92.

doi:10.1016/j.ijpsycho.2011.12.007

Tillmann, B., Bharucha, J. J., & Bigand, E. (2000). Implicit learning of tonality: A self-organizing approach. *Psychological Review*, 107(4), 885–913.

doi:10.1037//0033-295X.107.4.885

Tillmann, B., Poulin-Charronnat, B., & Bigand, E. (2014). The role of expectation in music: From the score to emotions and the brain. *Wiley Interdisciplinary Reviews: Cognitive Science*, 5(1), 105–113. doi:10.1002/wcs.1262

Trainor, L. J., & Trehub, S. E. (1994). Key membership and implied harmony in Western tonal music: developmental perspectives. *Perception & Psychophysics*, 56(2), 125–132. doi:10.3758/BF03213891

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

- Vuvan, D., Zendel, B. R., & Peretz, I. (2018). Random Feedback Makes Listeners Tone-Deaf. *Scientific Reports*, 8(1), 7283. doi:10.1038/s41598-018-25518-1
- Wang, L., Hagoort, P., & Jensen, O. (2018). Language Prediction Is Reflected by Coupling between Frontal Gamma and Posterior Alpha Oscillations. *Journal of Cognitive Neuroscience*, 30(3), 432–447. doi:10.1162/jocn_a_01190
- Wlotko, E. W., & Federmeier, K. D. (2012). So that's what you meant! Event-related potentials reveal multiple aspects of context use during construction of message-level meaning. *NeuroImage*, 62(1), 356–66.
doi:10.1016/j.neuroimage.2012.04.054
- 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. doi:10.1037/a0029206
- Zendel, B. R., Lagrois, M.-E., Robitaille, N., & Peretz, I. (2015). Attending to Pitch Information Inhibits Processing of Pitch Information: The Curious Case of Amusia. *Journal of Neuroscience*, 35(9), 3815–3824.
doi:10.1523/jneurosci.3766-14.2015

Appendix A. MMN studies of melodic expectation

Fujioka, T., Trainor, L. J., Ross, B., Kakigi, R., & Pantev, C. (2005). Automatic Encoding of Polyphonic Melodies in Musicians and Nonmusicians. *Journal of Cognitive Neuroscience*, 17(10), 1578–1592.

doi:10.1162/089892905774597263

Fujioka, T., Trainor, L. J., Ross, B., Kakigi, R., & Pantev, C. (2004). Musical training enhances automatic encoding of melodic contour and interval structure. *Journal of Cognitive Neuroscience*, 16(6), 1010–1021.

doi:10.1162/0898929041502706

Habibi, A., Wirantana, V., & Starr, A. (2013). Cortical Activity During Perception of Musical Pitch. *Music Perception: An Interdisciplinary Journal*, 30(5), 463–479. doi:10.1525/mp.2013.30.5.463

Hsu, Y.-F., Hämäläinen, J. A., & Waszak, F. (2018). The processing of mispredicted and unpredicted sensory inputs interact differently with attention. *Neuropsychologia*, 111(August 2017), 85–91.

doi:10.1016/j.neuropsychologia.2018.01.034

Hsu, Y.-F., Le Bars, S., Hämäläinen, J. A., & Waszak, F. (2015). Distinctive Representation of Mispredicted and Unpredicted Prediction Errors in Human Electroencephalography. *Journal of Neuroscience*, 35(43), 14653–14660.

doi:10.1523/JNEUROSCI.2204-15.2015

- Magne, C., Schön, D., & Besson, M. (2006). Musician children detect pitch violations in both music and language better than nonmusician children: Behavioral and electrophysiological approaches. *Journal of Cognitive Neuroscience*, 18(2), 199–211. doi:10.1162/jocn.2006.18.2.199
- Maheu, M., Dehaene, S., & Meyniel, F. (2019). Brain signatures of a multiscale process of sequence learning in humans. *ELife*, 8, 1–24. doi:10.7554/elife.41541
- Mathias, B., Palmer, C., Perrin, F., & Tillmann, B. (2015). Sensorimotor Learning Enhances Expectations during Auditory Perception. *Cerebral Cortex*, 25(8), 2238–2254. doi:10.1093/cercor/bhu030
- Mathias, B., Tillmann, B., & Palmer, C. (2016). Sensory, Cognitive, and Sensorimotor Learning Effects in Recognition Memory for Music. *Journal of Cognitive Neuroscience*, 28(8), 1111–1126. doi:10.1162/jocn_a_00958
- Nan, Y., Knösche, T. R., & Friederici, A. D. (2009). Non-musicians' perception of phrase boundaries in music: A cross-cultural ERP study. *Biological Psychology*, 82(1), 70–81. doi:10.1016/j.biopsycho.2009.06.002
- Neuhaus, C., Knösche, T. R., & Friederici, A. D. (2006). Effects of musical expertise and boundary markers on phrase perception in music. *Journal of Cognitive Neuroscience*, 18(3), 472–493. doi:10.1162/jocn.2006.18.3.472

- Putkinen, V., Tervaniemi, M., Saarikivi, K., de Vent, N., & Huottilainen, M. (2014). Investigating the effects of musical training on functional brain development with a novel Melodic MMN paradigm. *Neurobiology of Learning and Memory*, 110, 8–15. doi:10.1016/j.nlm.2014.01.007
- Schön, D., & Besson, M. (2005). Visually Induced Auditory Expectancy in Music Reading: A Behavioral and Electrophysiological Study. *Journal of Cognitive Neuroscience*, 17(4), 694–705. doi:10.1162/0898929053467532
- Schön, D., Magne, C., & Besson, M. (2004). The music of speech: Music training facilitates pitch processing in both music and language. *Psychophysiology*, 41(3), 341–349. doi:10.1111/1469-8986.00172.x
- Schotanus, Y. P., Eekhof, L. S., & Willems, R. M. (2018). Behavioral and Neurophysiological Effects of Singing and Accompaniment on the Perception and Cognition of Song. In R. Parnicutt & S. Sattmann (Eds.), *Proceedings of ICMPC15/ESCOM10* (pp. 389–394). Graz, Austria: Centre for Systematic Musicology, University of Graz.
- Tervaniemi, M., Huotllainen, M., Brattico, E., Ilmoniemi, R. J., Reinlkainen, K., & Alho, K. (2003). Event-Related Potentials to Expectancy Violation in Musical Context. *Musicae Scientiae*, 7(2), 241–261. doi:10.1177/102986490300700203

- Tervaniemi, M., Rytkönen, M., Schröger, E., Ilmoniemi, R. J., & Näätänen, R. (2001). Superior formation of cortical memory traces for melodic patterns in musicians. *Learning and Memory*, 8(5), 295–300. doi:10.1101/lm.39501
- Tew, S., Fujioka, T., He, C., & Trainor, L. (2009). Neural representation of transposed melody in infants at 6 months of age. *Annals of the New York Academy of Sciences*, 1169, 287–290. doi:10.1111/j.1749-6632.2009.04845.x
- Trainor, L. J., McDonald, K. L., & Alain, C. (2002). Automatic and Controlled Processing of Melodic Contour Brain Activity. *Journal of Cognitive Neuroscience*, 14, 430–442. doi:10.1162/089892902317361949
- Trainor, L. J., Desjardins, R. N., & Rockel, C. (1999). A comparison of contour and interval processing in musicians and nonmusicians using event-related potentials. *Australian Journal of Psychology*, 51(3), 147–153. doi:10.1080/00049539908255352
- Trainor, L. J., McDonald, K. L., & Alain, C. (2002). Automatic and controlled processing of melodic contour and interval information measured by electrical brain activity. *Journal of Cognitive Neuroscience*, 14(3), 430–442. doi:10.1162/089892902317361949

Appendix B. Notation for melodic stems used in Chapter 2.

For each pair of melodies, the melodic stems used in the cloze task were comprised of the notes from melody onset until the note just prior to the asterisked note. The upper member of each pair is the authentic cadence (AC) stem; the lower member is the non-cadence (NC) stem.

Melodies with high or low constraint

Jason Carl Rosenberg

1

2

3

4

5

6

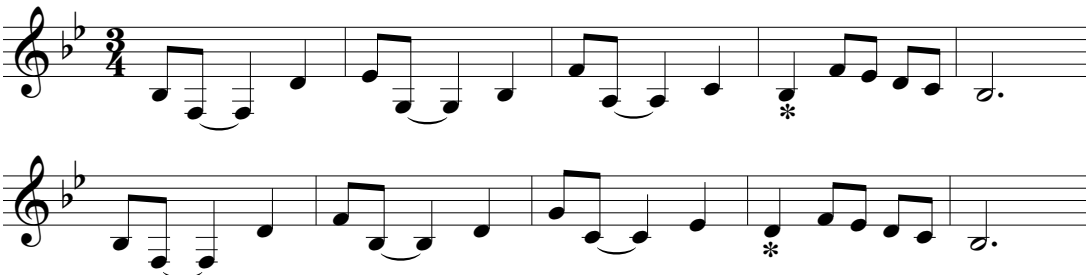
7

8

9

10

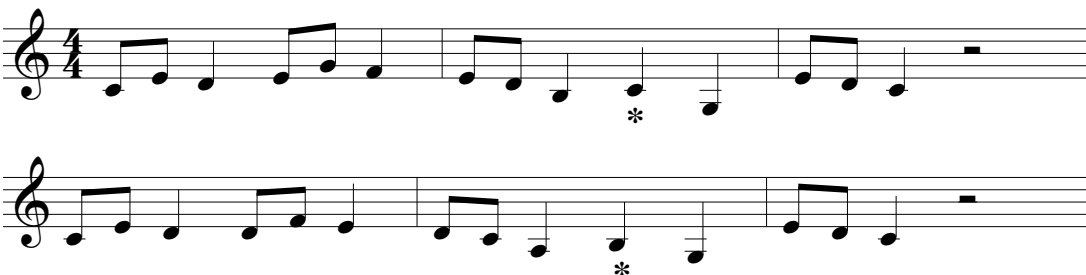
11



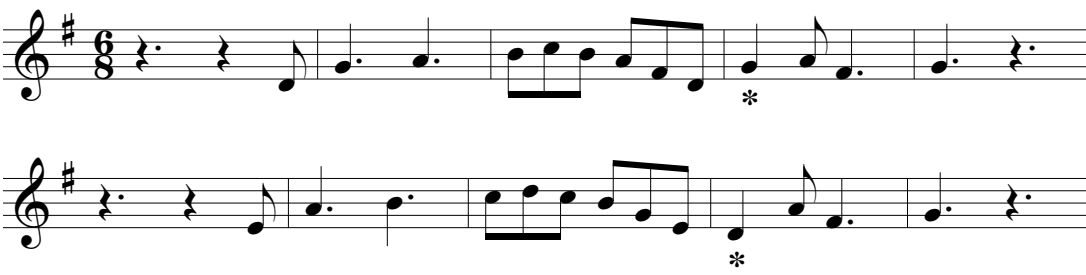
12



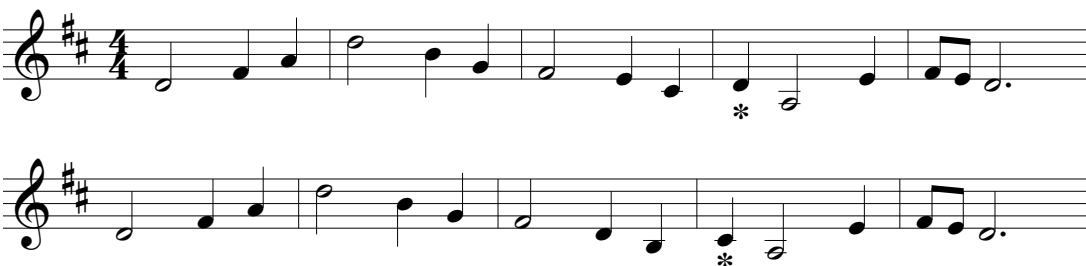
13



14



15



16

17

18

19

20

21

22

23

24

25

26

27

28

29

30

31

32

33

34

35

36

37

38

41

42

43

44

45

46

47

48